SOFTWARE TOOL ARTICLE

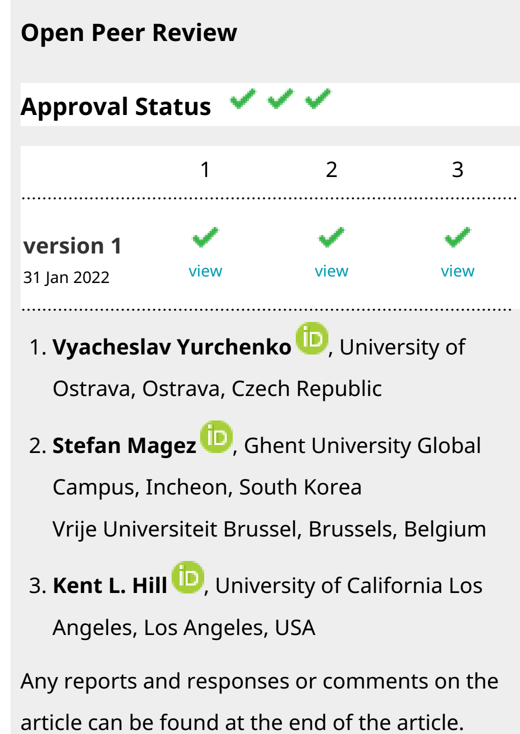
Visualisation of proteome-wide ordered protein abundances in *Trypanosoma brucei* [version 1; peer review: 3 approved]

Michele Tinti, Michael A. J. Ferguson

Wellcome Centre for Anti-Infectives Research, School of Life Sciences, University of Dundee, Dundee, DD1 5HN, UK

v1 **First published:** 31 Jan 2022, **7**:34
https://doi.org/10.12688/wellcomeopenres.17607.1

Latest published: 31 Jan 2022, **7**:34
https://doi.org/10.12688/wellcomeopenres.17607.1



Abstract
Background: *Trypanosoma brucei* is a protozoan parasite and etiological agent of human and animal African trypanosomiasis. It has a complex life cycle, but the most studied cellular types are the *in vitro* cultivated bloodstream- and procyclic-forms. These correspond to the replicating, mammalian host bloodstream-dwelling, slender trypomastigotes and tsetse vector midgut-dwelling procyclic lifecycle stages, respectively. Several proteomics studies have reported the differential abundance of proteins between these *in vitro* cultivated cell types. However, there are no datasets providing protein abundance, from most to least abundant, within and between both cell types.
Methods: We used MaxQuant software 1.6.10.4 to reprocess a recent large-scale proteomics experiment from our laboratory and extracted intensity-based quantifications of the bloodstream and procyclic form proteomes.
Results: We created a web interface to visually explore protein abundances within and between the *in vitro* cultivated *T. brucei* bloodstream and procyclic form proteomes.
Conclusions: The protein abundance visualization tool, searchable by protein name(s) and attribute(s), is likely to be useful to the trypanosome research community. It will allow users to contextualise their proteins of interest in terms of their abundances in the *T. brucei* bloodstream and procyclic form proteomes.

Keywords
Trypanosoma brucei, proteomics, bloodstream form, procyclic form, quantification, iBAQ, web application

Corresponding author: Michael A. J. Ferguson (majferguson@dundee.ac.uk)

Author roles: Tinti M: Conceptualization, Data Curation, Methodology, Software, Writing – Original Draft Preparation, Writing – Review & Editing; **Ferguson MAJ**: Conceptualization, Funding Acquisition, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This work was supported by Wellcome [101842; an Investigator award to M.A.J.F.].
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2022 Tinti M and Ferguson MAJ. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Tinti M and Ferguson MAJ. **Visualisation of proteome-wide ordered protein abundances in *Trypanosoma brucei* [version 1; peer review: 3 approved]** Wellcome Open Research 2022, **7**:34 https://doi.org/10.12688/wellcomeopenres.17607.1

