

Analysis of iTRAQ data using Mascot and Peaks quantification algorithms

Carla M.R. Lacerda, Lei Xin, Iain Rogers and Kenneth F. Reardon

Abstract

The field of proteomics has been developing rapidly toward quantification of proteins. Despite the variety of experimental techniques available for peptide and protein labelling, there are few commercially available analytical tools with the ability to interpret data from any mass spectrometer. In this study, we compare two software packages, Mascot and Peaks, for the analysis of iTRAQ data from ESI-Q/TOF mass spectrometry. In the case of a six-protein mixture combined in a known proportion, the output of the Peaks algorithm deviated from the correct result by 14% on average, while the error of the Mascot quantification was nearly 200%. When the software were used to analyse iTRAQ data from a complex protein sample, the quantification results agreed within 20% for only 26% of the quantified proteins, showing significant differences in the two quantification algorithms. This comparison and analysis revealed major intricacies in peptide and protein quantification that must be taken into consideration for software development.

Keywords: quantitative proteomics; iTRAQ; Mascot; Peaks

INTRODUCTION

Proteomics tools are applied to study the response of cells and tissues to environmental perturbations (nutrients [1], xenobiotics [2], infection [3]) and genetic alterations (cancer [4, 5], metabolic engineering [6, 7]) by identifying proteins with altered expression levels. Often, however, it is important to obtain more than a qualitative description of a proteome, and thus there has been an increase in efforts to develop quantitative proteomics methods. Quantitative proteomics can help characterize pathway regulation and complex system networks by providing protein concentration information corresponding to different cellular states. Such information is of utmost importance in the developing field of systems biology [8].

Current experimental approaches for quantitative proteomics include *in vivo* and *in vitro* methods [9–11]. Stable isotope labelling with amino acids (SILAC) [12, 13] or with small nitrogen- or

carbon-containing nutrients [14–16] can be incorporated in the growth medium to directly label proteins while synthesized *in vivo*. *In vitro* quantification methods include labelling of cysteine residues (isotope-coded affinity tags, ICAT) [17, 18] and labelling of lysines and N-termini (isobaric tags for relative and absolute quantification, iTRAQ) [19–21]. Other less commonly used quantification methods were reviewed by Lau *et al.* [22].

The iTRAQ technology has a significant advantage over other methods due to its capability of multiplexing up to eight samples in one experimental setup [23]. Another positive aspect includes unbiased peptide labelling, since iTRAQ isobaric tags label lysine side groups and all free amino-terminal groups of the peptides present in a sample. The iTRAQ tags consist of a reporter group, a balance group and a peptide reactive group that covalently binds to the peptides. The balance group gives all tags the same mass during peptide

Corresponding author. Kenneth F. Reardon, Department of Chemical and Biological Engineering, Colorado State University, Fort Collins, CO 80523, USA. Tel: 970 491 6505; Fax: 970 491 7369; Email: kenneth.reardon@colostate.edu

Carla M.R. Lacerda is a PhD candidate in the Department of Chemical and Biological Engineering at Colorado State University, Fort Collins, CO, USA.

Lei Xin is a Software Developer for Bioinformatics Solutions Inc., Waterloo, ON, Canada.

Iain Rogers is the V.P. of Business Development for Bioinformatics Solutions Inc., Waterloo, ON, Canada.

Kenneth F. Reardon is Professor and Associate Department Head in the Department of Chemical and Biological Engineering at Colorado State University, Fort Collins, CO, USA.

mass fingerprinting. In the collision-induced dissociation stage of a tandem mass spectrometer, there is a neutral loss of the balance group, and the reporter groups are detected in the second MS. The tandem mass spectra include contributions from each sample, and the individual contributions of each sample can be measured by the intensity of the reporter ion peaks.

The data analysis for iTRAQ requires special software packages that perform protein identification as well as quantification of reporter ions. Such quantification procedures vary, and different results can be obtained from different software, depending upon how peak intensity is calculated for peptide quantification, and on how peptide abundances are averaged for protein quantification. Currently, iTRAQ data analysis is supported by Mascot (Matrix Science Inc., Boston, MA, USA), I-Tracker [24], Libra (<http://tools.proteomecenter.org/Libra.php>), Pro Quant (Applied Biosystems, Foster City, CA, USA), SpectrumMill (Agilent Inc., Santa Clara, CA, USA) and by other software supplied by spectrometer vendors. In this study, we present a comparison between the quantification schemes of Mascot and Peaks, a quantification software from Bioinformatics Solutions Inc. (Waterloo, ON, Canada) slated for commercial release in 2008.

