Description of DDA:Cox2014 - Cox et al., 2014[1]

Data processing and quantification

The 8 raw files beginning with "20130510_EXQ1_IgPa_QC_UPS" (the "Dynamic Range Benchmark Dataset") for PRIDE PXD000279 were downloaded. In Skyline-daily (version 3.5.1.9327), using the results (msms.txt) produced by MaxQuant (v1.5.0.0) with the Andromeda search engine, as described in Cox2014_DDA_MaxQuant_description.pdf in RMSV000000255.1(MassIVE Reanalysis ID), a spectral library was built with 0.99 cut-off score and ambiguous spectrum matches not included. A new document was created with default settings plus the new library, Oxidation (M), Acetyl (N-Term), Acetyl (S), and Acetyl (T), precursors charge states 2-4, ion type "p" and chromatogram extraction of 3 isotope peaks was set to use 15 ppm mass tolerance applying to all centroided MS1 spectra within 5 minutes of MS/MS IDs.

The 3 (human, E. coli, contaminants) proteome sequence (FASTA) files mentioned in the original paper methods were imported into the new document, with repeated peptides and empty proteins removed, resulting in 2269 proteins, 20,091 peptides, and 25,920 precursors.

The 8 runs were then imported into the template for chromatogram extraction, retention time alignment and automatic peak picking. The MSstats report was exported for further analysis.

Statistical analysis

Proteins that have human protein ids, but were not spiked in yeast background, were removed. The MSstats report from Skyline and annotation for condition and biological replicate were used for SkylinetoMSstatsFormat function in MSstats v3.10.6, in order to pre-process before statistical analysis. Truncated peaks are replaced with NA missing value. Only unique peptides were used and shared peptides are removed (useUniquePeptide=TRUE). Features that have less than 3 measurements across MS runs were removed (fewMeasurements="remove"). Proteins which have only one PSM were removed (removeProtein_with1Feature=TRUE). Three isotopic peaks per feature and run were summed. Medians of log2 intensities across MS runs were equalized for normalization (normalization='equalizeMedians'). Intensity=0 was considered as censored missing values (censoredInt='0'). Differential abundance analysis for all possible pairwise comparisons was performed. R script for statistical analysis with MSstats is available (Cox2014_DDA_Skyline_Rscript.R).

1 Cox, J. *et al.* Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics* **13**, 2513-2526, doi:10.1074/mcp.M113.031591 (2014).