First published: 31 Jan 2022, **7**:34 https://doi.org/10.12688/wellcomeopenres.17607.1

Introduction

The protozoan parasite *Trypanosoma brucei* is transmitted to its human and animal hosts by the tsetse fly (*Glossina* species), which is found only in sub-Saharan Africa[1]. The parasites replicate as the procyclic form (PCF) in the tsetse midgut. Some of these differentiate into the replicating epimastigote form as they migrate to the tsetse salivary glands. These differentiate into non-dividing metacyclic trypomastigote forms that are adapted for transmission to the mammalian host during a tsetse blood-meal. Once in the host, the parasites differentiate into replicating slender bloodstream form (BSF) trypomastigotes and some of these further differentiate into non-dividing stumpy forms that are adapted for transmission to the tsetse vector[1]. If left untreated, the parasites invade the human central nervous system causing character disintegration, coma and death[1]. The proteomes of *in vitro* cultivated BSF and PCF cells have been analysed quite extensively[2–5]; however, the focus of such studies has been the determination of the differential protein abundances between the two lifecycle forms rather than ranked-order relative protein abundance values. The latter can be particularly useful when assessing protein functions. With this in mind, we decided to re-process a deep proteome study recently published in our laboratory[6] to extract **i**ntensity-**b**ased **a**bsolute **q**uantification (iBAQ)[7] values for the *in vitro* cultivated BSF and PCF lifecycle stages. The iBAQ method estimates protein abundances in a complex proteome by integrating all the peptide intensities measured by mass spectrometry for each detected protein group and dividing them by the number of theoretical observable tryptic peptides (i.e., between six and 30 amino acids) contained within them.

Methods

We reprocessed with MaxQuant software 1.6.10.4[8] the protein turnover dataset of *T. brucei* described in Tinti *et al.*, 2019[6] deposited at the PRIDE database[9] with accession number: PXD007115. We used the same parameters described in 6 except for: (i) Using the protein sequences version 46 for the TREU927 clone and version 52 for the Lister427_2018 clone, downloaded from TriTrypDB[10]. (ii) The protein sequences of the TREU927 clone were also supplemented with the BES1/TAR40 protein sequences downloaded from NCBI (accession number: FM162566). (iii) The iBAQ option in the MaxQuant software was selected.

The Tinti *et al.*, 2019 paper[6] reported the *T. brucei* Lister strain 427 *in vitro* cultivated BSF and PCF protein half-lives computed from a label-chase experiment using a **S**table **I**sotope **L**abeling using **A**mino acids in **C**ell culture (SILAC) approach[11]. BSF and PCF parasites were labelled to steady-state in medium SILAC culture media (M) and chased into light SILAC culture media (L). The experiment consisted of seven and nine time points for the BSF and PCF samples, respectively, with three biological replicates each. The samples of the time course experiments were also mixed 1:1 with BSF or PCF parasite lysates, as appropriate, labelled to steady-state in heavy SILAC culture media (H) to provide an internal standard for normalisation. Total proteomes from each biological replicate time point were digested with trypsin and the resulting peptides were separated into 10 sub-fractions for LC-MS/MS analysis. The total dataset therefore represents a total of 210 and 270 individual LC-MS/MS runs for the BSF and PCF samples, respectively, providing particularly deep and robust proteomes. After data re-processing with MaxQuant[8] we considered only the iBAQ values of the H labelled samples retrieved from the proteinGroups.txt file. After equalizing the median values of all the replicates, any missing values were replaced by sampling from a random distribution centred around the minimum value. Finally, the median iBAQ values of the heavy-labelled peptides for each protein were taken for the BSF and PCF replicates. We also computed the numerical data ranks from the median iBAQ values starting from 1 (the least abundant) to n (the most abundant) where n is the number of the protein group identified. The gene id of the leading protein of the protein groups assembled by MaxQuant[8] was used to report the gene id description. The script to process the data were developed in Jupyter Notebook version 0.5.1 using the Python SciPy version 1.4.1 ecosystem[12].