Peaks is a peptide mass spectrometry data analysis package that is well-known for its *de novo* sequencing algorithms, which use tandem mass spectrometry spectra to determine peptide sequences based on the measured fragmentation pattern without reference to sequence database. Peaks also contains database search algorithms for identification of peptides. These employ *de novo* sequences both for identification of homologues and in parallel with peptide fragment fingerprinting. The algorithms for data processing focus on peak detection, charge recognition, peptide identification and results validation. For quantitative proteomics applications, Bioinformatics Solutions developed the quantification algorithm evaluated in this study.

The goal of this study was to perform a comparison between the quantification algorithms of Peaks and Mascot, by evaluating similarities and differences in the results of these two approaches. Mascot was chosen as a widely used software, not associated with a particular mass spectrometer vendor. Two experiments were designed to accomplish these goals. The first was the quantification of a simple protein mixture (Applied Biosystems

six-protein mix provided in the iTRAQ Reagent Methods Development Kit). The second experiment evaluated protein expression in complex samples from a strain of *Burkholderia cepacia* exposed to different toxic heavy metals.

EXPERIMENTAL PROCEDURES

Simple protein mixture preparation and analysis

The six-protein mix and solutions used here were provided by the iTRAQ Reagent Methods Development Kit (Applied Biosystems). One vial of protein mix (containing 129 μg or 3.71 nmol) was used according to the iTRAQ reagents protocol from Applied Biosystems. Briefly, sample was dissolved in 20 μl of dissolution buffer and 1 μl of denaturant. After full sample dissolution, 2 μl of reducing reagent were added, and sample was incubated for 1 h at 60°C. Cysteines were blocked by the addition of 1 μl of cysteine blocking reagent followed by room temperature incubation for 10 min. The sample was then digested overnight with trypsin (10 μg to a final concentration of 0.3 mg/ml) at 37°C. After digestion, 34 μl of dissolution buffer were added to the resulting peptide mixture, which was then divided into two vials, labelled with 114 and 117 separately. The labelled peptides were mixed in the ratio 114:117 = 1:3. The peptide mix was purified with cation exchange cartridge—labelled peptides were diluted with 4.4 ml SCX buffer A (described in the upcoming section). The eluate from the cation exchange cartridge was vacuum dried and re-dissolved in 1 ml of 0.1% trifluoroacetic acid (TFA). The resulting concentration of this solution was about 1.8 pmol/ μl and 1800 fmol were injected into a Waters ESI-Q/TOF LC-MS/MS for analysis.

Complex protein mixture preparation and analysis

Burkholderia cepacia strain Cd44, identified by 16S rDNA sequencing, was isolated from a soil bacterial community based on its high cadmium resistance, i.e. ability to grow on 0.44 mM Cd^{2+} . After 24 h of growth in a defined mineral medium, cultures of strain Cd44 were subjected to additions of metals with different toxicities (0.44 mM Cd^{2+} , 0.38 mM Cr^{3+} and 0.62 mM Fe^{3+}) and one was kept without metals to serve as control. After 2 h of metal

exposure, cells were harvested and their proteins extracted and immediately digested prior to a shotgun proteomics workflow. Tryptic digestion of 100 µg of each sample followed the iTRAQ reagents protocol from Applied Biosystems (described in the previous section). iTRAQ labelling also followed the iTRAQ reagents protocol, with samples labelled according to the following scheme: control (no metal shock), 114; chromium shock, 115; iron shock, 116 and cadmium shock, 117. The labelled peptides were mixed and then fractionated by SCX chromatography using a Polysulphoethyl A column 100 mm × 2.1 mm × 5 µm, 200 Å (PolyLC Inc., Columbia, MD, USA). Buffer A was 5 mM potassium phosphate in 25% acetonitrile, pH 3, and Buffer B contained 500 mM potassium chloride added to the Buffer A composition. A 35 min linear gradient at 0.2 ml/min was used. Fractions were collected every minute, and purified with C18 PepClean spin columns (Pierce Biotechnology Inc., Rockford, IL, USA), according to the manufacturer's protocol. C18-cleaned samples were vacuum-dried and resuspended in 15 µl of 0.1% formic acid.

