

RELATIVE PROTEIN QUANTITATION WITH POST TRANSLATIONAL MODIFICATIONS IN MASS SPECTROMETRY BASED PROTEOMICS

Jens Allmer

Molecular Biology and Genetics, Izmir Institute of Technology
Gulbahce Campus, Urla, Izmir, 35430, Turkey
phone: + (90) 232 750 7517, fax: + (90) 232 750 7509, email: jens@allmer.de
web: <http://jens.allmer.de>

ABSTRACT

Mass spectrometry has become the tool of choice for most investigations in proteomics. Identification of proteins from complex mixtures has long been achieved and is now routinely used in countless high throughput studies. Quantitation by mass spectrometry is comparably newer and many different strategies have been proposed. One such strategy quantitates the difference in protein expression level among samples via extracted ion chromatograms, or spectral counts or a combination thereof. Another strategy involves mass modifications of the analytes in one or more of the samples under investigation. MSMAG has been developed as an extension to 2DB and it has been shown that it can aid in quantitation of data from experiments employing label-free quantitation. Recently, it has been extended to allow for analysis of data based on labelling strategies. This also makes it possible to quickly visualize and investigate inherent mass differences as presented by post translational modifications.

1. INTRODUCTION

Proteomics aims to elucidate the protein complement of a genome taking into account the spatial and temporal expression patterns of proteins. Historically one gene was thought to produce one protein; today we know that due to alternative splicing, RNA editing, and protein splicing this is hardly the case and one gene usually gives rise to multiple proteins with different sequences [1,2]. This creates a large number of possible gene products which is further increased by post translational modifications (PTMs) in which chemical groups are attached to amino acids. Adding PTMs to amino acids in the sequence does not lead to a new protein since the sequence remains the same but rather creates a new protein species [3,4]. Systems biology is interested in all protein species expressed by a genome, their interactions, expression patterns, locations, and their absolute and relative quantity under all possible conditions [5]. While the goal seems elusive it is important to create the tools necessary to enable steps towards the overall aim of systems biology.

Mass spectrometry (MS) has become the tool of choice when studying proteins [6]. Due to the complexity of MS data many software tools [7-9] have been proposed for analysis and different data storage facilities have been developed [10-12]. We recently developed 2DB, an application

for storage, analysis and presentation of results from MS/MS analyses which was first introduced at the 2007 HIBIT and whose improvements have been presented at the 2009 HIBIT symposium [13]. In addition to protein identification, which MS has long been used for, protein quantities and their relative differences among samples can now be examined. Two general strategies are employed for quantitation. One of these involves differential labelling of peptides in different samples with a marker which changes the mass such that the mass to charge ratio (m/z) difference between labelled and unlabelled analytes can be differentiated by a mass spectrometer [14-16]. Labelling implies an additional effort which involves extra cost and an increase in labour time as well as a potential raise in sample complexity [17] whereas label-free quantitation may be done without extra effort. Relative protein quantities can be determined without the need for adding a label to the analytes based on the notion that protein abundance and number of spectra and intensity of precursor ions are correlated [18-20]. One method to perform label-free quantitation with liquid chromatography (LC) MS/MS data is to perform spectral counting [21-23]. In addition to spectral counting, the count can be weighted by the total ion current (TIC) [24,25] or the reported score of the identification software [26]. MSMAG is an extension to 2DB and can be used for label-free quantitation essentially as a new feature in the analysis and presentation module while the underlying data model remains unchanged [25]. Since PTMs are also modeled in the database, it was natural to extend the quantitation facility and allow for quantitation of labeled analytes. Viewing labels as an arbitrary change in the measurable mass to charge ratio (m/z) opens the possibility to treat inherent PTMs as if they were labels. With this notion, MSMAG has been extended to allow for visualization of the distribution of PTMs over the fractions of a sample. Additionally, the interactive quantitation of labels and PTMs among samples has been enabled and thus relative quantities of modified and unmodified analytes can be investigated. A dataset has been created *in silico* to test the new features. This dataset will be presented in the next section.

2. PRELIMINARIES

MSMAG is a part of a larger system and provides just one of several analysis and display possibilities of mass spectrometry based proteomics data stored in 2DB.

2.1 Data Import and MSMAG Analysis Steps

- First an instance of 2DB needs to be created if none is available (see 2DB manual for details).
 - This step involves import of protein sequence information for target organisms.
 - Setting confidence thresholds for the MS/MS identification software used.
- Then data from MS/MS experiments can be imported into the database. Different formats such as standards like pepXML are supported in this step (<http://tools.proteomecenter.org/wiki/index.php?title=Formats:pepXML>).
- The data can be amended with location information as for example an image of the source gel.
- In regards to analysis with MSMAG, the identified PTMs are automatically extracted from these results (mass differences only).
- PTMs whose mass have not been named in 2DB can be given a name using 2DB functions (for readability).
- After all data has been imported, the significantly identified proteins can be investigated using 2DB functions.
- For in-depth analysis MSMAG can be selected from the Experiments menu of 2DB.
- Within MSMAG the first step is to select the experiments in 2DB that shall be analyzed.
- With that selection it is possible to either perform an analysis of the peptide and protein distribution over the experiment or investigate the differential peptide/ protein expression among selected experiments.
- The expression profile displays the proteins, their associated peptides, and their associated mass spectra (Figure 1).
- The quantitation choice (Figure 3) needs several settings before data can be displayed.
- For normalization of data housekeeping proteins can be selected from peptides shared among samples.
- The PTMs that shall be considered need to be selected from the list of available PTMs.
- The pooling type can be set (Settings – Pooling Type).
- Also some mathematical calculations can be selected or deselected in the Settings menu.
- View allows for modifying the output and allows for showing raw data instead of calculated ratios for instance.
- The most important setting is under Run - Quantitation Mode. Here the appropriate quantitation

