Technical Note

Protein Quantification by Progenesis LC-MS

Overview

- Inferring protein expression changes from measurements made at the peptide level is the final critical step in LC-MS data analysis for proteomics research[1,2,3]
- Here we describe how Progenesis LC-MS performs this key process on label-free data and the advantages it gives to generate reliable measures of relative protein expression
- We also provide published evidence supporting the accuracy and reliability of how Progenesis LC-MS generates protein level quantification[4,5,6]

Introduction

Current instruments for generating labelled or label-free LC-MS data can measure peptide ions in your proteomics samples with high accuracy and sensitivity. The challenge then, is how to recombine all these individual peptide ion measurements and make sense of what that means at the protein level. Several solutions exist including:

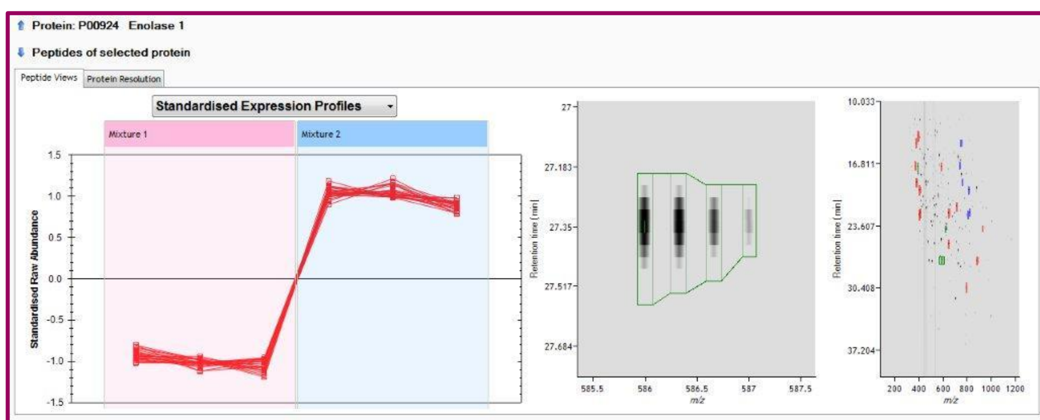
- Setting rules so a certain number of pre-defined features must all change in a statistically significant way, and in the same direction, for a protein to be considered differentially expressed[3]
- Averaging intensities of all peptide ions detected for a specific protein and comparing between groups[3]

These approaches have issues when it comes to analysing real samples, such as:

- Not every peptide is measured in every run; some may not fragment or generate reliable identifications[1,2,3]
- Differences in physicochemical properties mean not all peptide signals from a specific protein correlate well with the overall protein abundance[3]
- Peptide sequences are not always unique; a peptide can be common to more than one protein[2,3,2]

Progenesis LC-MS

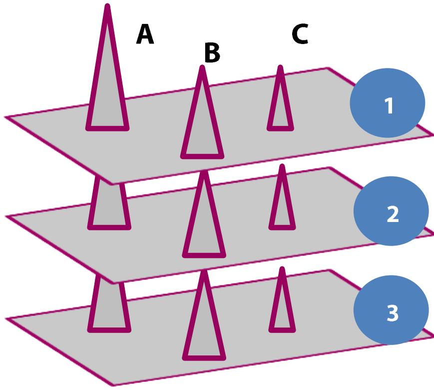
Detection and quantification of all peptide ions is followed by expression analysis and identification of the proteins from which they originate. **Protein abundance is calculated from the sum of all unique normalised peptide ion abundances for a specific protein on each run**. Alignment and co-detection of features means you have exactly the same number of quantified peptides identified on all runs so you can compare the sum of ion abundances between groups.



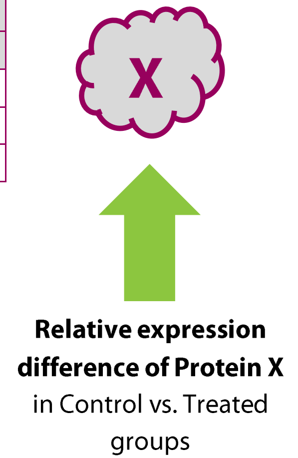
Protein quantification and relative expression changes in Progenesis LC-MS. All peptide ions displayed in the ion intensity map, far right, were quantified and then identified as being unique to Enolase I. Protein quantification is automatically generated by summing all peptide ion isotope abundances (see example feature in green) unique to a specific protein in each run. Relative protein expression (fold changes) and reliability of the measured differences (ANOVA p-value) are calculated between each group, in this case Mixture 1 vs. Mixture 2, using the summed peptide ion abundances. The expression profile (far left) shows all relative peptide ion abundances within group and between groups.

How does Progenesis LC-MS quantify protein abundance?

