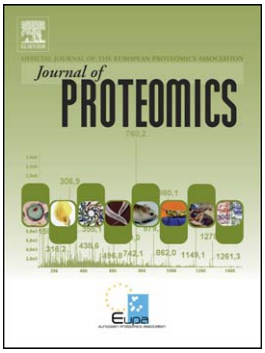

Computational and statistical methods for high-throughput analysis of post-translational modifications of proteins

Veit Schwämmle, Thiago Verano-Braga, Peter Roepstorff

PII: S1874-3919(15)30071-3
DOI: doi: 10.1016/j.jprot.2015.07.016
Reference: JPROT 2221

To appear in: *Journal of Proteomics*

Received date: 4 May 2015
Revised date: 9 July 2015
Accepted date: 13 July 2015

Please cite this article as: Schwämmle Veit, Verano-Braga Thiago, Roepstorff Peter, Computational and statistical methods for high-throughput analysis of post-translational modifications of proteins, *Journal of Proteomics* (2015), doi: 10.1016/j.jprot.2015.07.016

Computational and statistical methods for high-throughput analysis of post-translational modifications of proteins

Veit Schwämmle[1], Thiago Verano-Braga[1,2] and Peter Roepstorff[1]*

[1]Protein Research Group, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark

[2]National Institute of Science and Technology in Nanobiopharmaceutics (INCT-Nanobiofar), Department of Physiology and Biophysics, Federal University of Minas Gerais, Belo Horizonte, Minas Gerais, Brazil

[*]Corresponding author: veits@bmb.sdu.dk

Abstract

The investigation of post-translational modifications (PTMs) represents one of the main research focuses for the study of protein function and cell signaling. Mass spectrometry instrumentation with increasing sensitivity, improved protocols for PTM enrichment and recently established pipelines for high-throughput experiments allow large-scale identification and quantification of several PTM types. This review addresses the concurrently emerging challenges for the computational analysis of the resulting data and presents PTM-centered approaches for spectra identification, statistical analysis, multivariate analysis and data interpretation. We furthermore discuss the potential of future developments that will help to gain deep insight into the PTM-ome and its biological role in cells.

Introduction

Molecular complexity increases enormously from genome to proteome; while human cells contain around 20,000 protein coding genes [1], 1,000,000 proteoforms are estimated to exist and posttranslational modifications (PTMs) play an important role on such diversity [2]. PTMs change the physicochemical properties of target proteins, leading to structural changes and therefore affecting their localization, activity and binding partners. Virtually all cellular processes are modulated by PTMs and their dysregulation is associated with several pathologies, highlighting the relevance of PTM studies. At present, there are 1,001 protein