



Label-free quantitative proteomics in *Candida* yeast species: technical and biological replicates to assess data reproducibility

Gaëlle Lelandais, Thomas Denecker, Camille Garcia, Nicolas Danila, Thibaut Léger, Jean-Michel Camadro

► To cite this version:

Gaëlle Lelandais, Thomas Denecker, Camille Garcia, Nicolas Danila, Thibaut Léger, et al.. Label-free quantitative proteomics in *Candida* yeast species: technical and biological replicates to assess data reproducibility. BMC Research Notes, 2019, 12 (1), pp.470. 10.1186/s13104-019-4505-8 . hal-02389332

HAL Id: hal-02389332

<https://cnrs.hal.science/hal-02389332>

Submitted on 2 Dec 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

DATA NOTE

Open Access



Label-free quantitative proteomics in *Candida* yeast species: technical and biological replicates to assess data reproducibility

Gaëlle Lelandais^{1*} , Thomas Denecker¹, Camille Garcia², Nicolas Danila³, Thibaut Léger² and Jean-Michel Camadro^{2,3}

Abstract

Objective: Label-free quantitative proteomics has emerged as a powerful strategy to obtain high quality quantitative measures of the proteome with only a very small quantity of total protein extract. Because our research projects were requiring the application of bottom-up shotgun mass spectrometry proteomics in the pathogenic yeasts *Candida glabrata* and *Candida albicans*, we performed preliminary experiments to (i) obtain a precise list of all the proteins for which measures of abundance could be obtained and (ii) assess the reproducibility of the results arising respectively from biological and technical replicates.

Data description: Three time-courses were performed in each *Candida* species, and an alkaline pH stress was induced for two of them. Cells were collected 10 and 60 min after stress induction and proteins were extracted. Samples were analysed two times by mass spectrometry. Our final dataset thus comprises label-free quantitative proteomics results for 24 samples (two species, three time-courses, two time points and two runs of mass spectrometry). Statistical procedures were applied to identify proteins with differential abundances between stressed and unstressed situations. Considering that *C. glabrata* and *C. albicans* are human pathogens, which face important pH fluctuations during a human host infection, this dataset has a potential value to other researchers in the field.

Keywords: Mass spectrometry, Label-free quantitative proteomics, *Candida glabrata*, *Candida albicans*, Alkaline pH

Objective

Studying proteome dynamics is a key step in systems biology projects. In this context, label-free bottom-up shotgun MS-based proteomics produces quantitative analyses of proteomes. This technique has emerged from significant improvements achieved by mass spectrometry (MS) instrumentation, chromatographic separation systems and a stronger correlation between the relative measured ion intensity and the original molecule abundance in the electrospray ionization process [1–3]. Members of our research team were involved in functional

genomics studies in pathogenic yeasts *Candida glabrata* and *Candida albicans* [4–8]. We observed how the experimental design is a critical step to empower the statistics used to assess the robustness of the results.

“How many replicates is enough?” is certainly one of the most frequently asked questions in wet laboratories. This question is especially critical in situations where the experiments are expensive, and/or the preparation of the biological samples is challenging. Here, our objective was to assess the robustness of the results arising from label-free bottom-up shotgun MS-based proteomics performed in *C. glabrata* and *C. albicans*, in case of technical and biological replicates. If the importance of biological replicates was indisputable when we started

*Correspondence: gaelle.lelandais@u-psud.fr

¹ CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), Univ. Paris-Sud, Gif-Sur-Yvette, France

Full list of author information is available at the end of the article



this project, the interest for technical replicates was more questionable.

We induced proteome modifications applying an alkaline pH stress to *Candida* cells grown in minimal liquid medium. Our final dataset comprises quantitative proteomics for 24 samples (two species, three time-courses, two time points and two runs of mass spectrometry, see below) [9, 10]. We believe it could be useful for other researchers, either interested in a statistical exploitation of the results (to model for instance the variability of protein quantifications associated with biological or technical replicates respectively) or interested in a better understanding of the cellular mechanisms which underly adaptation of pathogenic yeasts to pH changes, a key process during a human host infection [11].