Control runs



Control runs	
Peptides unique to Protein X	
Run 1	Σ peptide ion abundances A, B, C
Run 2	Σ peptide ion abundances A, B, C
Run 3	Σ peptide ion abundances A, B, C

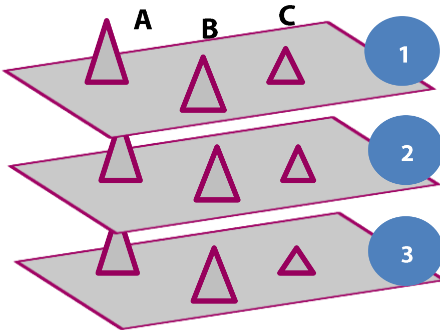


ANOVA p-value & fold change calculated from **SUM of unique peptide ion abundances per run** for a specific protein

Relative expression difference of Protein X in Control vs. Treated groups

x

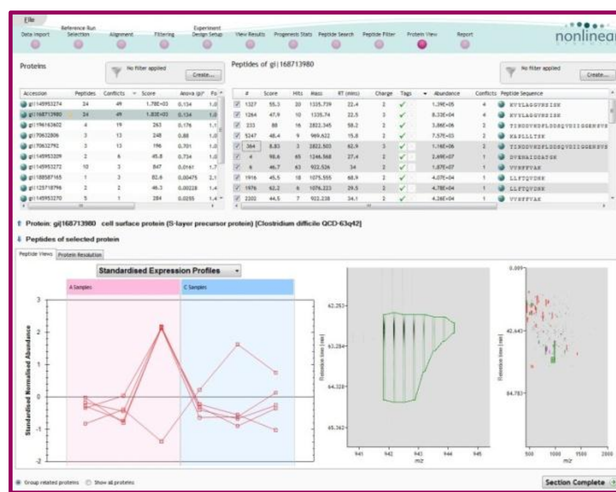
Treated runs



Treated runs	
Peptides unique to Protein X	
Run 1	Σ peptide ion abundances A, B, C
Run 2	Σ peptide ion abundances A, B, C
Run 3	Σ peptide ion abundances A, B, C

Features of protein quantification

1. Co-detection of features provides ion abundance measures for the same peptides from each protein on every run, with no missing values
2. Summing means that ion abundances with higher signal, and therefore less "noisy", have more weight; making the final protein quantification more reliable
3. When two protein identifications share common peptides the software uses unique peptides for quantification
4. Searching peptides against databases can return multiple entries that are actually the same or related proteins, so the software automatically groups similar proteins into one quantification result
5. Fractionated proteins are quantified in the same way but include all peptides from every fraction



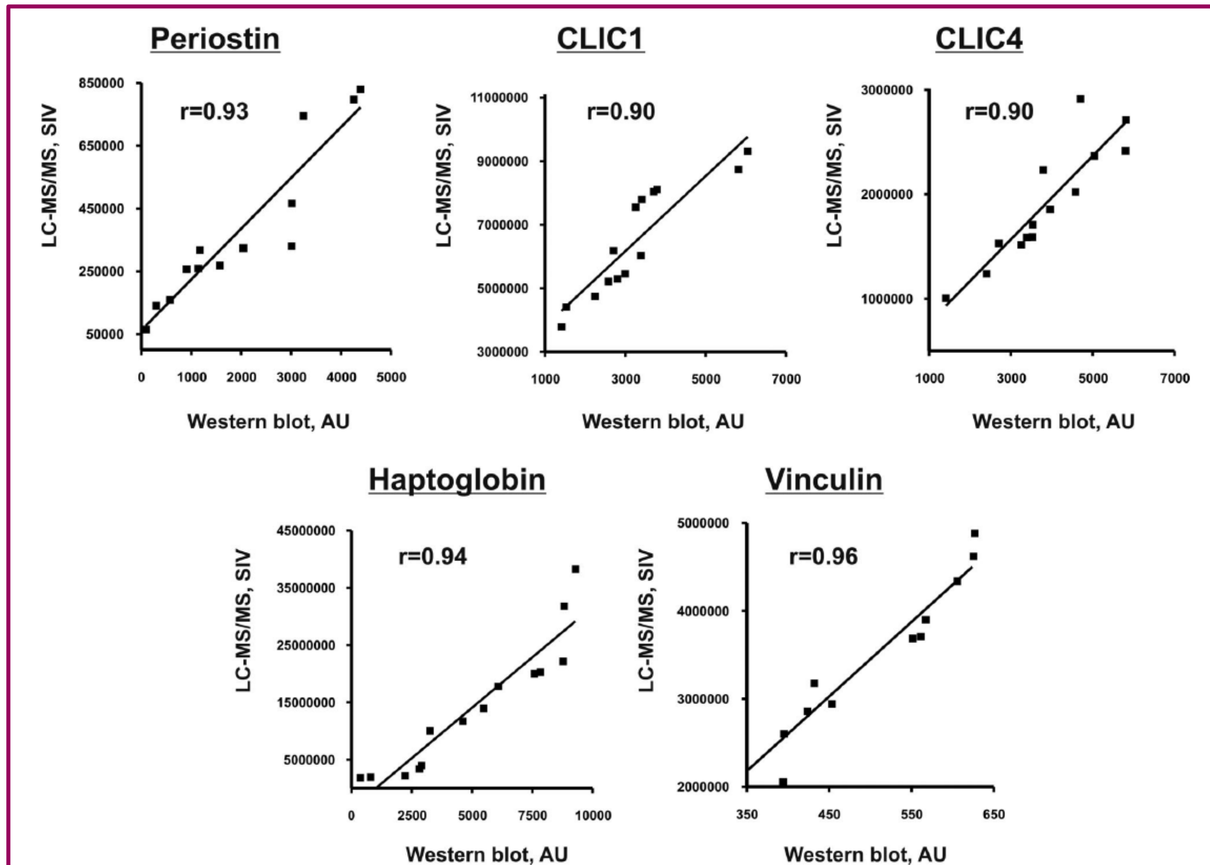
The results are shown in the Protein View above, which includes: ● Maximum mean abundance for a protein within an experiment. ● Fold-change and p-value for protein expression differences between groups. ● The group with highest and lowest mean protein abundance. ● Number of peptides quantified-then-identified per protein. ● A combined "protein score" from combining all individual peptide scores. ● Expression profiles for peptides that contribute to protein quantification.