method needs to be chosen before selecting Run – Quantitation.

- The results can be viewed in MSMAG, copied from the provided data tables, saved in different formats for later offline analysis and exported as images.

3. MATERIALS AND METHODS

In a study on simulating spectra from LC-MS data, Schulz-Trieglaff et al. created large *in silico* datasets based on mouse proteins [27]. When introducing MSMAG the data accompanying its introduction were based on that dataset since no other, more targeted, experimental dataset was freely available. Modifying the dataset to include PTMs was achieved by randomly modifying the peptide identifications in the dataset in regard to their precursor mass and their abundance. Three different modifications were used which were named ctLA, ctLB, and ctLC and were set to create a shift in precursor mass by 5, 10, and 21 Dalton (Da), respectively. The modifications are assumed to have no influence on the chemical properties of the amino acid and shall therefore not lead to a shift in the LC elution time. Each identified peptide was selected in turn and a random number between 1 and 10 was generated. If the random number was 3 or less, the peptide was modified by generating another random number in-between 1 and 3 to select one of the three possible modifications. For each identified peptide the abundance was adjusted in case a modification was added. The TIC was adjusted between 70 and 160% of its original value; again by generating a random number in that range. The modification was added to the identification and the original identification was removed from the dataset. This leads to a serration of data with less identification per unmodified peptide. With some quantitation methods, it is essential to have both a modified and an unmodified version of the peptide within the pool of identifications. With this method of generating the data set, this cannot be guaranteed and it seems likely that the number of quantifiable identifications may be rather low.

The data was imported into an instance of 2DB (<http://www.biolnk.com/2DB>) where it can be accessed publicly. The experiments were entitled A-LUin1 to F-LUin1. This dataset shall serve as an example in the following sections.

4. RESULTS AND DISCUSSION

4.1 Protein and Peptide Expression Profiling

Data that has been imported into 2DB and can be analyzed using MSMAG which can be accessed by selecting Experiments - MSMAG.

Selecting just one Experiment (here equal to one sample) allows for analysis of the protein and peptide distribution among the fractions of the experiment (Figure 1). In Figure 1, the occurrence of the protein with the JGI (<http://genome.jgi-psf.org/Chlre4/Chlre4.home.html>) accession number 137778 in the fractions of the experiment A-LUin1 is shown in the top pane. Most of the protein can be found in fractions 6 through 9. The protein list in the left

