



CHALMERS
UNIVERSITY OF TECHNOLOGY

Benchmarking accuracy and precision of intensity-based absolute quantification of protein abundances in *Saccharomyces cerevisiae*

Downloaded from: <https://research.chalmers.se>, 2021-08-31 11:35 UTC

Citation for the original published paper (version of record):

Sánchez, B., Lahtvee, P., Campbell, K. et al (2021)

Benchmarking accuracy and precision of intensity-based absolute quantification of protein abundances in *Saccharomyces cerevisiae*

Proteomics, 21(6)

<http://dx.doi.org/10.1002/pmic.202000093>

N.B. When citing this work, cite the original published paper.

timsTOF *Pro*

The New Standard for Shotgun Proteomics



Powered by PASEF

While delivering revolutionary improvements in speed, specificity and sensitivity, the timsTOF Pro with TIMS-based PASEF technology also delivers accurate CCS values for all peptides separated in the ion mobility dimension. By doing so, it ushers in the era of 4D-Proteomics and gives scientists the tools to dig deeper into the complex biology revealed by the deep proteome.

- Near 100% duty cycle using dual TIMS technology
- Unrivalled MS/MS speed >100Hz
- Discover more proteins using PASEF

For more information please visit www.bruker.com/timstofpro

TIMS-QTOF MS

Innovation with Integrity

For research use only. Not for use in clinical diagnostic procedures

TECHNICAL BRIEF

Benchmarking accuracy and precision of intensity-based absolute quantification of protein abundances in *Saccharomyces cerevisiae*

Benjamín J. Sánchez^{1,2}  | Petri-Jaan Lahtvee³  | Kate Campbell^{1,2}  |
 Sergo Kasvandik³  | Rosemary Yu^{1,2}  | Iván Domenzain^{1,2}  |
 Aleksej Zelezniak¹  | Jens Nielsen^{1,2,4} 

¹ Department of Biology and Biological Engineering, Chalmers University of Technology, Gothenburg, Sweden

² Novo Nordisk Foundation Center for Biosustainability, Chalmers University of Technology, Gothenburg, Sweden

³ Institute of Technology, University of Tartu, Tartu, Estonia

⁴ Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Lyngby, Denmark

Correspondence

Jens Nielsen, Department of Biology and Biological Engineering, Chalmers University of Technology, Gothenburg, Sweden.
 Email: nielsenj@chalmers.se

Funding information

Horizon 2020 Framework Programme, Grant/Award Numbers: 668997, 686070; Comisión Nacional de Investigación Científica y Tecnológica, Grant/Award Number: 6222/2014; Eesti Teadusagentuur, Grant/Award Number: PUT1488P; Novo Nordisk Fonden; Knut och Alice Wallenbergs Stiftelse

Abstract

Protein quantification via label-free mass spectrometry (MS) has become an increasingly popular method for predicting genome-wide absolute protein abundances. A known caveat of this approach, however, is the poor technical reproducibility, that is, how consistent predictions are when the same sample is measured repeatedly. Here, we measured proteomics data for *Saccharomyces cerevisiae* with both biological and inter-batch technical triplicates, to analyze both accuracy and precision of protein quantification via MS. Moreover, we analyzed how these metrics vary when applying different methods for converting MS intensities to absolute protein abundances. We demonstrate that our simple normalization and rescaling approach can perform as accurately, yet more precisely, than methods which rely on external standards. Additionally, we show that inter-batch reproducibility is worse than biological reproducibility for all evaluated methods. These results offer a new benchmark for assessing MS data quality for protein quantification, while also underscoring current limitations in this approach.

KEYWORDS

absolute proteomics, batch effect, iBAQ, mass spectrometry, UPS2

Mass spectrometry (MS) is currently the main technology used for predicting genome wide protein copy number per cell, thanks to its high sensitivity, specificity, and multiplexing capacity [1]. Among the different MS technologies available, quantitative label-free methods are becoming increasingly popular, due to their relative ease of use and

cost-effectiveness, particularly when compared to more expensive and laborious methods, such as isotope-labeled peptide based approaches [2]. In quantitative label-free methods, normalization of the raw data is a critical step when predicting protein absolute abundance [3–6]. Two fundamental metrics for assessing the quality of these predictions are: (i) accuracy, that is, how far away from the true value the prediction is, and (ii) precision, that is, how variable different predictions are when the same measurement is repeated (also referred to as reproducibility).

Abbreviations: FC, Fold Change; FC_m, median absolute Fold Change; iBAQ, intensity-Based Absolute Quantification; MS, Mass Spectrometry; PC, Principal Component; R², coefficient of determination; SILAC, Stable Isotope Labeling by Amino acids in Cell culture; TPA, Total Protein Approach; UPS, Universal Proteomics Standard

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2021 The Authors. *Proteomics* published by Wiley-VCH GmbH

There are several factors that affect the precision and accuracy of absolute protein abundance predictions generated via MS. These are: (i) the intrinsic biological nature of the proteome, with the dynamic range of intracellular protein abundance being able to span several orders of magnitude; (ii) the physicochemical nature of amino acids: as peptide molecules can have different ionization properties, this can lead to two similarly abundant molecules having different capacities for detection by the MS; and (iii) the differences in MS instrumentation (e.g., Orbitraps versus time-of-flight instruments), chromatography and experimental protocols. All of the above factors yield only modest results in MS-based analyses when comparing predictions to the true protein concentrations values [7–9], and is highly likely to contribute to a large level of variability, the latter which is observed across different proteomics studies [7, 10, 11].