Samples were separated on an Agilent ZORBAX reverse-phase column (C18 wide-pore, 50 mm × 0.075 mm internal diameter) in an Agilent 1200 HPLC system, containing a nanospray directly connected to the mass spectrometer ionization source. The HPLC buffers were 0.1% formic acid in 3% acetonitrile (Buffer A) and 0.1% formic acid in 97% acetonitrile (Buffer B). The method used a 45 min gradient at 0.6 µl/min starting with 5% Buffer B for 5 min, followed by 30 min of ramping to 70% Buffer B. Two other 5 min steps followed at 90% and 5% Buffer B. Mass spectrometry was accomplished by an ESI-Q/TOF instrument (Agilent 6510, with Chip Cube sample inlet), with the following settings: gas temperature 300°C; drying gas 5 l/min and capillary voltage of 1875 V. The auto MS/MS mode was used, with positive ion polarity and profile data storage for the simple mixture and centroid data storage for the complex mixture. For data centroiding, the following thresholds were used: for MS absolute of 200 intensity counts and relative of 0.01%, for MS/MS absolute of 5 intensity counts and relative of 0.01%. Acquisition ranges were 300–2000 m/z for MS (4 spectra/s) and 59–1800 m/z for MS/MS (5 spectra/s). Collision energy used a slope of 4 (V/100 Da). Four precursor

ions were collected per cycle, with active exclusion. Internal reference mass standard was used for Chip operation.

Data preparation procedures

All datasets were exported to the mzData format and imported into Peaks Studio 4.5 sp2. Since a mass spectrometer may take an MS/MS scan of the same peptide several times, replicate spectra were merged (using the Peaks 'data refine' tool with a parent ion m/z tolerance of 0.1 and a retention time window of 30 s) to provide a more complete fragmentation pattern, better signal to noise ratio and a more representative intensity value for each reporter ion, before submission to both Peaks and Mascot. The Peaks 'data refine' tool uses an additive approach to merging. Providing the same merged dataset to each quantification algorithm also allowed for easier post-analysis comparison. Data were acquired in centroid mode, and parent charge states were assigned by the instrument, so MS/MS spectra were not subjected to further pre-processing prior to searching.

Identification procedures

Database searches were conducted on Mascot and Peaks using the following parameters: all bacterial species in the NCBI nr database, trypsin digestion allowing for up to two missed cleavages, methionine oxidation and iTRAQ on N-termini and lysines, iTRAQ 4-plex quantification, 0.1 Da tolerance for MS and MS/MS, peptide charges of +1, +2 and +3 and monoisotopic masses.

Quantification procedures

The data analysis stage of iTRAQ quantification relies on accurate computation of the intensities of reporter ion peaks in MS/MS spectra and rigorous statistical analysis of relative reporter ion intensities from multiple peptides in computation of protein expression ratios. The calculation of peak intensity, though trivial to the human eye, requires complex tasks of a computer algorithm, and appears to be one aspect in which Peaks and Mascot differ. When starting with profile data, the Peaks algorithm uses a dynamic baseline subtraction algorithm to recognize how much of the peak's area can be attributed to noise, recognizes the shape of a real peak and what part of the peak can be attributed to an interfering signal, bounds the width of the peak and determines at what point the peak sinks into the noise.

The pre-processing scheme used by Mascot was previously described by Berndt *et al.* [25], and is based on a sequence of steps, including peak detection, isotope distribution fit and subtraction of fit. To avoid bias created by any pre-processing algorithm, previously centroided data were used for both Peaks and Mascot.

The calculation of a protein expression ratio from the peptide ratios relies on accurate peptide identifications and accurate peptide ratios. Since these are not guaranteed, and are subject to some randomness, differences between the two algorithms evaluated here are noteworthy. The Peaks quantification algorithm performs a statistical analysis at the protein level. First, Peaks computes a weighting for each peptide ratio to allow peptide ratios from high quality spectra to be considered more reliable than those from spectra where reporter ions are barely visible. Second, outliers are removed before protein ratio computation (using Dixon's Test [26]) in Peaks since some peptide ratios can be erroneous—e.g. those resulting from false positive peptide identifications and/or chemical aberrations. Finally, to facilitate review of results, Peaks displays the protein expression ratio along with some measures of the success of the statistical analysis: standard deviation of the peptide ratio values as well as the coefficient of variation (CV) corresponding to this standard deviation.