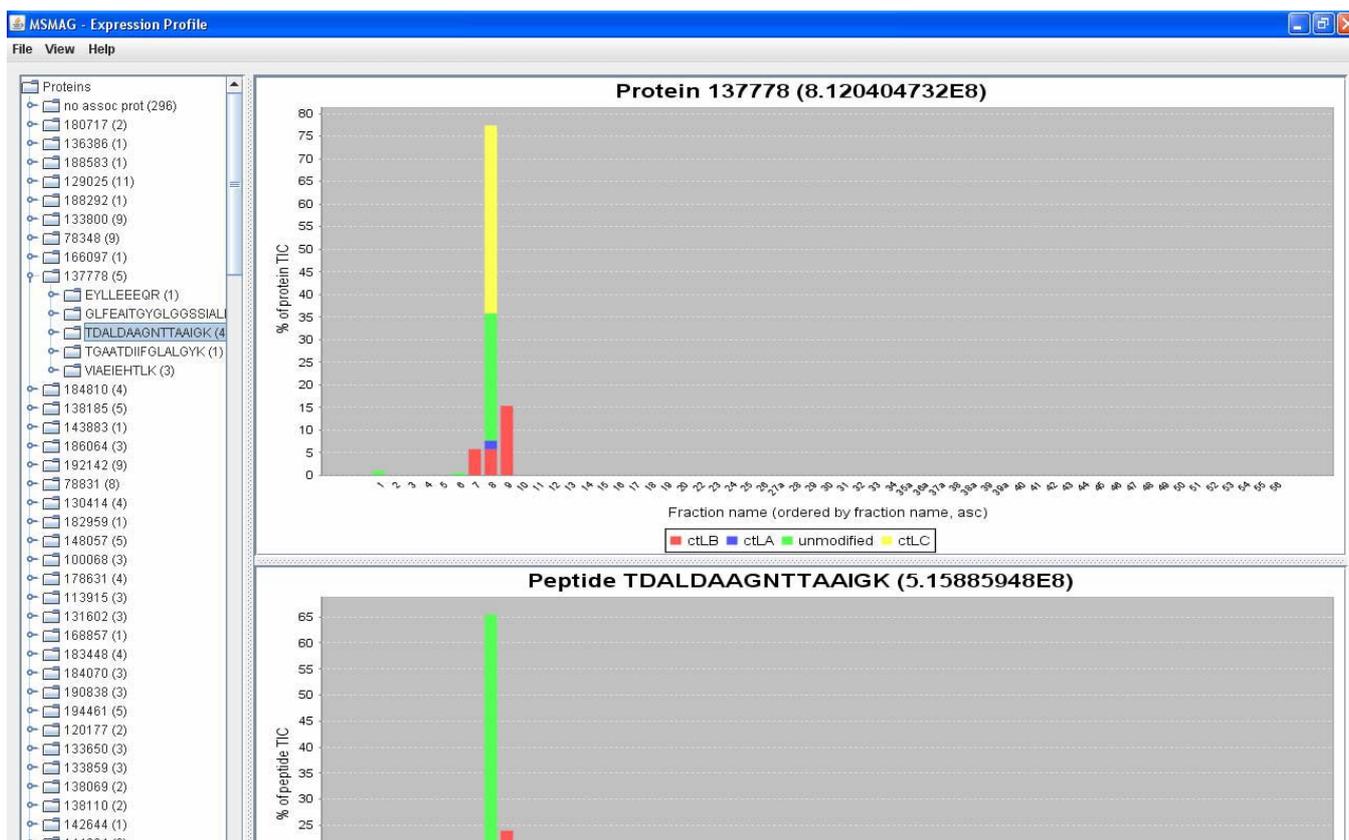


Figure 1: Screenshot of MSMAG displaying the protein and peptide distribution of the experiment A-LUin1. Modified peptides are coloured (see colour code below panes). The protein (137778) shows all possible modifications in fraction 8. However the selected peptide is only modified by ctLC in the fraction which means that the other peptides supporting the protein must carry the remaining modifications to account for the modifications to the protein in that fraction. Selecting them reveals their modifications and thus confirms this assumption.

pane shows the selected protein with the selected peptide highlighted. The selected peptide shows a similar distribution as the protein and is modified by ctLC in fraction 8 whereas it carries the modification ctLB in fractions 7 and 9. Selecting other peptides constituting the protein reveals the remaining modifications for the very heterogeneously modified fraction 8 (not shown). The first peptide carries the modification ctLA whereas the second peptide is modified by ctLC, but not quantitatively in fraction 8 (not shown). This assessment can help in identifying interesting fractions that warrant further investigation. While the motivation to perform the expression analysis in MSMAG was at least three fold (proteolytic control of proteins under different experimental conditions, finding samples for pooling, and detecting of proteotypic peptides). The program is now even more useful with the additional knowledge about the distribution of PTMs and/or labels that it provides. It is now possible to understand why a protein is found in different fractions since the differential PTMs of the protein in different fractions can be considered.

4.2 Re-establishing Peptide LC-Profiles

Although PTMs may change the chemical properties of the modified peptides and thus lead to different elution times in

LC-MS/MS experiments, such modifications merely present an additional mass attached to an amino acid for MS analysis. Thus for analyses such as those possible with the combination of 2DB and MSMAG altered chemical properties have no influence since the analyses are based on results from spectra identified by one or more database search algorithms. The modified chemical behaviour can, however, be investigated using the newly established LC-Profile display (Figure 2). In Figure 2 the LC profile has an unexpected distribution which is due to the way the data was generated and is not to be expected with experimental data. Using the LC-Profile view allows for a deeper analysis of the underlying raw data which was not possible with the prior version of MSMAG. Depending on the LC separation, peptides with modifications could be expected to elute in a certain order which can, for example, be verified using the LC-Profile display. The relative abundance of the peptide with different modifications can be derived from this re-established LC profile similar to using the expression profiling view but with information about elution time and the measured TIC.

4.3 Quantitation with Post Translational Modifications

Since PTMs and labels are modelled in the same way within 2DB, it is possible for MSMAG to retrieve information from

