



ISOQuant - an integrated bioinformatics pipeline for evaluation and reporting of data independent (LC-MS^E) label-free quantitative proteomics data

Jörg Kuharev, Pedro Navarro, Ute Distler, Hansjörg Schild, Stefan Tenzer

Institute for Immunology, University Medical Center of the Johannes Gutenberg University Mainz, Mainz, Germany

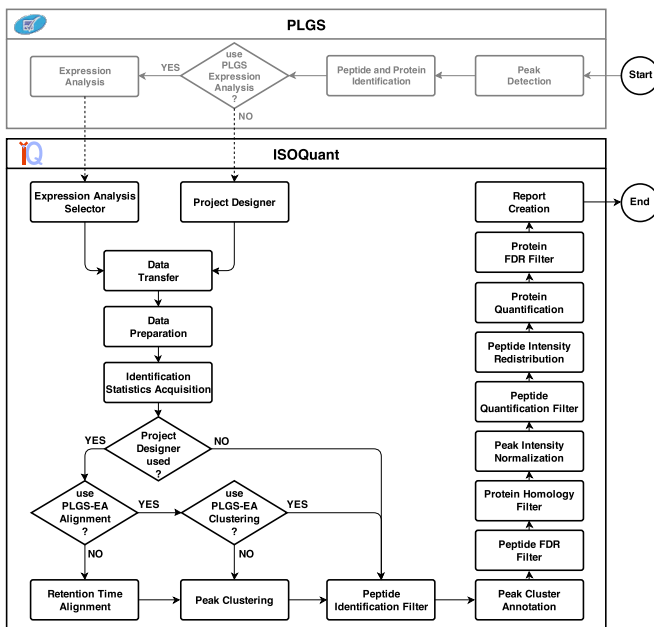
Objectives

Recent generation high-resolution mass spectrometers supporting ion-mobility separation (IMS) combined with advanced acquisition schemes such as UDMSE enable the reliable identification and quantification of over 4,000 proteins in a single nanoUPLC-MS run [1]. Large amounts of generated data induce new requirements for the data processing software. MS^E/HDMSE/UDMSE data require processing by proprietary vendor software ProteinLynx GlobalSERVER (PLGS) [2]. Preprocessed data require further analysis to confidently identify and quantify regulated proteins, typically done by stepwise application of different dedicated algorithms and exporting to common data formats for result interpretation and sharing with the community. We present ISOQuant, an integrated open source software pipeline for in-depth evaluation and standardized reporting of MS^E/HDMSE/UDMSE based label-free proteomics data [1].

Methods

Tryptic HeLa cell lysate was analyzed in MS^E/HDMSE and UDMSE acquisition modes on a Waters Synapt G2-S mass spectrometer using 90 and 180 min gradients. PLGS (version >2.5, Waters Corporation) was used for raw data processing and for peptide and protein identification. ISOQuant was used for downstream analysis.

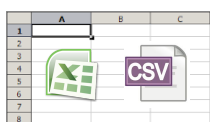
Data Analysis Workflow



Reports

ISOQuant exports analysis results to a set of uniform report formats. Generated MS Excel and CSV spreadsheets, as well as HTML reports allow to easily browse and share analysis results and display basic data quality metrics. Additionally, HUPO PSI standardized mzIdentML files are generated for an easy submission to PRIDE Archive - proteomics data repository [3][4].

Spreadsheets



HTML



mzIdentML



Availability

Please visit www.isoquant.net to obtain current version of ISOQuant. Installation packages are available for Windows, Mac OSX and UNIX/Linux operating systems.



Bibliography

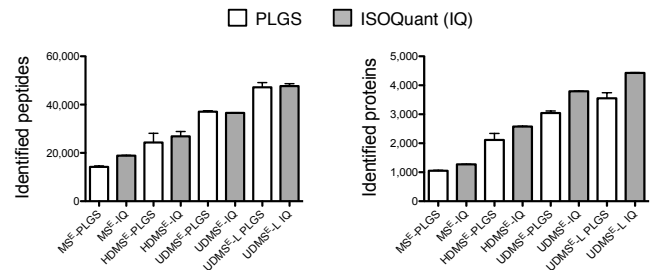
- Distler, U., Kuharev, J., Navarro, P., Levin, Y., Schild, H., & Tenzer, S. (2013). DTTI time-specific collision energies enable deep-coverage data-independent acquisition proteomics. *Nature Methods*, advance online publication. doi:10.1038/nmeth.2767
- Waters: ProteinLynx Global SERVER (PLGS), at <http://www.waters.com/waters/nav.htm?cid=513821>
- Vizcaino, J. A., Côté, R. G., Csordas, A., Dhanes, J. A., Fabregat, A., Foster, J. M., ... Hermjakob, H. (2013). The Proteomics Identifications (PRIDE) database and associated tools: status in 2013. *Nucleic Acids Research*, 41(D1), D1063-D1069. doi:10.1093/nar/gks1262
- Reisinger, F., Krishna, R., Ghali, F., Pflös, D., Hermjakob, H., Antonio Vizcaino, J., & Jones, A. R. (2012). mzIdentML API: A Java interface to the mzIdentML standard for peptide and protein identification data. *PROTEOMICS*, 12(6), 790-794. doi:10.1002/pmic.2011100277

Acknowledgements

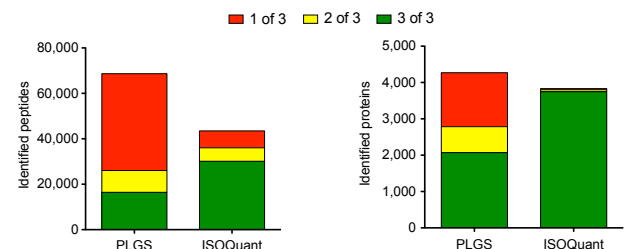
This work was supported by the Deutsche Forschungsgemeinschaft (SFB490, Z3) and the Research Center for Immunotherapy (FZI) of the UMC of the Johannes Gutenberg University Mainz. We are thankful to Yasset Pérez-Riverol for supporting with the mzIdentML specifications.

Results

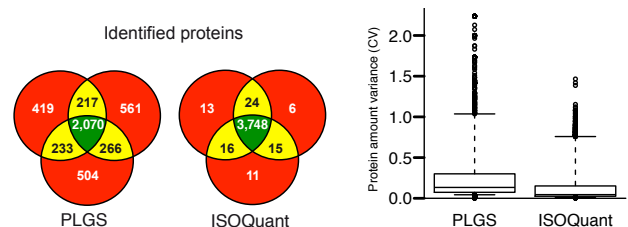
Identified peptides and proteins from tryptic digest of HeLa lysate analyzed in MS^E, HDMSE and UDMSE acquisition modes before/after ISOQuant analysis at 1% FDR.



ISOQuant reduces the overall number of identifications but increases the number of identified peptides per technical replicate as well as the overlap between technical replicates on peptide and protein level.



The increased overlap among technical replicates of identified proteins (at 1% FDR) for UDMSE goes along with the reduction of the technical replicate variance of the calculated protein amounts.



We produced a model dataset to evaluate the performance of label-free protein quantification. Tryptic digests of HeLa, yeast and E.coli proteomes were combined in exactly known ratios and analyzed in multiple technical replicates.

