PL10

Biomarker Discovery: From Hypervariable Data to Biomarker Candidates

Rainer Bischoff • *Peter Horvatovich* • *Natalia Govorukhina* • *Christin Christin* • *Berend Hoekman* • *Therese Rosenling, Nicolas Abello* • *Jos Hermans* • *University of Groningen* • *Frank Suits (IBM TJ Watson Research Center, Yorktown Heights, NY, USA)* • *Age Smilds, Huub Hoefsloot (University of Amsterdam)* • *Ate van der Zee (University Medical Center Groningen)* • *Theo Luider (Erasmus Medical Center Rotterdam)*

The biomarker discovery phase is characterized by the analysis of a limited number of well-classified samples. 'Omics' analyses generally lead to a large number of signals relative to a limited number of analyses. The main goal of the discovery phase is to find statistically significant differences in protein profiles between groups of pre-classified samples. The fact that there are many more measured proteins relative to the number of samples requires that the obtained statistical models be carefully validated to avoid that observed differences are due to chance. In this presentation I will show developments in the area of time alignment between LC-MS data sets (1-3) and statistical data analysis based on designed and real-life data sets.

Pre-analytical factors are particularly critical when it comes to proteomics analyses. Studies from the literature give guidelines on how to approach this complex and largely unsolved problem but each type of biological sample requires specific studies. Based on our work on serum and cerebrospinal fluid, I will present results highlighting stability issues related to coagulation time, time laps between sampling and freezing as well as with respect to freeze-thaw cycles (4;5). Applying the concept of 'fractional factorial design' to this complex, multi-parameter problem will be discussed.

While most of our present studies have been obtained by 'label-free' LC-MS, there is merit to using comparative, stable isotope labeling to correct variations in the quantitative response of the mass spectrometer and more generally of the analytical method as such. We have recently developed a novel stable isotope labeling approach that allows for extensive multiplexing as well as for an enhanced response in electrospray ionization, notably for low molecular weight, hydrophilic metabolites (6). Further data about this approach will be shown and a future outlook of its possibilities will be given.

Biomarker discovery may lead to so-called "biomarker candidates" that require further validation. There are a number of ways of biomarker validation, many of which rely on immunochemical techniques such as immunohistochemistry or ELISA. More recently the combination of liquid chromatography with mass spectrometry using triple quadrupole mass analyzers in the Multiple Reaction Monitoring (MRM) mode is gaining acceptance. This approach has the advantage of being able to quantify multiple biomarker candidates in a single analysis without the need for well-characterized antibodies. Initial data from the area of matrix-metalloproteases and other biomarker candidates will be presented.

Finally, discovering and validating biomarker candidates in disease-related tissue is an area of great interest. In this context I will present results from our work with biopsies from cervical cancer patients that were analyzed by comparative proteomics after laser microdissection.

1. Christin C, Hoefsloot HC, Smilde AK, Suits F, Bischoff R, Horvatovich P. Time Alignment Algorithms based on Selected Mass Traces for Complex LC-MS Data. J Proteome Res 2010;9:1483-95.

2. Christin C, Smilde AK, Hoefsloot HCJ, Suits F, Bischoff R, Horvatovich PL. Optimized Time Alignment Algorithm for LC-MS Data: Correlation Optimized Warping Using Component Detection Algorithm-Selected Mass Chromatograms. Analytical Chemistry 2008;80:7012-21.

3. Suits F, Lepre J, Du P, Bischoff R, Horvatovich P. Two-Dimensional Method for Time Aligning Liquid Chromatography-Mass Spectrometry Data. Anal Chem 2008;80:3095-104.

4. Govorukhina NI, de Vries M, Reijmers TH, Horvatovich P, van der Zee AGJ, Bischoff R. Influence of clotting time on the protein composition of serum samples based on LC-MS data. Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences 2009;877:1281-91.

5. Rosenling T, Slim CL, Christin C, Coulier L, Bosman J, Shi S et al. The effect of pre-analytical factors on stability of the proteome and selected metabolites in cerebrospinal fluid (CSF). J Proteome Res 2009;8:5511-22.

6. Abello N, Geurink PP, Toorn MV, Oosterhout AJ, Lugtenburg J, Marel GA et al. Poly(ethylene glycol)-Based Stable Isotope Labeling Reagents for the Quantitative Analysis of Low Molecular Weight Metabolites by LC-MS. Anal Chem 2008;80:9171-80.