Implementation

We reprocessed a large proteomics dataset of *T. brucei* to extract iBAQ values and ranked the proteins by their abundance values. We focused on two of the most used reference protein databases of *T. brucei*; the TREU927 clone that is better annotated with respect to other clones and the Lister 427 clone that is the most used in laboratory experiments and which was the source of the original proteomics data[6]. We then assembled a web application to compare the iBAQ and protein rank values extracted from the proteomics data for the PCF and BSF lifecycle stages (Figure 1). A tour of the web application starts on page loading and guides the user to the main functionalities. Briefly, we provided two interactive scatter plots, one for the rank values and one for the iBAQ values (Figure 1). Also, an interactive data table allows the data to be searched using protein names or descriptions (Figure 1). Finally, we coded a bar plot of the normalized iBAQ values without data imputation to visually assess: (i) Data reproducibility and (ii) The number of data points used to compute the iBAQ medians (Figure 1). We implemented two identical web applications to visualise the protein abundance rank in BSF and PCF lifecycle stages, except that one is based on the TREU927 proteome and the other on the Lister427_2018 proteome. The source code of the web application was developed in JavaScript D3 version 5 and JavaScript C3 version 0.7.20 for the scatter plots and bar plot visualisation, and DataTables v1.10.21 for the rendering of the data in tabular format. Custom JavaScript code was used to make the scatterplots, bar plot and the data table responsive with respect to each other and to the user interactions.

Operation

The websites are hosted at Netlify (Netlify.com): https://tbrucei-ibaq.netlify.app/web_data/e927/ for the TREU927 proteome and https://tbrucei-ibaq.netlify.app/web_data/e427/ for the Lister 427_2018 proteome. It should be relatively straightforward to clone the git repository (git clone mtinti/tbrucei_ibaq) and run the application locally by opening the application folder with Microsoft's Visual Studio Code software (version 1.63.1) with the "golive" plugin (version 5.7.2). We recommend 1.6 GHz or faster processor and 1 GB of RAM.

