

Managing Proteomics Data from Generation and Data Warehousing to Central Data Repository

H. Thiele, J. Glandorf, P. Hufnagel

Bioinformatics, Bruker Daltonik GmbH, Bremen, Germany

Introduction

The tremendous amount of data from today's expression proteomics requires a database solution with data-warehousing and data-mining capabilities. ProteinScape provides a bioinformatics platform for in-house proteome studies as well as for large scale approaches. The growing requirement for protein pre-fractionation to obtain more precise quantitative protein information is uniquely addressed in ProteinScape. Entire workflows of pre-fractionation, detailed LC/MS/MS separation and post-processing with bioinformatics tools are merged and can be easily controlled and reviewed.

Methods

ProteinScape is a bioinformatics platform addressing the requirements for biomarker discovery, protein identification and quantification. It supports various discovery workflows through a flexible analyte hierarchy, various database search engines and quantification approaches including a label-free strategy. All current label chemistries for protein quantification are fully supported (ICPL, SILAC, iTRAQ, ICAT, and C-term ¹⁸O/¹⁶O-C-term labeling). The support includes multiplexed quantification (e.g., ICPL triplex, iTRAQ or SILAC 4plex). It enables the use of isobaric or non-isobaric label chemistries and it permits the targeted analysis of proteins in complex mixtures. Interactive validation of protein quantification based on raw LC/MS data is now simple and straight forward.

Proceedings of The Joint **2nd Pacific Rim International Conference on Protein Science** and **4th Asian-Oceania Human Proteome Organization**, Cairns- Australia, 22-26
June 2008

Results

ProteinScape has a number of dedicated viewers that permit the evaluation and validation on each level of proteomics experiments. BioTools integrates with ProteinScape for advanced sequence validation, PTM discovery, de novo sequencing and MS-BLAST searches for full structure elucidation functionalities. Integrated quantification workflows that utilize labeling and label-free technologies require greatly reduced analysis and validation time.

The European Commission-funded ProDaC consortium (<http://www.fp6-prodac.eu/>) will finalize data storage and documentation standards, implement conversion tools and establish standardized submission pipelines into a central data repository. This contains export from local LIMS systems like ProteinScape to standard file formats or direct upload into PRIDE. With respect to this a tool is already implemented to upload ProteinScape data sets into PRIDE.

Proceedings of The Joint 2nd Pacific Rim International
Conference on Protein Science and 4th Asian-Oceania
Human Proteome Organization, Cairns- Australia, 22-26
June 2008