A detailed description of the Mascot approach to iTRAQ quantification can be found on the Mascot website (http://www.matrixscience.com/help/quant_reporter_help.html). Briefly, identification and quantification are performed at the peptide level. Ratios for peptide matches are reported depending on peptide modification state, minimum precursor charge, strength of the peptide match, minimum number of fragment ion pairs, among other criteria. The Mascot result report assigns peptide matches to protein hits, and the ratios for individual peptide matches are combined to determine ratios for the protein hits (usually by weighted average). A ratio for a protein hit is only reported if two peptide matches are found. Standard deviations are only reported if the ratios for the peptide matches are consistent with a normal distribution. The Mascot result report assigns the peptide matches to protein hits, and the ratios for individual peptide matches are combined to determine ratios for the protein hits. The methods provided for calculating a protein ratio from a set of peptide ratios are median,

average or weighted average, where for each component, the intensity values of the set of peptides are summed and the protein ratio(s) calculated from the summed values. Ratios for peptide matches are only reported if various quality criteria are fulfilled, the most important of which are peptide modification state, minimum precursor charge, strength of the peptide match (defined in terms of either a minimum score, a maximum expected value or the score being at or above either the identity threshold or the homology threshold) and method-specific criteria, such as a minimum number of fragment ion pairs for multiplex. A ratio for a protein hit is only reported if the minimum number of peptide matches, is achieved. Testing for outliers and reporting a standard deviation for the protein ratio can only be performed if the peptide ratios are consistent with a sample from a normal distribution.

In the Peaks quantification workflow, supporting spectra is identified for each peptide match. Peaks searches through each spectrum in the supporting spectra set, and then calculates the intensity for each reporter ion. Peptide match ratios are calculated from reporter intensities and outliers are removed according to the weights. For each protein, all supporting peptides have their weights normalized to one and then weighted averages are calculated.

RESULTS AND DISCUSSION

Simple sample mixture

Table 1 lists the identifications and quantifications obtained from both Mascot and Peaks. High-score identifications were obtained from the Swiss-Prot database for beta-lactoglobulin (human and ovine), beta-galactosidase (*Escherichia coli*), apotransferrin (human), serum albumin (bovine) and lysozyme (chicken). No significant identification was obtained for alpha-lactalbumin, the sixth protein present in the protein mix, with either software. Good protein coverage and number of queries matched were found for all proteins identified in this mixture. Table 1 also summarizes the protein quantification results obtained for this mixture. The average protein quantification from Peaks was 14% lower than the expected value of 0.33 for the 114:117 reporter ion ratio, whereas Mascot provided values that averaged nearly three times the expected value. Protein-level quantification is based on peptide quantification, and also relies on an accurate peptide weighting strategy,

Table 1: Six-protein mixture identification and comparison of Peaks and Mascot protein-level quantification

Accession #	Mass (Da)	Score (%)	Coverage (%)	Queries matched	Ratio from Mascot quantification	Ratio from Peaks quantification	Peaks coefficient of variance
P02754 LACB_BOVIN	19 883	99	40	25	0.992	0.316	0.115
P67975 LACB_OVIMU	18 151	99	28	26	0.994	0.313	0.110
P00722 BGAL_ECOLI	116 483	99	16	21	1.067	0.237	0.098
P02787 TRFE_HUMAN	77 050	99	13	16	0.928	0.271	0.147
P02769 ALBU_BOVIN	69 294	99	13	13	0.866	0.370	0.252
P00698 LYSC_CHICK	16 239	99	23	3	1.014	0.214	0.152

Two preparations of this protein mixture were combined to provide a ratio of 0.33.

which will take into consideration the ionization method used, among other factors. Data analysis suggests that peptide quantification and weighting are the two main causes of the differences in protein quantification found in this study. However, the observed quantification differences might also be attributed to the composition of this particular six-protein mixture. Table 1 further presents the CV calculated by Peaks on the basis of the individual peptide quantification, which is useful to assess the accuracy of the quantification. This information is unavailable from Mascot. An expanded version of Table 1, containing all peptide quantification values, can be found in Supplementary Table 1.

