curacy has inspired the development of new quantitative approaches [2]. In particular, label free hybrid identity/pattern-based approaches have gained popularity in the biomarker development field [3]. Such approaches use pattern recognition algorithms to *ex post facto* assign the identity of LC-MS peaks against databases of peptide sequence, mass, and retention time built from multiple experiments. Here we employ the accurate mass and time (AMT) strategy to identify and quantify peptides/proteins from unfractionated CSF samples using msInspect platform [4]. Our analysis workflow extends and combines existing open-source platforms for LC−MS/MS (Trans Proteomic Pipeline) and LC−MS (msInspect) data analysis. A peptide accurate mass and LC elution time AMT database was initially generated using MS/MS following extensive multidimensional LC separations to provide the basis for subsequent peptide identifications. Our CSF AMT database contains >2,000 entries from a total number of ~7,500 identified peptides, which translates into 558 confident CSF protein identifications (Figure 2). In conclusion, hybrid identity/pattern-based approaches are amenable for high throughput quantitative proteome analyses in biomarker discovery pipelines.

References

[1] Zougman, A et al. Integrated Analysis of the Cerebrospinal Fluid Peptidome and Proteome. *J Proteome Res* 2007; 7: 386-399.

[2] Mueller, LK et al. An Assessment of Software Solutions for the Analysis of Mass Spectrometry Based Quantitative Proteomics Data. *J Proteome Res* 2008; 7: 51-61.

[3] Jaffe, J et al. PEPPeR, a Platform for Experimental Proteomic Pattern Recognition. *Mol Cell Proteomics* 2006; 5:1927-1941.

[4] May, D et al. A Platform for Accurate Mass and Time Analyses of Mass Spectrometry Data. *J Proteome Res* 2007; 6: 2685-2694.

A comparison of quantitative proteomics methodologies on a differential experiment on test samples

Joan Josep Bech-Serra, Núria Colomé, Marta Monge, Francesc Canals

Laboratori de Proteòmica, Programa de Recerca en Oncologia Mèdica. Institut de Recerca. Hospital Vall d'Hebron, Barcelona

The field of proteomics includes a wide number of technologies aimed to the analysis of large number of proteins, representing ideally the entire proteome, in the same experiment. In its early stages , mass spectrometry-based proteomics was successfully used to the qualitative characterization of complex mixtures of proteins, leaving the quantitative aspects in a secondary role, essentially because of the lack of suitable technologies for quantitative analysis of such complex mixtures. However, in the last years, a wide number of strategies have been developed in order to give to the proteomics field robust tools to obtain reliable quantitative data. Those strategies include both 2D-gel based and non-based methodologies. In the first group, 2D-DIGE methodology is based on the labeling with different fluorochromes, enabling the simultaneous separation of different samples in the same 2D-gel, which allows to overcome reproducibility problems inherent to the 2D procedure and internal standardization of spot volume measurements, providing a robust quantitative comparison technique. Alternatively, a number of gel-free quantitative techniques can be used. Some of them are based on non-isobaric isotope labeling, such as ICAT, ICPL or SILAC labeling methodologies, the quantitative data being obtained from the from the MS data, by integrating extracted ion chromatograms of the heavy-light peptide pair masses. A second group which includes iTRAQ or TMT reagents is based on isobaric labeling that introduce different reporter fragments in the derivatized peptides. Quantitative data is in this case obtained at the MS-MS level, from the relative intensities of the reporter fragment ions. Finally, a third group of label-free methods uses different approaches for quantitative comparison of LC-MS data.

In order to evaluate and compare the performance of different quantitative proteomics methods, we have prepared a set of two test samples to be compared quantitatively. The samples consist on a matrix of cytoplasmatic E. Coli proteins, of medium complexity (around 100 proteins), to which four standard mammalian proteins have been spiked in different amounts, ranging from the fmol to the pmol level per microgram of total protein.. The ratios of the spiked proteins between the two samples have been chosen to range form 1.5:1 to 5:1.