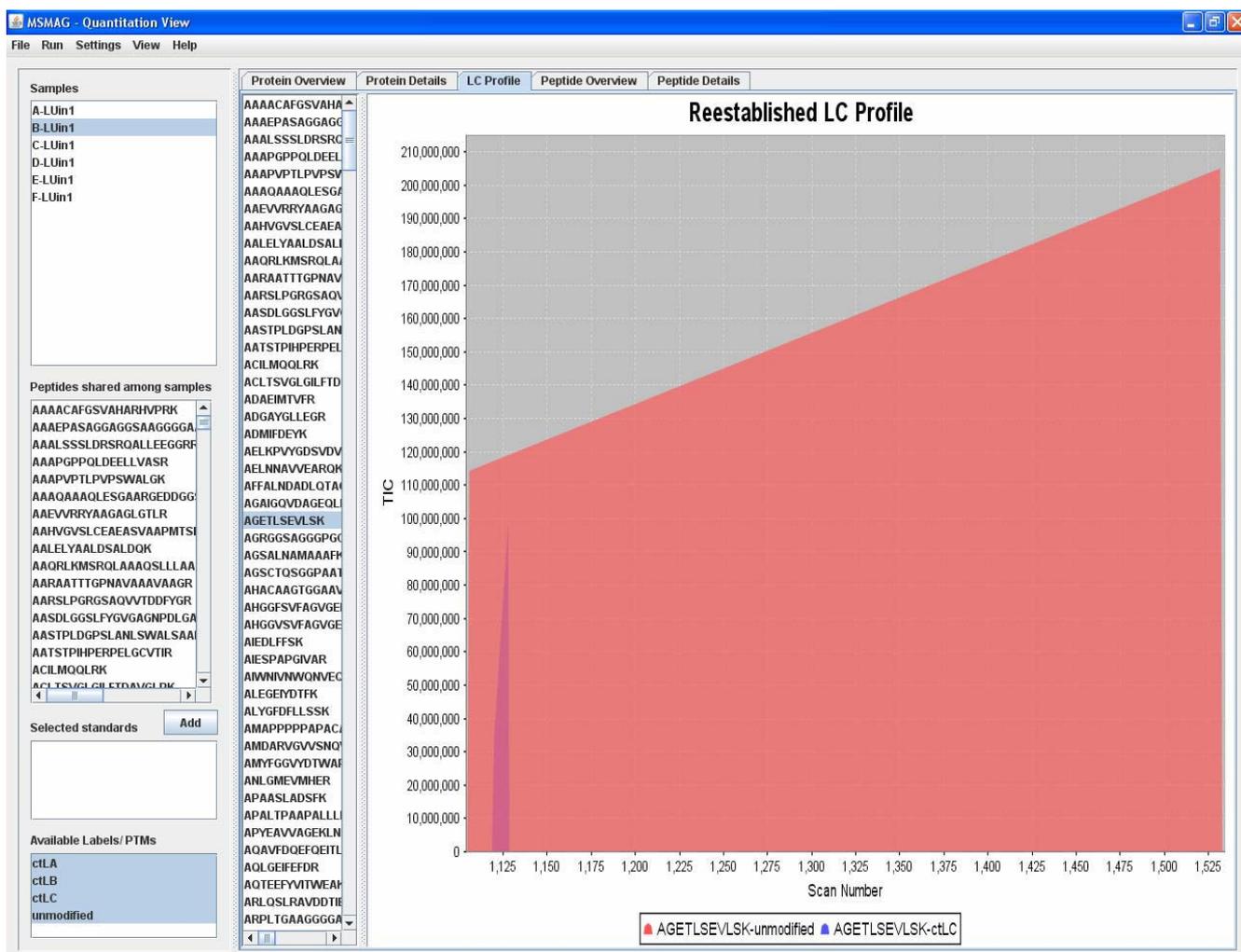


Figure 2: Re-established LC profile for the selected peptide AGETLSEVLSK. The peptide modified by ctLC (blue) elutes at about the same time but has a much lower area under the peak as compared to the unmodified peptide (red). With real experimental data, large shifts in elution time are to be expected but have not been included in the model for this data set.

2DB in a structured manner independent of whether a modifying PTM or label is present. This abstracts from the need to differentiate among diverse modifying agents. This simplification enables the user to perform quantitation based on labels or based on PTMs. Clearly, there must be a way to select which modification(s) ought to be used in a quantitation analysis. The new, modification aware, version of MSMAG therefore provides a box (Figure 4, bottom left corner) that lists all modifications that were mapped to peptides in the experiments that were selected for quantitation. From these modifications the ones that should be used in the analysis need to be selected (holding down the control key and clicking several modifications, allows for multiple selections). For performing label-free analysis no modifications need to be selected; for performing analyses involving labels, however, at least one of the modifications needs to be selected.

Once this selection is done, the quantitation method needs to be selected which can be chosen from Run – Quantitation

Mode. Available quantitation modes include the following possibilities:

- Label free comparison
- Labelled sample versus unlabelled sample
- Labelled sample versus labelled sample (similar to label free comparison)
- Compare among labels within a sample

Figure 3 shows the label-free comparison of the samples A-LUin1 to F-LUin1 which have a slight increasing trend from A to F. Label-free comparison was possible in the previous version of MSMAG and does not present a new development. Improvements are, for example, the comparison of labelled versus unlabelled samples and the comparison of samples containing both labelled and unlabelled versions of the analytes at the same time (Figure 4). In Figure 4 no trend from experiment A-LUin1 to F-LUin1 can be seen. The trend which was barely visible in the label-free analysis disappears when doing the comparison of ctLA labelled and unlabelled