Complex sample mixture

For the complex mixture of *B. cepacia* proteins, database searches were performed as described in the ‘Experimental Procedures’ section and similar results for protein identifications were obtained from both software packages. However, challenges were faced when merging the two datasets, mostly due to the distinct scoring systems in which proteins and peptides are ranked differently. In order to compare the two quantification schemes, it was necessary to construct the results tables manually. Some of these results are presented in Supplementary Table 1. Supplementary Table 1 also presents peptide sequences and isoelectric points. This information was of interest in the case where similar peptide sequences or isoelectric points were found to be markers of differences in quantification. No such patterns were found for the peptide sample used for this quantification study. The full quantification results from Peaks are presented in Supplementary Table 2. The complete tables contain all the information required for a proteomics experiment, namely: protein name and accession numbers, the organism in which the protein was first identified,

protein and peptide scores, peptide sequence and protein sequence coverage, protein and peptide ratios and standard deviations and peptide weights.

The agreement between Mascot and Peaks quantification varied widely. Quantification results in good agreement (within 20%) were obtained for 26% of the proteins sampled in this study. Adequate agreement (within 50%) was found for 60% of the proteins and poor agreement (difference >100%) was found in 9% of the proteins analysed here. Scatter plots relating the quantification results of Peaks and Mascot are presented in Figure 1. Protein ratios determined by Mascot were generally higher than those from Peaks when ratios were higher than 1.0 (Figure 1A) but this trend was not evident for ratios below 1.0. No correlations (e.g. with pI or protein ratio) were found to distinguish the proteins with poor agreement from those that were quantified similarly using the two software packages.

Two pronounced cases of disagreement between the two methods are described here as examples. The first comes from a peptide (QAVEIAETNFNAA AAVATK) identified from phasin, for which the average peptide 115:114 ratio was 15.975 according to Mascot and 2.176 according to Peaks. This difference can be attributed to the various peptide quantification factors discussed previously. It is interesting to note that such a large value might be generated by an inefficient removal of outliers during the averaging of the reporter peaks. Visual inspection of a sample spectrum of the reporter genes for this peptide (Figure 2) reveals that the 115:114 peak ratio is closer to 2 than to 16. Interestingly, the overall protein ratio for phasin reported by Peaks was 5.8, which corresponded well to the value of 5.98 produced by Mascot. A second example comes from chaperonin GroEL. The 116:114 protein ratio was 0.275 according to Peaks and 3.77 according to Mascot. Representative peptide ratios are 0.437

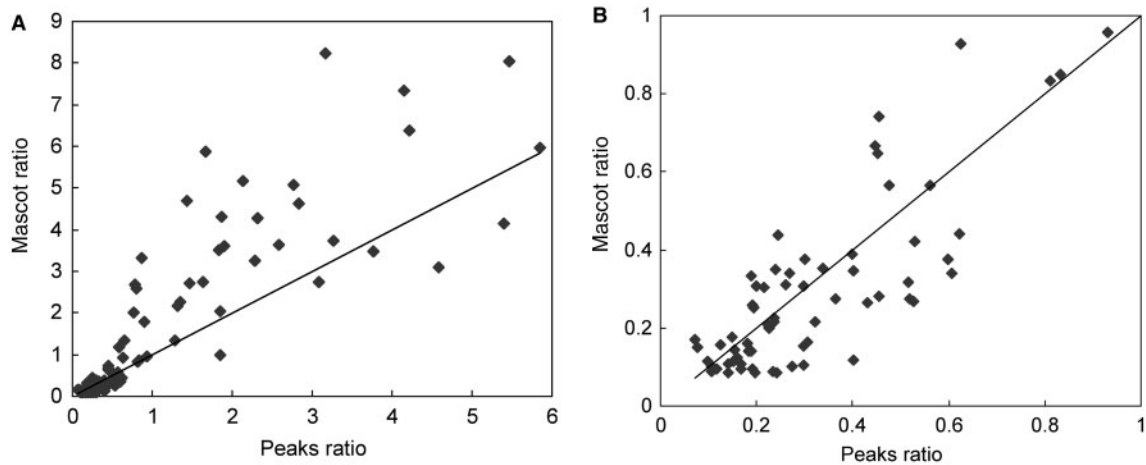


Figure 1: Scatter plots with Mascot and Peaks protein quantification results for (A) wide ratio range and (B) ratios < 1.0. Ratios are calculated relative to iTRAQ reporter ion 114. Lines (1 : 1) indicate perfect agreement between the two analyses.