These samples have been analyzed by different methodologies (Figure 1):

- 2D-DIGE, using four technical replicas of each sample on a 4 2D-gel experiment.

- Labeling with ICPL reagents at the protein level, followed by tryptic digestion and analysis by LC-MS. This is the standard procedure for this methodology, but has the drawback that only lysine containing peptides which have been modified by the reagent can be used for quantitation. This results often in a large part of the identified proteins lacking quantitative information.

- Labeling with ICPL at peptide level. In order to overcome this problem, we have explored an alternative procedure introducing the ICPL labeling step after tryptic digestion of the protein mixtures. This should result in a more comprehensive quantitative information, since all tryptic peptides will be derivatized at the N-terminus. The weakness of this strategy is that digestion of the two samples has to be run independently, which can introduce technical bias. To minimize the problem of a possible imbalance, both labeling schemes are run in parallel on sample aliquots. This way the balance correction derived from the protein level labeling experiment can be used to normalize the data from the second experiment.

ICPL labeled samples have been analyzed in four independent LC-MS runs in order to assess reproducibility.

- iTRAQ Labeling. The samples have been analyzed using an 8-plex iTRAQ reagent, using four different reporter masses for each of the samples, and analyzed by LC-MALDI.

- Label-free quantitation. Four LC-MS runs of a tryptic digest of each sample have used to quantitatively compare the two samples using Progenesis LC-MS software for alignment and statistical analysis of the LC-MS maps.

The results of all the quantitative proteomics methods used have been evaluated in terms of accuracy and reproducibility, and their performance and robustness is discussed.

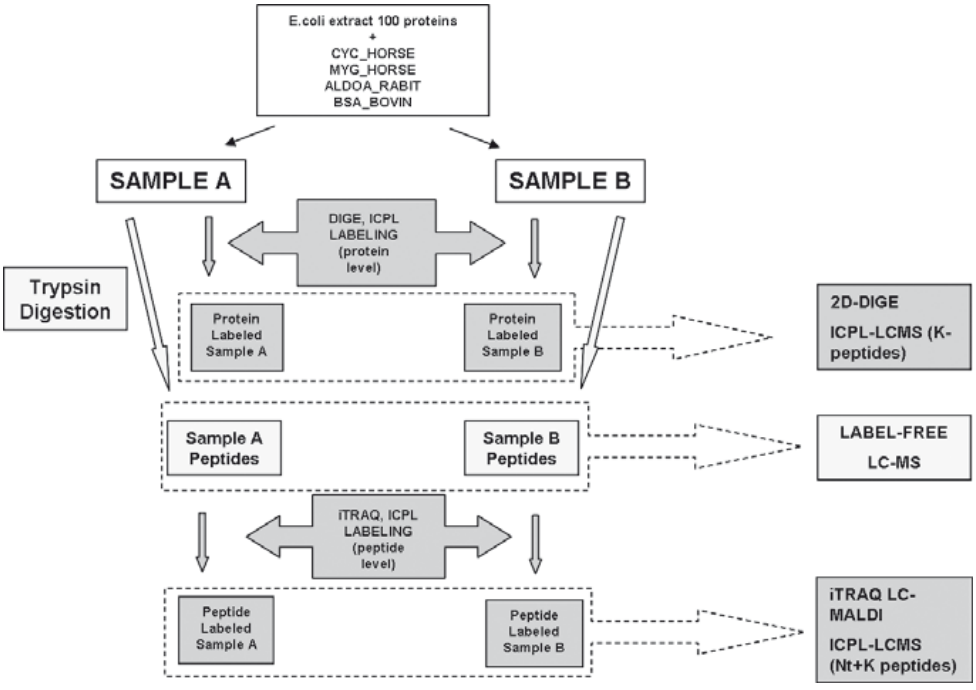

Figure 1. *Scheme of the methods compared for the relative quantification of two test samples.*