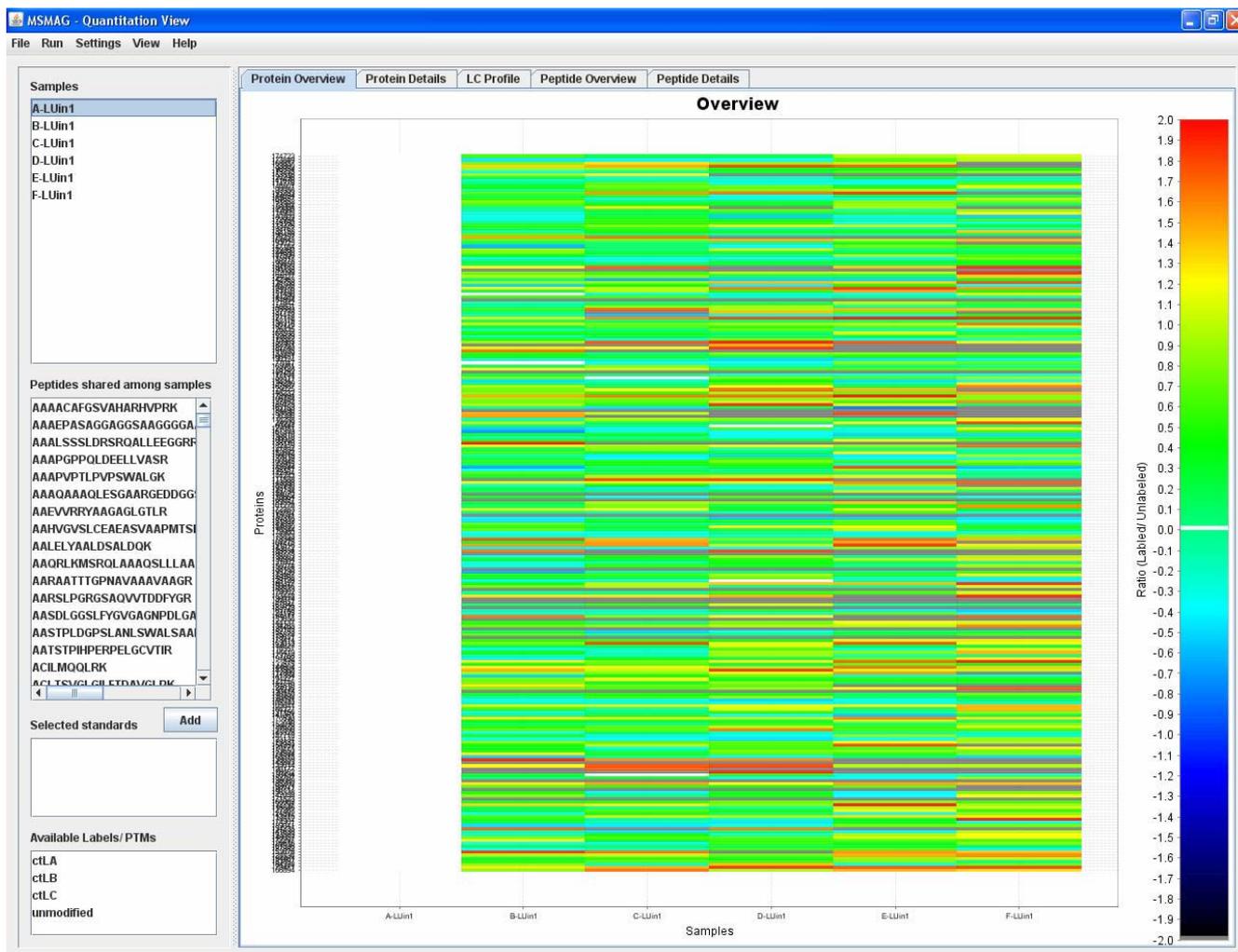


Figure 3: Label-free comparison of experiments A-LUin1 to F-LUin1 displaying all proteins. This can be regarded as an overview of the data. The heat map shows a slight trend from colder to warmer colours from the left-most sample to the right-most sample which is intentional by design of the data. A-LUin1 is white since it is compared with itself and thus should not display any up- or down-regulation.

proteins among experiments. This is to be expected from the set-up of the data. Modifications were added randomly to the data and thus the abundances of peptides, containing the label and those that do not, are random. Therefore, the ratio they produce is random and finally their ratio among experiments is also random. Another difference that is immediately visible is that the number of proteins is significantly reduced. This is due to the fact that only proteins that were found in all experiments are shown in Figure 4 whereas all proteins that were identified were shown in Figure 3. This can be selected via checking View – show All Proteins. Another reason why the number of proteins that can be assessed is significantly fewer is that only those proteins come into the first stage of analysis which have peptides carrying the modification ctLA and corresponding unlabelled peptides identifications in the same sample. In the second stage of analysis when experiments are compared, only those proteins which are present in all samples are further analyzed. Therefore, the number of proteins is significantly reduced. Allowing several labels will increase the number of proteins that can be analyzed but at

the cost of basing the analysis on a mixture of modified peptides that are compared to unmodified ones which are then compared among samples.

The heat maps seen in Figure 3 and Figure 4 are intended for visualization and quick assessment of the data. For a more thorough analysis, the data can be exported (File – Export), or copied from the table under the Protein Details tab. In addition to the protein heat map and the protein details table, similar information is available for the peptides supporting the proteins. A heat map for peptides is found under the tab Peptide Overview and Peptide Details contains the numerical information for the analysis and for each fraction separately. Options such as using peptide standards for normalization are not detailed in this work; they have been presented when MSMAG was first introduced [25].