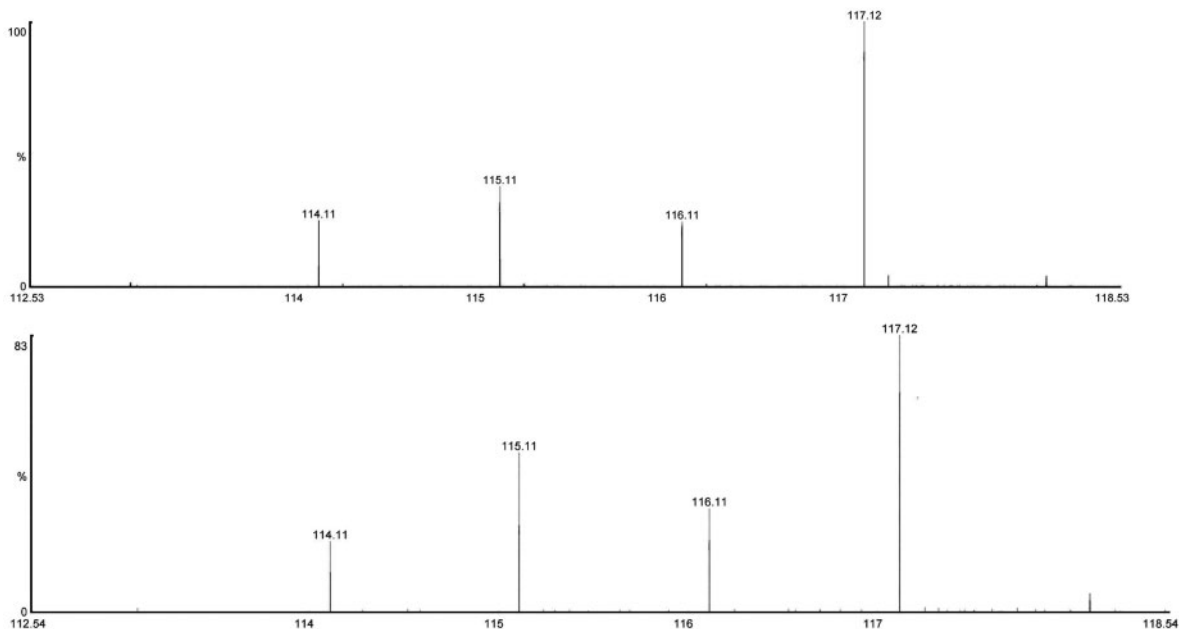


Figure 2: Portion of MS/MS spectra showing reporter ions for peptide QAVEIAETNFNAAAATK from phasin (in the complex protein mixture).

for Peaks and 1.466 for Mascot for peptide AAVEGIVAGGGVALIR (i.e. no agreement regarding up- or down-regulation) and 0.084 for Peaks and 0.655 for Mascot for peptide TALQNAASVAGLLLTDAAVAELPK. Other cases of differences in peptide and protein quantification can be found in Supplementary Table 1.

The correct protein expression ratios are not known for the complex mixture, so evaluation of the results must be performed by means other than calculation of standard errors. One such approach makes use of the experimental design. Since the

peptide mixture was divided into several fractions through SCX chromatography, we can expect to identify a protein in several fractions (usually represented by different peptides) and the same abundance ratios should be observed for that protein in each fraction. A basis for algorithm comparison is thus the degree of consistency of the protein ratio calculations across the fractions. Examples of such comparisons are shown in Figures 3 and 4. For phasin (gi|78066909|), the 116:114 ratio as calculated by Peaks is 1.8 with a variance of 0.1 across all fractions, the Peaks-calculated 115:114 ratio is

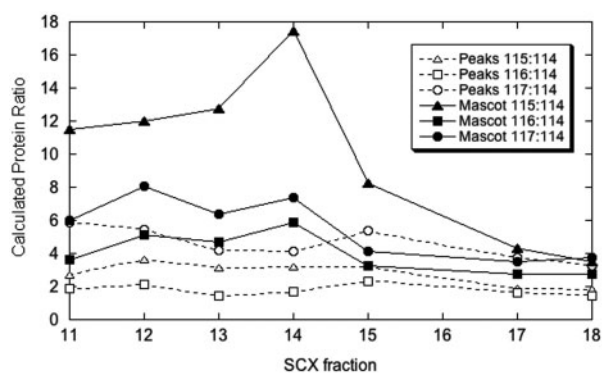


Figure 3: Phasin protein ratios (relative to the control, iTRAQ II4) across SCX fractions as calculated by Mascot and Peaks.