Data description

In this analysis, we performed in *Candida glabrata* (CGLAB) and *Candida albicans* (CALB) yeast species, three cultures referred as CTRL, ALK1 and ALK2. CGLAB and CALB strains are respectively the ones used in [4] and [7], and they were cultured in the same standard conditions as described in [4, 7]. Here, CTRL means “Control”, i.e. the cells were grown in minimal liquid medium. ALK means “alkaline pH stress”, i.e. the cells were subjected to an alkaline stress by adding 1 M of Tris base. This dose was appropriate to slightly affect cell growth without killing the cells. ALK1 and ALK2 referred to two biological replicates, i.e. independent cell growth cultures. T10 and T60 means respectively “time point 10 min” and “time point 60 min”, i.e. the time after stress induction at which the cells were collected for mass spectrometry experiments. These time points were chosen because the cells were then in the exponential phase. Finally, REP1 and REP2 referred to two technical replicates, i.e. independent MS acquisition from the same protein extract and trypsin digestion.

Overall, two datasets were associated to this paper note (Table 1). Data set 1 comprises 24 raw data files, obtained from a Q-Exactive Plus mass spectrometer coupled to

a Nano_LC Proseon 1000 equipped with an easy spray ion source (all from Thermo Fisher Scientific); 48 search files, obtained with the Proteome Discoverer software (Thermo Scientific, version 2.1) and the Mascot search engine (Matrix Science, version 2.5.1); 2 quantification files obtained with the Progenesis QI for Proteomics software (version 4.1, Waters) and 2 FASTA files obtained for the CGD website and used for the MS/MS identification step. Note that detailed descriptions of (i) sample processing protocol and (ii) data processing protocol can be found in [9]. Data file 2 explains the relationship between MS files and associated experimental conditions (CTRL, ALK1, ALK2, T10, T60, REP1 and REP2).

Limitations

We produced this dataset to assess our ability to properly quantify protein abundances in yeasts *Candida glabrata* and *Candida albicans*. An open question for us was the impact of technical replicates compared to biological replicates. We thus performed cell cultures under two different conditions (control and induced stress), collected cells at two separate time points (10 and 60 min) after stress induction, extracted the proteins, performed trypsin digestions and analysed the composition of samples by mass spectrometry. As a result, we were first able to observe a good coverage of proteome in yeasts *C. glabrata* and *C. albicans*, respectively. Between 1500 and 2000 proteins were identified in a reproducible way, representing ~30% of the total protein repositories in these species. It should be noted that a problem in two sample preparations occurred in *Candida glabrata*. Less than 250 proteins were found in technical replicates 1445007-Q3 and 1445007-Q9, which are CGLAB, ALK2, T10, REP1 and REP2 [10]. This is the main limitation for our data. Second, we observed that technical replicates were critical to increase the number of identified proteins, as ~25% of them were found in only one technical replicate. In this context, having a third technical replicate would have been of interest to see if better proteome coverage can still be obtained. Finally, we were able to

Table 1 Overview of the data files related to the study of label-free quantitative proteomics in *Candida* yeasts species, assessing data reproducibility in technical and biological replicates

Label	Name of data file	File types	Data repository and identifier (accession number)
Data set 1 [9]	Proteomics project in <i>Candida</i> yeast species	RAW files (.raw), Search files from Proteome Discoverer software (.pdResults and.xlsx), Quantification files from Progenesis QI for Proteomics software (.csv), FASTA files (.fasta)	http://identifiers.org/pride.project:PX014125
Data file 2 [10]	Detailed correspondence between MS files and experimental conditions	PDF (.pdf)	https://doi.org/10.5281/zenodo.3334949

observe very high positive correlation values (higher than 0.9) between abundances of proteins obtained from biological replicates. If this result is very encouraging, it may also reflect that our cell cultures were not totally “independent”. Indeed, they were performed simultaneously, starting from the same over-night pre-culture. We believe it could be interesting to replicate these experiments paying more attention to this last point, in the design of experiments.