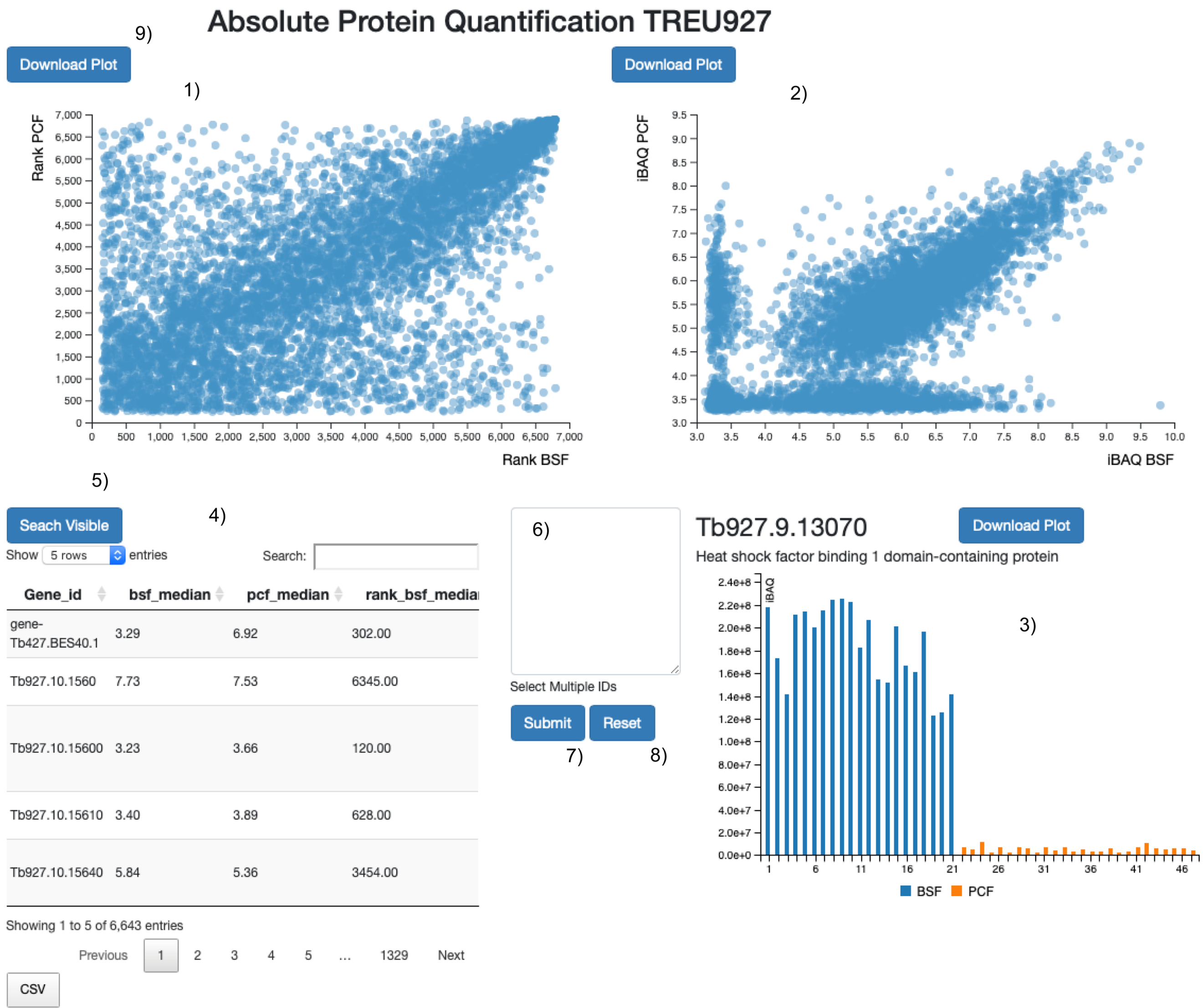


Figure 1. Web application layout. Screen shot of the web application user interface. 1) The protein abundance rank positions, from least to most abundant (1 – >7,000) according to iBAQ values (blue circles). These are plotted for the bloodstream form (x-axis) and the procyclic form (y-axis) proteomes. 2) The median protein iBAQ values themselves (blue circles) are plotted for the bloodstream form (x-axis) and the procyclic form (y-axis) proteomes. 3) Bar plot of all of the individual iBAQ values from the bloodstream form (blue) and procyclic form (amber) proteomes used to determine the median iBAQ values. 4) The main search table that reports the gene id of the leading protein id of the protein group (Gene id); the median iBAQ value in the bloodstream form (bsf_median) and procyclic form (pcf_median) proteomes; the median iBAQ values, rank transformed, in the bloodstream form (rank_bsf_median) and procyclic form (rank_bsf_median) proteomes; the gene id description (Desc); The protein ids of the protein group (Protein IDs). 5) One click of the button and all the proteins visible in the search table [4] are highlighted in the scatter plots [1 and 2]. 6) Text input area to search for gene ids in the two scatter plots [1 and 2]. 7) On click of the button the gene ids present in 8 are highlighted in the two scatter plots [1 and 2]. 8) On click of the button the two scatter plots [1 and 2] are reset. 9) The download button allows the user to save locally a scalable vector graphic (SVG) image of the plot.