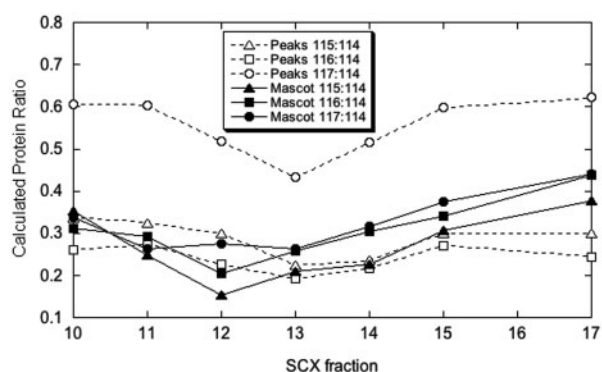


Figure 4: Chaperonin GroEL protein ratios (relative to the control, iTRAQ II4) across SCX fractions as calculated by Mascot and Peaks.

2.8 with a variance of 0.5 and the 117:114 ratio for this protein is 4.6 with a variance of 0.9 (Figure 3). The Mascot calculations show considerably higher variations across fractions (115:114 ratio of 10.0 with a variance of 24.4, 116:114 ratio of 4.0 with a variance of 1.5, 117:114 ratio of 5.6 with a variance of 3.3). However, when this analysis is performed for chaperonin GroEL (gi|107021936|ref|YP_620263.1|), both tools are found to provide consistently similar ratio calculations across the fractions, with a lower variance in Peaks results in two of the three cases (Figure 4). All these calculations are available in Supplementary Table 1.

CONCLUSIONS

This study reports the fundamental challenges involved in analysing quantitative shotgun proteomics data. The quantification performance of Mascot and Peaks were compared using iTRAQ datasets from a six-protein mixture and a complex protein

sample. The quantification results obtained from Peaks were found to be more accurate in the case of the simple mixture. The analysis of the complex protein mixture revealed substantial differences between the results from the two programs that are attributed to a variety of differences in the algorithms used by each (e.g. for calculation of reporter ion peak intensity for peptide quantification, and efficient weighting and removal of outliers for protein quantification). If profile data are used, additional differences would likely be observed, since the pre-processing steps performed by the two programs are different. This analysis also demonstrated that quantification software needs further development. As of yet, there is no consensus on how peptide and protein quantification should be accomplished, and direct comparison of results obtained by two software packages can be difficult. The results of this study suggest that quantitative shotgun proteomics data should be interpreted carefully.

SUPPLEMENTARY MATERIAL

Supplementary material is available at *Briefings in Functional Genomics and Proteomics* online.

Key Points

- iTRAQ quantification relies on accurate computation of the intensities of reporter ion peaks in MS/MS spectra and rigorous statistical analysis of relative reporter ion intensities from multiple peptides in computation of protein expression ratios.
- Automated calculation of reporter ion peak intensity is non-trivial and different algorithms can produce significantly different results.
- The calculation of a protein expression ratio from the peptide ratios depends upon accurate peptide identifications and accurate peptide ratios, and thus results are strongly affected by the weighting of peptide ratios and the exclusion of outliers.
- The evaluation of protein quantification results across different shotgun fractions can provide insights into the reliability of the data.

Acknowledgements

We would like to acknowledge Dr Richard Reisdorph and Dr Nichole Reisdorph of the National Jewish Research Center (Denver, CO, USA) for technical support. This work was supported in part by a National Science Foundation Grant (BES-0329514) to KFR.

References

1. Sarry JE, Kuhn L, Le Lay P, *et al.* Dynamics of *Arabidopsis thaliana* soluble proteome in response to different nutrient culture conditions. *Electrophoresis* 2006; 27:495–507.

