Proteomics data analysis

MS/MS data were analysed for protein identifications using MaxQuant 1.6.14 [[1](#_ENREF_1)] with the in-built Andromeda search engine [[2](#_ENREF_2)]. The raw files were searched against the [[3](#_ENREF_3)], using the T. brucei brucei 927 annotated protein sequences from TriTrypDB release 46 [[4](#_ENREF_4)] supplemented with the T. brucei brucei 427 VSG221 (Tb427.*BES40*.*22*) protein sequence. The mass tolerance was set to 4.5 ppm for precursor ions and trypsin was set as proteolytic enzyme with two missed cleavage allowed. Carbamidomethyl on cysteine was set as fixed modifications. SILAC labelling [[5](#_ENREF_5)] of heavy arginine (Arg-6) and Lysine (Lys-6) were specified. Oxidation of methionine and Acetylation of Protein N-term were set as variable modification. The false-discovery rate for protein and peptide level identifications was set at 1%, using a target-decoy based strategy. The minimum peptide length was set to seven amino acids and protein quantification was performed on unique plus razor peptides [[6](#_ENREF_6)]. “Reverse Hits”, “Only identified by site” and “Potential contaminant” identifications were filtered out. Only protein groups with at least two unique peptide sequences and Andromeda protein score greater than 5 were selected for further analysis. The high confidence glycosomal resident proteins were extracted from Güther et al, 2014 [[7](#_ENREF_7)]. The gene symbols were retrieved from the TriTrypDB database. The iBAQ values were extracted from the MaxQuant output and visualised with bar graphs.

The analysis pipeline was implemented in python using the SciPy packages (<https://www.scipy.org/>) [[8](#_ENREF_8)] and Jupyter notebook (https://jupyter.org/).

Data Availability

The raw mass spectrometry files are deposited in the PRIDE database [[9](#_ENREF_9)], with accession id XXXXXX.

The analysis pipeline is available in GitHub (https://github.com/mtinti/nucleotide\_sugar) and it is archived in Zenodo (https://doi.org/10.5281/zenodo.4289929). The analysis pipeline is reproducible using the mybinder app with the link reported in the GitHub repository.

1. Cox, J. and M. Mann, *MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification.* Nat Biotechnol, 2008. **26**(12): p. 1367-72.

2. Cox, J., et al., *Andromeda: a peptide search engine integrated into the MaxQuant environment.* J Proteome Res, 2011. **10**(4): p. 1794-805.

3. Tyanova, S., T. Temu, and J. Cox, *The MaxQuant computational platform for mass spectrometry-based shotgun proteomics.* Nat Protoc, 2016. **11**(12): p. 2301-2319.

4. Aslett, M., et al., *TriTrypDB: a functional genomic resource for the Trypanosomatidae.* Nucleic Acids Res, 2010. **38**(Database issue): p. D457-62.

5. Sury, M.D., J.X. Chen, and M. Selbach, *The SILAC fly allows for accurate protein quantification in vivo.* Mol Cell Proteomics, 2010. **9**(10): p. 2173-83.

6. Cox, J., et al., *A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics.* Nat Protoc, 2009. **4**(5): p. 698-705.

7. Guther, M.L., et al., *High-confidence glycosome proteome for procyclic form Trypanosoma brucei by epitope-tag organelle enrichment and SILAC proteomics.* J Proteome Res, 2014. **13**(6): p. 2796-806.

8. Eric, J., O. Travis, and P. Pearu, *SciPy: Open source scientific tools for Python.* 2001.

9. Perez-Riverol, Y., et al., *The PRIDE database and related tools and resources in 2019: improving support for quantification data.* Nucleic Acids Res, 2019. **47**(D1): p. D442-D450.