Use Cases

The user is presented with a responsive web application with four main components: two scatter plots (Figure 1.1 and 1.2), one bar plot (Figure 1.3) and a search table (Figure 1.4). The user can search for a protein id or attribute (for example, "mitochondrial") in the table search field, and the table will display the matched results. Hovering on the table rows (Figure 1.4) will highlight the position of the proteins in the two linked scatter plots (Figure 1.1 and 1.2) and pull the iBAQ values into the bar plot (Figure 1.3). Any protein visible in the search table (Figure 1.4) will be visualized at once in the two linked scatter plots (Figure 1.1 and 1.2) by clicking the "search visible" button (Figure 1.5). Multiple protein ids can be pasted into the text search field (Figure 1.6) and similarly visualized

at once in the two linked scatter plots (Figure 1.2 and 1.3) by clicking the "submit" button (Figure 1.7). The scatter plot visualisations can be reset by clicking on the "reset" button (Figure 1.8).

The user can hover over any of the circles in the scatter plots (Figure 1.2 and 1.3) to visualize further information on the proteins they represent, such as the protein id, protein description and the x and y iBAQ or protein abundance rank order values in BSF and PCF cells, respectively. Any protein hovered over will be highlighted in both scatter plots (Figure 1.2 and 1.3) and in the search table (Figure 1.4), and the user will see the corresponding iBAQ values in the bar plot (Figure 1.3). By clicking on any protein circle, the user can annotate it with custom text, with the gene id presented as the default option. The annotation can be moved around to better fit in the scatter plot area.

The user can drag a rectangle on either of the two scatter plots (Figure 1.2 and 1.3) to activate a zoom functionality. Only the circles contained in the dragged area will be visualised in the scatter plot and in the search table. The user can download the scatter plots and bar plot as Scalable Vector Graphics (SVG) by clicking on the "download plot" button (Figure 1.9).

Conclusions

We present an effort to rank the *T. brucei* proteome by absolute abundance using the iBAQ values and we provided a visualisation tool to explore the data (Figure 1). The iBAQ values have been shown to perform reasonably well to determine absolute abundance but it is dependent on the quality of the peptide ionization and the number of peptides identified during the mass spectrometry analysis[7,13]. The iBAQ and other quantification methods use a data acquisition protocol in mass-spectrometry named data-dependent acquisition (DDA)[14]. Recent advances in another type of data acquisition in mass spectrometry, named data-independent acquisition (DIA), promise to increase the number of quantified peptides and consequently improve the protein quantification[15]. It is likely that in the future we will replace our estimates of protein abundance with a dataset originating from DIA experiments, but for the time being we believe that our strategy provides a good and useful approximation.

Data availability

No data are associated with this article.

Accession numbers

PRIDE Project: Proteome turnover in bloodstream and procyclic form *Trypanosoma brucei* measured by quantitative proteomics. Accession number PXD007115; https://identifiers.org/pride.project:PXD007115

TriTrypDB: Proteome of *Trypanosoma brucei* procyclic form mitochondrial enriched fraction. Accession number TREU927; https://identifiers.org/tritrypdb:TREU927

TriTrypDB: ChIP-Seq of H4K10ac, bromodomain protein (BDF3), and four histone variants mark the boundaries of polycistronic transcription units in *Trypanosoma brucei*. Accession number TBLister 427-2018; https://identifiers.org/tritrypdb:TbLister 427-2018

NCBI Protein: *Trypanosoma brucei* Lister 427 surface glycoprotein expression site BES1/TAR40 from bloodstream. Accession number FM162566; https://identifiers.org/ncbiprotein:FM162566

Software availability

Source code available from: https://github.com/mtinti/tbrucei_ibaq including analysis pipeline, links to the raw data and code used to extract the data for the web applications. The code to reproduce the data extraction form the protein group files can be inspected using the mybinder badge in GitHub.

