Automated workflow composition in mass spectrometry-based proteomics

Numerous software utilities operating on mass spectrometry (MS) data are described in the literature and provide specific operations as building blocks for the assembly of on-purpose workflows. Working out which tools and combinations are applicable or optimal in practice is often hard. Thus researchers face difficulties in selecting practical and effective data analysis pipelines for a specific experimental design. Results: We provide a toolkit to support researchers in identifying, comparing and benchmarking multiple workflows from individual bioinformatics tools. Automated workflow composition is enabled by the tools’ semantic annotation in terms of the EDAM ontology. To demonstrate the practical use of our framework, we created and evaluated a number of logically and semantically equivalent workflows for four use cases representing frequent tasks in MS-based proteomics. Indeed we found that the results computed by the workflows could vary considerably, emphasizing the benefits of a framework that facilitates their systematic exploration.

General information
Publication status: Published
Organisations: IT Service, Leiden University, Utrecht University, University of Southern Denmark
Contributors: Palmblad, M., Lamprecht, A. L., Ison, J., Schwämmle, V.
Pages: 656-664
Publication date: 15 Feb 2019
Peer-reviewed: Yes

Publication information
Journal: Bioinformatics
Volume: 35
Issue number: 4
ISSN (Print): 1367-4803
Ratings:
Web of Science (2019): Indexed yes
Original language: English
Electronic versions:
HKKR_bty646.pdf
DOIs:
10.1093/bioinformatics/bty646
Source: Scopus
Source-ID: 85061363387
Research output: Contribution to journal › Journal article – Annual report year: 2019 › Research › peer-review