2. Leonard JF, Courcol M, Mariet C, *et al.* Proteomic characterization of the effects of clofibrate on protein expression in rat liver. *Proteomics* 2006;**6**:1915–33.
3. Scharlaken B, de Graaf DC, Memmi S, *et al.* Differential protein expression in the honey bee head after a bacterial challenge. *Arch Insect Biochem Physiol* 2007;**65**:223–37.
4. Kuruma H, Egawa S, Oh-Ishi M, *et al.* Proteome analysis of prostate cancer. *Prostate Cancer Prostatic Dis* 2005;**8**:14–21.
5. Somiari RI, Somiari S, Russell S, *et al.* Proteomics of breast carcinoma. *J Chromatogr B Analyt Technol Biomed Life Sci* 2005;**815**:215–25.
6. Lee PS, Lee KH. Engineering HlyA hypersecretion in *Escherichia coli* based on proteomic and microarray analyses. *Biotechnol BioEng* 2005;**89**:195–205.
7. Park SJ, Lee SY, Cho J, *et al.* Global physiological understanding and metabolic engineering of microorganisms based on omics studies. *Appl Microbiol Biotechnol* 2005;**68**:567–79.
8. Aggarwal K, Lee KH. Functional genomics and proteomics as a foundation for systems biology. *Brief Funct Genomic Proteomic* 2003;**2**:175–84.
9. Chen X, Sun LW, Yu YB, *et al.* Amino acid-coded tagging approaches in quantitative proteomics. *Expert Rev Proteomics* 2007;**4**:25–37.
10. Fröhlich T, Arnold GJ. Proteome research based on modern liquid chromatography – tandem mass spectrometry: separation, identification and quantification. *J Neural Transm* 2006;**113**:973–94.
11. Heck AJR, Krijgsveld J. Mass spectrometry-based quantitative proteomics. *Expert Rev Proteomics* 2004;**1**:317–26.
12. Mann M. Functional and quantitative proteomics using SILAC. *Nat Rev Mol Cell Biol* 2006;**7**:952–8.
13. Ong SE, Blagoev B, Kratchmarova I, *et al.* Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* 2002;**1**:376–86.
14. Beynon RJ, Pratt JM. Metabolic labeling of proteins for proteomics. *Mol Cell Proteomics* 2005;**4**:857–72.
15. Palmblad M, Bindschedler LV, Cramer R. Quantitative proteomics using uniform N-15-labeling, MASCOT, and the trans-proteomic pipeline. *Proteomics* 2007;**7**:3462–9.
16. Snijders APL, de Vos MGJ, Wright PC. Novel approach for peptide quantitation and sequencing based on N-15 and C-13 metabolic labeling. *J Proteome Res* 2005;**4**:578–85.
17. Vaughn CP, Crockett DK, Lim MS, *et al.* Analytical characteristics of ICAT-LC-MS/MS for quantitative proteomics studies. *J Mol Diagn* 2005;**7**:691–2.
18. Wu WW, Wang GH, Baek SJ, *et al.* Comparative study of three proteomic quantitative methods, DIGE, iCAT, and iTRAQ, using 2D gel- or LC-MALDI TOF/TOF. *J Proteome Res* 2006;**5**:651–8.
19. Boehm AM, Putz S, Altenhofer D, *et al.* Precise protein quantification based on peptide quantification using iTRAQ (TM). *BMC Bioinformatics* 2007;**8**:214–31.
20. Melanson JE, Avery SL, Pinto DM. High-coverage quantitative proteomics using amine-specific isotopic labeling. *Proteomics* 2006;**6**:4466–74.
21. Stensjo K, Ow SY, Barrios-Llerena ME, *et al.* An iTRAQ-based quantitative analysis to elaborate the proteomic response of *Nostoc* sp. PCC 7120 under N-2 fixing conditions. *J Proteome Res* 2007;**6**:621–35.
22. Lau KW, Jones AR, Swainston N, *et al.* Capture and analysis of quantitative proteomic data. *Proteomics* 2007;**7**:2787–99.
23. Choe L, D’Ascenzo M, Relkin NR, *et al.* 8-Plex quantitation of changes in cerebrospinal fluid protein expression in subjects undergoing intravenous immunoglobulin treatment for Alzheimer’s disease. *Proteomics* 2007;**7**:3651–60.
24. Shadforth IP, Dunkley TPJ, Lilley KS, *et al.* i-Tracker: for quantitative proteomics using iTRAQ (TM). *BMC Genomics* 2005;**6**:145–50.
25. Berndt P, Hobohm U, Langen H. Reliable automatic protein identification from matrix-assisted laser desorption/ionization mass spectrometric peptide fingerprints. *Electrophoresis* 1999;**20**:3521–6.
26. Dixon WJ. Processing data for outliers. *Biometrics* 1953;**9**:74–89.