Archived source code at time of publication: Zenodo: mtinti/tbrucei_ibaq: v0.2 add zenodo. https://doi.org/10.5281/zenodo.5526824[16]

This project contains the following data:

427

- proteinGroups_ibaq.txt.gz (output file of MaxQuant for the lister 427_2018 clone)

927

- proteinGroups_ibaq.txt.gz (output file of MaxQuant for the TREU927 clone)

Anti-infectives-RGB_200pc.png

README.md (GitHub file describing the repository)

analysis_427.ipynb (The code to extract the data for the lister 427_2018 clone)

analysis_927.ipynb (The code to extract the data for the TREU927 clone)

postBuild (configuration file for mybinder)

requirements.txt (python packages used to extract the data from the protein group files)

utilities.py (python functions used to extract the data from the protein group files)

web_data

- E427 (the web application code for the lister 427_2018 clone)
- E927 (the web application code for the TREU927 clone)

Licence: MIT

References

1. Cox FE: **History of sleeping sickness (African trypanosomiasis).** *Infect Dis Clin North Am.* 2004; **18**(2): 231–45.
PubMed Abstract | Publisher Full Text

2. Dejung M, Subota I, Bucerius F, *et al.*: **Quantitative Proteomics Uncovers Novel Factors Involved in Developmental Differentiation of *Trypanosoma brucei*.** *PLoS Pathog.* 2016; **12**(2): e1005439.
PubMed Abstract | Publisher Full Text | Free Full Text

3. Urbaniak MD, Guther ML, Ferguson MA: **Comparative SILAC proteomic analysis of *Trypanosoma brucei* bloodstream and procyclic lifecycle stages.** *PLoS One.* 2012; **7**(5): e36619.
PubMed Abstract | Publisher Full Text | Free Full Text

4. Urbaniak MD, Martin DM, Ferguson MA: **Global quantitative SILAC phosphoproteomics reveals differential phosphorylation is widespread between the procyclic and bloodstream form lifecycle stages of *Trypanosoma brucei*.** *J Proteome Res.* 2013; **12**(5): 2233–44.
PubMed Abstract | Publisher Full Text | Free Full Text

5. Gunasekera K, Wüthrich D, Braga-Lagache S, *et al.*: **Proteome remodelling during development from blood to insect-form *Trypanosoma brucei* quantified by SILAC and mass spectrometry.** *BMC Genomics.* 2012; **13**: 556.
PubMed Abstract | Publisher Full Text | Free Full Text

6. Tinti M, Güther MLS, Crozier TWM, *et al.*: **Proteome turnover in the bloodstream and procyclic forms of *Trypanosoma brucei* measured by quantitative proteomics [version 1; peer review: 3 approved].** *Wellcome Open Res.* 2019; **4**: 152.
PubMed Abstract | Publisher Full Text | Free Full Text

7. Schwanhäusser B, Busse D, Li N, *et al.*: **Global quantification of mammalian gene expression control.** *Nature.* 2011; **473**(7347): 337–42.
PubMed Abstract | Publisher Full Text

8. Tyanova S, Temu T, Cox J: **The MaxQuant computational platform for mass spectrometry-based shotgun proteomics.** *Nat Protoc.* 2016; **11**(12): 2301–2319.
PubMed Abstract | Publisher Full Text

9. Perez-Riverol Y, Csordas A, Bai J, *et al.*: **The PRIDE database and related tools and resources in 2019: improving support for quantification data.** *Nucleic Acids Res.* 2019; **47**(D1): D442–D450.
PubMed Abstract | Publisher Full Text | Free Full Text

10. Aslett M, Aurrecoechea C, Berriman M, *et al.*: **TriTrypDB: a functional genomic resource for the Trypanosomatidae.** *Nucleic Acids Res.* 2010; **38**(Database issue): D457–62.
PubMed Abstract | Publisher Full Text | Free Full Text

11. Ong SE, Blagoev B, Kratchmarova I, *et al.*: **Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics.** *Mol Cell Proteomics.* 2002; **1**(5): 376–86.
PubMed Abstract | Publisher Full Text