progenesis

Nonlinear Dynamics Ltd **www.nonlinear.com**
Keel House | Garth Heads | Newcastle upon Tyne | NE1 2JE | UK | tel: +44 (0)191 230 2121
Nonlinear USA Inc I 2530 Meridian Parkway I 3rd Floor I Durham I NC 27713 I USA
tel: +1 919 806 4401 I toll free: +1 866 435 7872 I fax: +1 919 806

Western Blots confirm protein quantification measured by Progenesis LC-MS[4]

Label-free proteomics identified significant protein expression differences in 25 proteins (out of 300 detected) relevant to the disease pathway of pulmonary arterial hypertension (PAH) between surgical samples of control patients and those with the disease. Several proteins were selected for validation of expression changes by Western Blotting. This included proteins with: highest and lowest relative abundance changes, proteins with more than one member of its family displaying increased abundance and a structural protein that was not significantly different. **There was high correlation (r >0.9, linear regression analysis) between abundances determined by LC-MS/MS and Western Blot analysis.**



Relationship between proteomic and Western blot analyses. Measurements of relative protein levels for each sample determined by proteomics (sum of peptide ion intensity values [SIV]) and densitometry of Western blots (arbitrary units [AU]). In each case, there is a strong correlation between the 2 methods of quantification, confirming the initial proteomics observations using LC-MS/MS analyses.

Quantification of protein expression differences validated by immunohistochemistry and pathway analysis[5]

Label-free LC-MS/MS found 334 proteins (from 893 detected) differentially expressed ($p<0.05$, 2-fold change) between samples from Equine recurrent uveitis (ERU) disease with healthy controls. Significantly regulated proteins were subjected to GO analyses, protein network analyses (STRING) and pathway overrepresentation analyses. This grouped expression changes between healthy and control samples, which confirmed expected hypothesis and findings from other studies.

Selected candidate markers for ERU were validated on large sample sizes by immunohistochemistry on paraffin-embedded tissue specimen and image analyses. Results matched differences measured by LC-MS. The work also assessed variation between cumulative peptide intensities for each identified protein. The average coefficient of variation for all proteins identified with ≥ 2 peptides was 17.1% for healthy tissues and 11.2% for disease.