Abbreviations

C. glabrata and CGLAB: *Candida glabrata*; *C. albicans* and CALB: *Candida albicans*; MS: mass spectrometry; CTRL: control; ALK1 and ALK2: alkaline stress 1 and 2; T10 and T60: time point 10 min and time point 60 min; REP1 and REP2: replicate 1 and replicate 2.

Acknowledgements

We would like to thank the research team “Mitochondria, Metals and Oxidative Stress” at the Jacques Monod institute for helpful discussions.

Authors' contributions

GL and JMC designed the experiments, GL and ND performed yeast cell cultures, GL and TD performed statistical and bioinformatics analyses, CG performed protein extractions, TL and JMC performed mass spectrometry experiments, TL uploaded the dataset on PRIDE, GL wrote the manuscript. All authors read and approved the final manuscript.

Funding

This work was funded by the Agence Nationale pour la Recherche (CANDIHUB project, Grant Number ANR-14-CE14-0018-02). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Availability of data and materials

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository [10] with the dataset identifier PXD014125 [9]. Please see Table 1 for details and links to the data.

Ethics approval and consent to participate

Not applicable.

Consent to publish

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹ CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), Univ. Paris-Sud, Gif-Sur-Yvette, France. ² Mass Spectrometry Laboratory, CNRS, Institut Jacques

Monod, UMR 7592, Université de Paris, 75205 Paris, France. ³ CNRS, Institut Jacques Monod (IJM), Univ. Paris Diderot, Paris, France.

Received: 14 June 2019 Accepted: 20 July 2019

Published online: 01 August 2019

References

- Léger T, Garcia C, Videliér M, Camadro J-M. Label-free quantitative proteomics in yeast. In: Methods in molecular biology (Clifton, NJ). 2016. p. 289–307.
- Bantscheff M, Schirle M, Sweetman G, Rick J, Kuster B. Quantitative mass spectrometry in proteomics: a critical review. *Anal Bioanal Chem*. 2007;389(4):1017–31.
- Neilson KA, Ali NA, Muralidharan S, Mirzaei M, Mariani M, Assadourian G, et al. Less label, more free: approaches in label-free quantitative mass spectrometry. *Proteomics*. 2011;11(4):535–53.
- Merhej J, Thiebaut A, Blugeon C, Pouch J, Ali Chaouche MEA, Camadro J-M, et al. A network of paralogous stress response transcription factors in the human pathogen *Candida glabrata*. *Front Microbiol*. 2016;7:645.
- Merhej J, Delaveau T, Guitard J, Palancade B, Hennequin C, Garcia M, et al. Yap7 is a transcriptional repressor of nitric oxide oxidase in yeasts, which arose from neofunctionalization after whole genome duplication. *Mol Microbiol*. 2015;96(5):951–72.
- Benchouaia M, Ripoche H, Sissoko M, Thiébaut A, Merhej J, Delaveau T, et al. Comparative transcriptomics highlights new features of the iron starvation response in the human pathogen *Candida glabrata*. *Front Microbiol*. 2018;9:2689.
- Léger T, Garcia C, Ounissi M, Lelandais G, Camadro J-M. The Metacaspase (Mca1p) has a dual role in farnesol-induced apoptosis in *Candida albicans*. *Mol Cell Proteomics*. 2015;14(1):93–108.
- Léger T, Garcia C, Collomb L, Camadro J-M. A simple light isotope metabolic labeling (SLIM-labeling) strategy: a powerful tool to address the dynamics of proteome variations in vivo. *Mol Cell Proteomics*. 2017;16(11):2017–31.
- Lelandais G, Denecker T, Garcia C, Danila N, Léger T, Camadro J-M. Label-free quantitative proteomics in *Candida* yeast species: technical and biological replicates to assess data reproducibility. *PRIDE*. 2019. <http://identifiers.org/pride.project:PX014125>.
- Lelandais G. Label-free quantitative proteomics in *Candida* yeast species: technical and biological replicates to assess data reproducibility. *ZENODO*. 2019. <https://doi.org/10.5281/zenodo.3334949>.
- Fernandes TR, Segorbe D, Prusky D, Di Pietro A. How alkalization drives fungal pathogenicity. *PLOS Pathog*. 2017;13(11):e1006621.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