12. Virtanen P, Gommers R, Oliphant TE, *et al.*: **SciPy 1.0: fundamental algorithms for scientific computing in Python.** *Nat Methods.* 2020; **17**(3): 261–272.
PubMed Abstract | Publisher Full Text | Free Full Text

13. Krey JF, Wilmarth PA, Shin JB, *et al.*: **Accurate label-free protein quantitation with high- and low-resolution mass spectrometers.** *J Proteome Res.* 2014; **13**(2): 1034–1044.
PubMed Abstract | Publisher Full Text | Free Full Text

14. Johnson D, Boyes B, Fields T, *et al.*: **Optimization of data-dependent acquisition parameters for coupling high-speed separations with LC-MS/MS for protein identifications.** *J Biomol Tech.* 2013; **24**(2): 62–72.
PubMed Abstract | Publisher Full Text | Free Full Text

15. He B, Shi J, Wang X, *et al.*: **Label-free absolute protein quantification with data-independent acquisition.** *J Proteomics.* 2019; **200**: 51–59.
PubMed Abstract | Publisher Full Text | Free Full Text

16. Tinti M: **mtinti/tbrucei_ibaq: v0.2 add zenodo.** *Zenodo.* 2022.
http://www.doi.org/10.5281/zenodo.5526824

Open Peer Review

Current Peer Review Status:

Version 1

Reviewer Report 15 March 2022

https://doi.org/10.21956/wellcomeopenres.19475.r48676

© **2022 Hill K.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Kent L. Hill

Department of Microbiology, Immunology, and Molecular Genetics, University of California Los Angeles, Los Angeles, CA, USA

The authors provide a 'reprocessing' of published deep proteome datasets from PCF and BSF *Trypanosoma brucei*, allowing determination of relative protein abundance for each protein identified in a given life cycle stage. This is valuable information not readily accommodated by existing proteomic datasets for this organism. The authors further provide a user-friendly navigation tool that allows users to readily identify proteins based on relative abundance in PCF or BSF stages, as well as search for individual proteins (using geneID or description) to quickly obtain relative protein abundance for the protein of interest. The datasets and navigation tools are available for both the TREU927 and the Lister 427 clones of *T. brucei* (a feature that I expect will be much appreciated by many users).

These are useful data and very useful and informative data navigation tools. I expect this will be a valuable and oft-used resource for the research community.

One request for the paper is to adjust the position of the scatter plots in the manuscript figure (Fig 1.1 and 1.2, currently with 'Rank' plot on the left and 'iBAQ' on the right) to match their position in the online tool (currently 'Rank' plot on the right and 'iBAQ' on the left). Not critical, but the different layout is cause for brief confusion if alternating between manuscript and online tool.

Is the rationale for developing the new software tool clearly explained?

Yes

Is the description of the software tool technically sound?

Yes

Are sufficient details of the code, methods and analysis (if applicable) provided to allow replication of the software development and its use by others?

Yes

Is sufficient information provided to allow interpretation of the expected output datasets and any results generated using the tool?
Yes

Are the conclusions about the tool and its performance adequately supported by the findings presented in the article?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Molecular Parasitology, Cell Biology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 01 March 2022

https://doi.org/10.21956/wellcomeopenres.19475.r48579

© 2022 Magez S. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Stefan Magez
[1] Laboratory for Biomedical Research, Ghent University Global Campus, Incheon, South Korea
[2] Laboratory for Cellular and Molecular Immunology, Vrije Universiteit Brussel, Brussels, Belgium

This is going to be a very useful tool for many people with many different interests, including people in the field of immunoparasitology trying to understand how the host context changes protein expression in the parasite and for example how the host immune system is involved in the process of regulation the quorum sensing machinery of trypanosomes. Many more ideas will come to mind as people start to 'play' with this tool.

One suggestion:
- As some technical glitches will come to the surface while people are using the tool, it will be important to provide the user with a contact to whom to direct very practical questions (the corresponding author/PI might not be the person who will be able to free time to deal with this...)