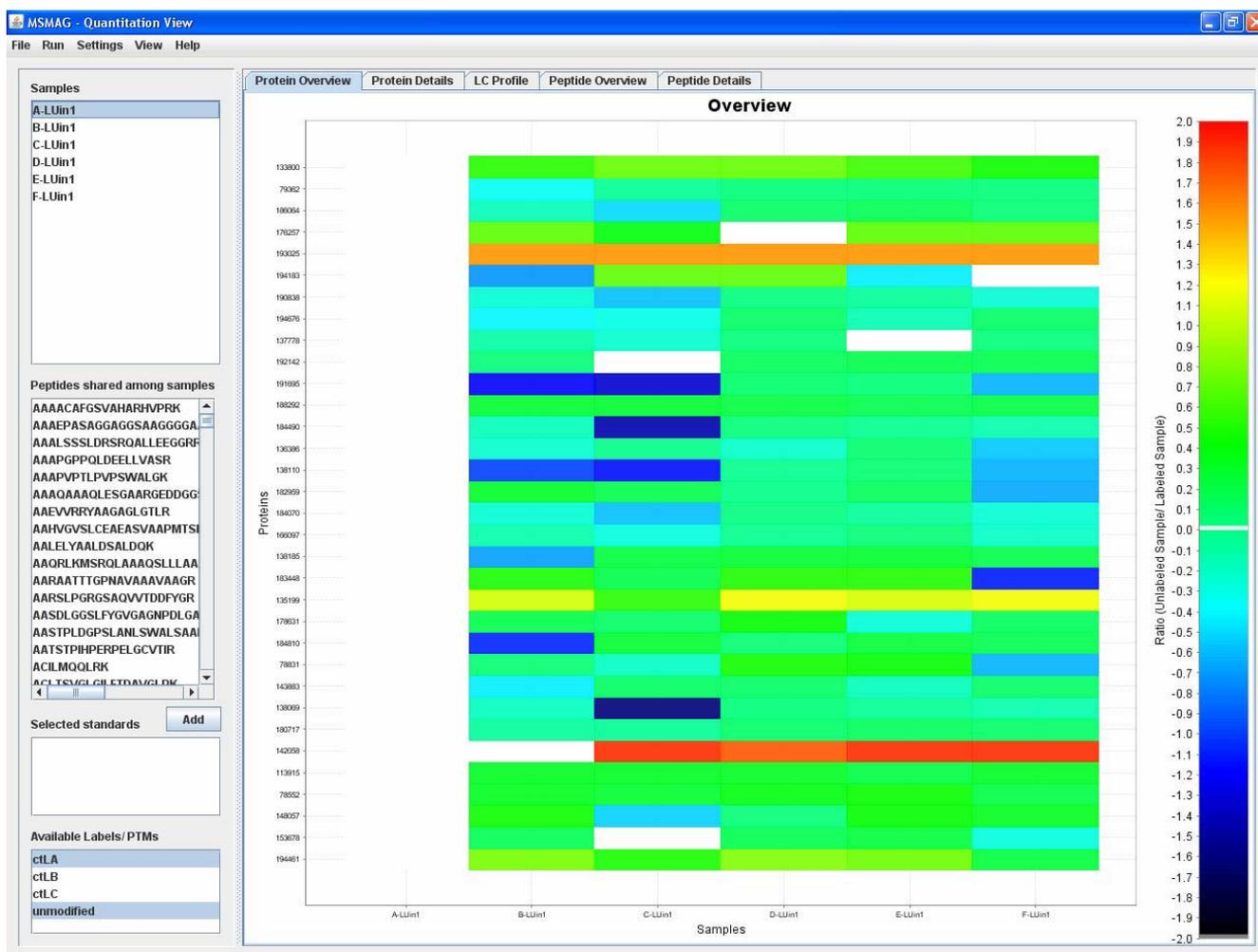


Figure 4: Comparison of Samples containing both, proteins with the ctLA modification and proteins without. The ratios are compared among the experiments A-LUin1 to F-LUin1. No trend can be seen which was expected since the modifications were added randomly.

5. CONCLUSION

A new version of MSMAG has been developed which

1. amends the expression profiling with information about PTMs
2. allows for new modes of quantitation
3. can perform quantitation with or without labels or PTMs
4. provides a new view of the re-established peptide LC profile with or without modifications

While only label-free quantitation was possible with the previous version of MSMAG, the new version has no such limitation and models both labels and PTMs as arbitrary mass shifts to peptides. Some new modes of quantitation were established to allow quantitation using different strategies. It is, however, envisioned that most commonly, label-free comparisons or a comparisons of labelled versus unlabeled sample will be used. The program is part of the 2DB distribution (<http://www.biolnk.com/2db>) and therefore free of charge. MSMAG's source code, as well as all other JAVA™ source

code that is part of 2DB, is currently not distributed along with the 2DB package but can be retrieved from the author via email.

