

Description of the UPS{pep/prot}2 datasets

Samuel Wieczorek*, Florence Combes*, Claire Ramus, Quentin Gai Gianetto,
Yohann Coute, Myriam Ferro, Christophe Bruley and Thomas Burger*

June 18, 2016

This dataset results from a controlled relative quantification proteomics experiment where Sigma UPS1 human proteins were spiked in a similar yeast lysate in 2 different concentrations (with a ratio of 2). As a consequence, it can be used to benchmark the quality of a statistical analysis: in the ideal case, after the differential analysis, only and all the human proteins should be selected.

1 Preparation and digestion of proteins and peptides

UPS1 was spiked in yeast extract in 8M urea, with a concentration of 10fmol for one condition, and of 5fmol for the other. The solution was diluted twice before 1h incubation with 10mM DTT at 37°C, 45min incubation with 55mM iodoacetamide in the dark and at room temperature. Proteins were then submitted to LysC digestion for 4 hours at 37°C. After dilution, proteins were submitted to overnight digestion with modified trypsin (Promega, sequencing grade) at 37°C. Digestion was stopped through acidification with trifluoroacetic acid. Resulting peptides were desalted using C18 MacroSpin columns (Harvard Apparatus).

2 Nano-LC-MS/MS analyses

Dried peptides were resuspended in 5% acetonitrile and 0.1% trifluoroacetic acid and analyzed by online nanoLC-MS/MS (Ultimate 3000, Dionex and Q-Exactive Plus, Thermo Scientific). Peptides were sampled on a 300 μm x 5 mm PepMap C18 precolumn and separated on a 75 μm x 250 mm C18 column (PepMap, Dionex). The nanoLC method consisted of a 120-minutes gradient at a flow rate of 300 nl/min, ranging from 5% to 37% acetonitrile in 0.1% formic acid for 114 min, before reaching 72% acetonitrile in 0.1% formic acid for the last 6 min. MS and MS/MS data were acquired using Xcalibur

*firstname.lastname@cea.fr

(Thermo Scientific). Spray voltage was set at 1.6 kV; heated capillary was adjusted to 250°C. Survey full-scan MS spectra ($m/z = 400\text{--}1600$) were acquired in the Orbitrap with a resolution of 70,000 after accumulation of 106 ions (maximum filling time 200 ms). MS/MS spectra were acquired for up to the ten most intense ions from the MS scan after higher energy collisional dissociation (accumulation of 105 ions, maximum filling time of 50 ms, resolution of 17,500).

3 Computational analyses

RAW files were processed using MaxQuant version 1.5.1.2. Spectra were searched against the Uniprot database (*Saccharomyces cerevisiae* (strain ATCC 204508 / S288c) taxonomy, June 2015 version), the UPS database and the frequently observed contaminants database embedded in MaxQuant. Trypsin was chosen as the enzyme and 2 missed cleavages were allowed. Precursor mass error tolerances were set respectively at 20 ppm and 4.5 ppm for first and main searches. Fragment mass error tolerance was set to 0.5 Da. Peptide modifications allowed during the search were: carbamidomethylation (C, fixed), acetyl (Protein N-ter, variable) and oxidation (M, variable). Minimum peptide length was set to 7 amino acids. Minimum number of peptides, razor + unique peptides and unique peptides were all set to 1. Maximum false discovery rates -calculated by employing a reverse database strategy - were set to 0.01 at peptide and protein levels. Raw intensity IBAQ and LFQ values were calculated using MaxQuant as previously described (2), from MS intensity of unique peptides.

4 Datasets

The package DAPARdata contains two datasets:

- USPpep2 which contains the quantitative data at the peptide level,
- USPprot2 which contains the quantitative data at the protein level.

Both datasets are available in CSV and MSnSet formats.