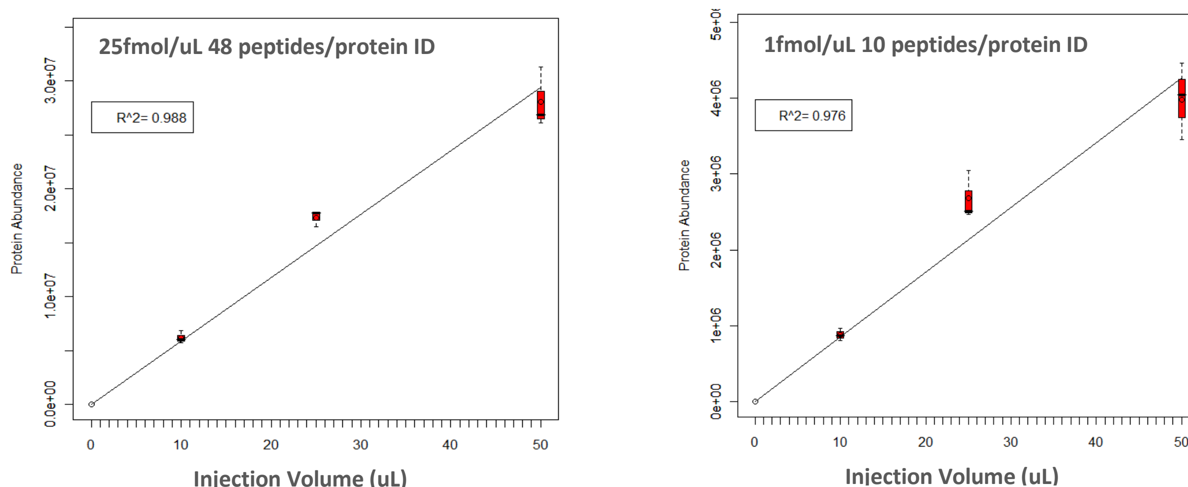
Nonlinear Dynamics Ltd
www.nonlinear.com
Keel House | Garth Heads | Newcastle upon Tyne | NE1 2JE | UK | tel: +44 (0)191 230 2121
Nonlinear USA Inc I 2530 Meridian Parkway I 3rd Floor I Durham I NC 27713 I USA
tel: +1 919 806 4401 I toll free: +1 866 435 7872 I fax: +1 919 806

Validating Progenesis LC-MS - quantifying known proteins in the 2-1000 fmol range[6]

This study validated the analysis approach by **quantifying and identifying nine known proteins in the 2-1000 fmol range.** We demonstrated that the **linearity of measured response is consistently high (>0.9) across this range**.

"Accurately quantify and identify a trypsinised mixture of proteins spanning a 2-1000 fmol concentration range using Progenesis LC-MS BUT without knowing how many or what proteins you're looking for"
Dr. Alexander Ivanov at the Harvard School of Public Health in Boston, USA

This was achieved! The linearity of response for measuring increasing amounts of all nine known proteins was assessed based on plots of calculated protein abundance vs. injection volume. Figure below shows two examples of such plots. The value of R^2 is not significantly affected by either the number of peptide ions or the protein concentration in the injected sample.



Best linear fit measured from plots of protein abundance vs. injection volume (10, 25 and 50uL) including box plots showing replicate variation for (A) Serotransferrin precursor R2 =0.988 (B) Alcohol dehydrogenase I R2 = 0.976.

Conclusion

Progenesis LC-MS offer advantages to applying *ad-hoc* rules or averaging intensities to infer protein abundance and measure relative protein expression changes. Detection and quantification of all peptide ions is followed by expression analysis and identification of the proteins they originate from. **Protein abundance is calculated from the sum of all unique normalised peptide ion abundances for a specific protein on each run**. Alignment and co-detection of features means you have exactly the same number of quantified peptides identified on all runs so you can compare sum of ion abundances between groups for calculating relative expression differences. The unique approach for label-free data analysis, protein quantification and measuring relative protein expression has been validated by published data. Visit www.nonlinear.com/lc-ms and **try it on your data today!**

References

1. Vaudel M, Sickmann A, Martens L. Peptide and protein quantification: a map of the minefield. Proteomics. 2010 Feb;10(4):650-70
2. Bantscheff M, Schirle M, Sweetman G, Rick J, Kuster B. Quantitative mass spectrometry in proteomics: a critical review. Anal Bioanal Chem. 2007 Oct;389(4):1017-31.
3. Clough T, Key M, Ott I, Ragg S, Schadow G, Vitek O. Protein Quantification in Label-Free LC-MS Experiments. J. Proteome Res., 2009, 8 (11), pp 5275–5284
4. Abdul-Salam VB, Wharton J, Cupitt J, Berryman M, Edwards RJ, Wilkins MR,. Proteomic analysis of lung tissues from patients with pulmonary arterial hypertension. Circulation 122:2058-2067
5. Hauck SM, Dietter J, Kramer RL, Hofmaier F, Zipplies JK, Amann B, Feuchtinger A, Deeg CA, Ueffing M. Deciphering membrane-associated molecular processes in target tissue of autoimmune uveitis by label-free quantitative mass spectrometry. Mol Cell Proteomics, Mol Cell Proteomics 9:2292-2305.
6. M. O'Gorman, A. R. Ivanov, P. Lavery, M. Bennett. Progenesis LC-MS - Validating a unique approach for label-free LC-MS data analysis. ASMS 2009

Nonlinear Dynamics Ltd www.nonlinear.com
Keel House | Garth Heads | Newcastle upon Tyne | NE1 2JE | UK | tel: +44 (0)191 230 2121
Nonlinear USA Inc I 2530 Meridian Parkway I 3rd Floor I Durham I NC 27713 I USA
tel: +1 919 806 4401 I toll free: +1 866 435 7872 I fax: +1 919 806