6. OUTLOOK

Currently, no quantitative measure is derived from the area under the re-established LC profiles but in the future it will be integrated into the analysis.

Some analyses compare the TIC of labelled and unlabelled fragment ions within one spectrum [28]. This is currently not supported by MSMAG but will be enabled in the near future.

Many mathematical and data mining approaches have been developed for micro array studies. In essence, the mathematical problem is the same and thus it is aimed that such regularization, feature extraction, and clustering techniques will be incorporated into the software.

7. ACKNOWLEDGMENTS

I would like to thank Anne Frary for proof reading the article and for making it more intelligible. This work has been

supported by the Izmir Institute of Technology with the BAP grant 2009IYTE30.

REFERENCES

- [1] K.S. Nair, A. Jaleel, Y.W. Asmann, K.R. Short, and S. Raghavakaimal, "Proteomic research: potential opportunities for clinical and physiological investigators," *American Journal of Physiology. Endocrinology and Metabolism*, vol. 286, Jun. 2004, pp. E863-874.
- [2] E.S. Lander, L.M. Linton, B. Birren, C. Nusbaum, M.C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, et al. "Initial sequencing and analysis of the human genome," *Nature*, vol. 409, Feb. 2001, pp. 860-921.
- [3] P. Jungblut, B. Thiede, U. Zimny-Arndt, E.C. Müller, C. Scheler, B. Wittmann-Liebold, and A. Otto, "Resolution power of two-dimensional electrophoresis and identification of proteins from gels," *Electrophoresis*, vol. 17, May. 1996, pp. 839-847.
- [4] H. Schluter, R. Apweiler, H. Holzhutter, and P. Jungblut, "Finding one's way in proteomics: a protein species nomenclature," *Chemistry Central Journal*, vol. 3, 2009, p. 11.
- [5] S. Oeljeklaus, H.E. Meyer, and B. Warscheid, "Advancements in plant proteomics using quantitative mass spectrometry," *Journal of Proteomics*, vol. 72, Apr. 2009, pp. 545-554.
- [6] R. Aebersold and M. Mann, "Mass spectrometry-based proteomics," *Nature*, vol. 422, Mar. 2003, pp. 198-207.
- [7] I. Shadforth, D. Crowther, and C. Bessant, "Protein and peptide identification algorithms using MS for use in high-throughput, automated pipelines," *Proteomics*, vol. 5, Nov. 2005, pp. 4082-95.
- [8] D.C. Chamrad, G. Korting, K. Stuhler, H.E. Meyer, J. Klose, and M. Bluggel, "Evaluation of algorithms for protein identification from sequence databases using mass spectrometry data," *Proteomics*, vol. 4, Mar. 2004, pp. 619-28.
- [9] E. Kapp, F. Schutz, L. Connolly, J. Chakel, J. Meza, C. Miller, D. Fenyo, J. Eng, J. Adkins, G. Omenn, and R. Simpson, "An evaluation, comparison, and accurate benchmarking of several publicly available MS/MS search algorithms: sensitivity and specificity analysis," *Proteomics*, vol. 5, Aug. 2005, pp. 3475-90.
- [10] K. Garwood, T. McLaughlin, C. Garwood, S. Joens, N. Morrison, C.F. Taylor, K. Carroll, C. Evans, A.D. Whetton, S. Hart, D. Stead, Z. Yin, A.J. Brown, A. Hesketh, K. Chater, L. Hansson, M. Mewissen, P. Ghazal, J. Howard, K.S. Lilley, S.J. Gaskell, A. Brass, S.J. Hubbard, S.G. Oliver, and N.W. Paton, "PEDRo: a database for storing, searching and disseminating experimental proteomics data," *BMC Genomics*, vol. 5, Sep. 2004, p. 68.
- [11] J. Hartler, G. Thallinger, G. Stocker, A. Sturm, T. Burkard, E. Korner, R. Rader, A. Schmidt, K. Mechtler, and Z. Trajanoski, "MASPECTRAS: a platform for management and analysis of proteomics LC-MS/MS data," *BMC Bioinformatics*, vol. 8, 2007, p. 197.
- [12] P. Pernet, A. Bruneel, B. Baudin, and M. Vaubourdolle, "PHProteomicDB: a module for two-dimensional gel electrophoresis database creation on personal web sites," *Genomics Proteomics Bioinformatics*, vol. 4, May. 2006, pp. 134-6.
- [13] J. Allmer, S. Kuhlger, and M. Hippler, "2DB: a Proteomics database for storage, analysis, presentation, and retrieval of information from mass spectrometric experiments," *BMC Bioinformatics*, vol. 9, Jul. 2008, pp. 302-313.
- [14] S.P. Gygi, B. Rist, S.A. Gerber, F. Turecek, M.H. Gelb, and R. Aebersold, "Quantitative analysis of complex protein mixtures using isotope-coded affinity tags," *Nature Biotechnology*, vol. 17, Oct. 1999, pp. 994-999.
- [15] J. Krijgsveld, R.F. Ketting, T. Mahmoudi, J. Johansen, M. Artal-Sanz, C.P. Verrijzer, R.H.A. Plasterk, and A.J.R. Heck, "Metabolic labeling of *C. elegans* and *D. melanogaster* for quantitative proteomics," *Nature Biotechnology*, vol. 21, Aug. 2003, pp. 927-931.
- [16] F. Wang, M. Ye, J. Dong, R. Tian, L. Hu, G. Han, X. Jiang, R. Wu, and H. Zou, "Improvement of performance in label-free quantitative proteome analysis with monolithic electrospray ionization emitter," *J Sep Sci*, vol. 31, Aug. 2008, pp. 2589-97.
- [17] F. Yang, N. Jaitly, H. Jayachandran, Q. Luo, M.E. Monroe, X. Du, M.A. Gritsenko, R. Zhang, D.J. Anderson, S.O. Purvine, J.N. Adkins, R.J. Moore, H.M. Mottaz, S.J. Ding, M.S. Lipton, D.G. Camp, H.R. Udseth, R.D. Smith, and S. Rossie, "Applying a targeted label-free approach using LC-MS AMT tags to evaluate changes in protein phosphorylation following phosphatase inhibition," *J Proteome Res*, vol. 6, Nov. 2007, pp. 4489-97.
- [18] R.E. Higgs, M.D. Knierman, V. Gelfanova, J.P. Butler, and J.E. Hale, "Comprehensive label-free method for the relative quantification of proteins from biological samples," *Journal of Proteome Research*, vol. 4, Aug. 2005, pp. 1442-1450.
- [19] J. Silva, R. Denny, C. Dorschel, M. Gorenstein, I. Kass, G. Li, T. McKenna, M. Nold, K. Richardson, P. Young, and S. Geromanos, "Quantitative proteomic analysis by accurate mass retention time pairs," *Anal Chem*, vol. 77, Apr. 2005, pp. 2187-200.
- [20] J. Silva, R. Denny, C. Dorschel, M. Gorenstein, G. Li, K. Richardson, D. Wall, and S. Geromanos, "Simultaneous qualitative and quantitative analysis of the *Escherichia coli* proteome: a sweet tale," *Mol Cell Proteomics*, vol. 5, Apr. 2006, pp. 589-607.
- [21] M.P. Washburn, D. Wolters, and J.R. Yates, "Large-scale analysis of the yeast proteome by multidimensional protein identification technology," *Nature Biotechnology*, vol. 19, Mar. 2001, pp. 242-247.
- [22] H. Liu, R.G. Sadygov, and J.R. Yates, "A model for random sampling and estimation of relative protein abundance in shotgun proteomics," *Analytical Chemistry*, vol. 76, Jul. 2004, pp. 4193-4201.