In all, very impressed with the initiative! And a very welcome gift to the trypanosome research community.

Is the rationale for developing the new software tool clearly explained?
Yes

Is the description of the software tool technically sound?

Yes

Are sufficient details of the code, methods and analysis (if applicable) provided to allow replication of the software development and its use by others?

Yes

Is sufficient information provided to allow interpretation of the expected output datasets and any results generated using the tool?

Yes

Are the conclusions about the tool and its performance adequately supported by the findings presented in the article?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Immunoparasitology, diagnostic development and vaccine research - focused on trypanosomosis.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 04 Mar 2022

Michele Tinti, School of Life Sciences, University of Dundee, Dundee, UK

We indeed hope that this tool will be useful for the T. brucei community. We will add a mail contact to report issues with the application.

Competing Interests: No competing interests were disclosed.

Reviewer Report 08 February 2022

https://doi.org/10.21956/wellcomeopenres.19475.r48550

© 2022 Yurchenko V. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Vyacheslav Yurchenko

Life Science Research Centre, Faculty of Science, University of Ostrava, Ostrava, Czech Republic

The presented work of Tinti & Ferguson reports visualization of proteome-wide ordered protein abundances in *T. brucei*. I should mention that I noticed and read this paper even before I was asked to review it. My overall impression of the newly developed tool is that it is destined to

become a standard part of any serious gene-function analysis in *Trypanosoma brucei*. On a side note, I particularly liked that "little tour" offered by the program. I do have some questions and suggestions on how it might be improved.

1. Would it be possible to add and visualize proteomic data from other datasets (not only from Tinti *et al.* 2019)? I envision that one resource combining all the available data may be an ultimate goal.

2. I am slightly confused by the note about BES1/TAR40 locus (~ 60 kb, GenBank FM162566). Why did you supplement TREU927 dataset by LISTER427 sequence?

3. This did not work when I tried it: "The user can drag a rectangle on either of the two scatter plots (Figure 1.2 and 1.3) to activate a zoom functionality. Only the circles contained in the dragged area will be visualised in the scatter plot and in the search table."

4. The provided netlife links (on page 3) redirected me to other web-addresses. Please replace them with direct links.

Is the rationale for developing the new software tool clearly explained?

Yes

Is the description of the software tool technically sound?

Yes

Are sufficient details of the code, methods and analysis (if applicable) provided to allow replication of the software development and its use by others?

Yes

Is sufficient information provided to allow interpretation of the expected output datasets and any results generated using the tool?

Yes

Are the conclusions about the tool and its performance adequately supported by the findings presented in the article?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Molecular parasitology, omics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 09 Feb 2022

Michele Tinti, School of Life Sciences, University of Dundee, Dundee, UK

1)Indeed, it is possible. We find it convenient to use this tool to visualise differential

expression (RNA-seq) or abundance (proteomics) analysis using volcano and MA plots. However, creating a resource to display all the available data is beyond our reach. We will contact the TriTrypDB database and ask if they will consider implementing this visualisation for the datasets deposited there.

2)The 427 strain is commonly used for experiments in the lab, and the BES1 locus is available for the T. brucei 427 strain. However, the BES1 locus is absent in the sequenced genome of the T. brucei 927 strain that is commonly used for data analysis, as it provides a better gene annotation.

3)I tried the functionality in Firefox, Chrome and Safari using a mac laptop. I suspect it might be a windows system issue. I will do more tests to find the problem.

4) We will ask the publisher to change the links of the apps if possible. Unfortunately, I had to change the web hosting company after submitting the paper for an issue I didn't expect.

Competing Interests: No competing interests were disclosed.

Comments on this article

Version 1

Reader Comment 14 Feb 2022

Christine Clayton, ZMBH, Heidelberg, Germany

Incredibly useful, thank-you!

Competing Interests: none