- [23] A. Gilchrist, C.E. Au, J. Hiding, A.W. Bell, J. Fernandez-Rodriguez, S. Lesimple, H. Nagaya, L. Roy, S.J.C. Gosline, M. Hallett, J. Paiement, R.E. Kearney, T. Nilsson, and J.J.M. Bergeron, "Quantitative proteomics analysis of the secretory pathway," *Cell*, vol. 127, Dec. 2006, pp. 1265-1281.
- [24] J.M. Asara, H.R. Christofk, L.M. Freemark, and L.C. Cantley, "A label-free quantification method by MS/MS TIC compared to SILAC and spectral counting in a proteomics screen," *Proteomics*, vol. 8, Mar. 2008, pp. 994-999.
- [25] J. Allmer, "Label-free quantitation, an extension to 2DB," *Amino Acids*, Jul. 2009.
- [26] N. Allet, N. Barrillat, T. Baussant, C. Boiteau, P. Botti, L. Bougueleret, N. Budin, D. Canet, S. Carraud, D. Chiappe, N. Christmann, J. Colinge, I. Cusin, N. Dafflon, B. Depresle, I. Fasso, P. Frauchiger, H. Gaertner, A. Gleizes, E. Gonzalez-Couto, C. Jeandenans, A. Karmime, T. Kowall, S. Lagache, E. Mahé, A. Masselot, H. Mattou, M. Moniatte, A. Niknejad, M. Paolini, F. Perret, N. Pinaud, F. Ranno, S. Raimondi, S. Reffas, P. Regamey, P. Rey, P. Rodriguez-Tomé, K. Rose, G. Rossellat, C. Saudrais, C. Schmidt, M. Villain, and C. Zwahlen, "In vitro and in silico processes to identify differentially expressed proteins," *Proteomics*, vol. 4, Aug. 2004, pp. 2333-2351.
- [27] O. Schulz-Trieglaff, N. Pfeifer, C. Gropl, O. Kohlbacher, and K. Reinert, "LC-MSsim - a simulation software for liquid chromatography mass spectrometry data," *BMC Bioinformatics*, vol. 9, 2008, p. 423.
- [28] B. Naumann, A. Busch, J. Allmer, E. Ostendorf, M. Zeller, H. Kirchhoff, and M. Hippler, "Comparative quantitative proteomics to investigate the remodeling of bioenergetic pathways under iron deficiency in *Chlamydomonas reinhardtii*," *Proteomics*, vol. 7, Nov. 2007, pp. 3964-79.