Studies that compute absolute protein abundance commonly address biological reproducibility by running biological replicates in the same MS batch [7, 12, 13]. However, awareness of how the MS instrument itself impacts protein abundance, that is, technical reproducibility, has been less studied. This can be determined by running the same biological sample in the same batch [14], or in separate batches [15]; with the latter often referred to as “the batch effect.” As different normalization/scaling methods can be used to predict protein abundance from raw MS intensities [16], it is interesting to study how these methods propagate the inter-batch technical variability into uncertainty in the final protein abundance predictions. In this study, we analyze both accuracy and technical precision of intensity-based absolute quantification of a proteomics dataset from *S. cerevisiae* and show how prediction quality can be improved using different normalization/scaling methods. In particular, we show that a simple rescaling method [5] performs as accurately as but more precisely than alternatives that rely on the use of costly external standards.

We generated a proteomics dataset using the *S. cerevisiae*'s strain CEN.PK113-7D, containing both biological triplicate and technical replicate samples. Samples were obtained from aerobic glucose-limited chemostats at a dilution rate of 0.1/h and were mixed with an internal standard, using stable isotope labeling by amino acids in cell culture (SILAC). Here, a lysine auxotrophic strain was grown in medium supplemented with double labelled heavy ^{15}N , ^{13}C -lysine (Cambridge Isotope Laboratories Inc.); samples were then mixed in a 1:1 ratio with each of the other non-labelled (“light”) samples. The internal standard was also mixed with an external standard of known concentrations, in a ratio of 6:1.1. The external standard used here was the Proteomics Dynamic Range Standard Set (UPS2) mix (Merck), consisting of 48 human proteins in a dynamic concentration range from 500 amol to 50 pmol. All mixed samples were stored at -80°C until their analysis, wherein they were similarly processed; the latter step being crucial in order to isolate variability from either the biological source or the MS equipment, and not from other sources such as sample preparation differences.

For proteome identification, samples were digested with 1:50 LysC overnight at room temperature. Peptides were separated on an Ultimate 3000 RSLCnano system (Dionex), eluted to a Q Exactive Plus (Thermo Fisher Scientific) tandem mass spectrometer and identified with the MaxQuant 1.4.0.8 software package [17], maintaining the

peptide-spectrum match and the protein false discovery rate below 1% using a target-decoy approach. Each sample was measured six times: on three separate batches of the MS instrument (with a time difference of 12 and 30 days), and each time twice, using Top5 and Top10 data-dependent acquisition strategies, wherein only the top five or ten highest intensity peptide peaks per one MS full scan were selected for MS/MS analysis, respectively (additional details on the experimental setup can be found in the Supplementary Material).

Using the described data as a reference, we then evaluated the ability of four different methods for transforming the MS intensity computed by MaxQuant (which corresponds to the sum of all associated peptide intensities) to protein abundances of the internal standard. The first method, known as intensity based absolute quantification (iBAQ) [3], normalizes each protein MS intensity by the corresponding number of theoretically observable peptides, then infers the abundances of each internal standard protein using a linear model generated from the external standard (normalized protein MS intensity vs. known protein quantities). As this method yields abundances that do not always add up to equal amounts of protein injected per sample (Figures S1–S2), a second method was also assessed that rescales all abundances from iBAQ to equal the total injected mass. The third method tested was the total protein approach (TPA) [18], which bypasses the need for an external standard and instead assumes that the sum of MS intensities of all detected proteins multiplied by the corresponding molecular weights should be proportional to the total amount of protein injected. Finally, the fourth method tested was a variation of the TPA method [5], which first normalizes protein intensities with the number of theoretically observable peptides.

$$\log_{10}(P_{i1}) = m_{ES} \cdot \log_{10}\left(\frac{A_i}{N_i}\right) + n_{ES} \quad [\text{Method 1}]$$

$$P_{i2} = \frac{P_{i1}}{\sum_i (MW_i \cdot P_{i1})} \quad [\text{Method 2}]$$

$$P_{i3} = \frac{A_i}{\sum_i (MW_i \cdot A_i)} \quad [\text{Method 3}]$$

$$P_{i4} = \frac{\frac{A_i}{N_i}}{\sum_i \left(MW_i \cdot \frac{A_i}{N_i} \right)} \quad [\text{Method 4}]$$

where P_{ij} is the predicted absolute abundance of protein i by method j [fmol/ μg protein], m_{ES} and n_{ES} are the parameters of the external standard curve, A_i is the sum of all peptide intensities associated to protein i , N_i is the number of theoretically observable peptides for protein i , and MW_i is the molecular weight of protein i [kDa]. The abundances of the biological triplicates were also computed differently depending on the method: For Method 1, the corresponding internal standard abundance (i.e., heavy fraction) was used, together with the normalized H/L ratios obtained from each sample run [19]. For Method 2, the same transformation as for the internal standard was used, that is, using the protein predictions from Method 1. Finally, for Methods 3 and 4, the transformation used for the internal standard was used as well, only this time on the light fraction.

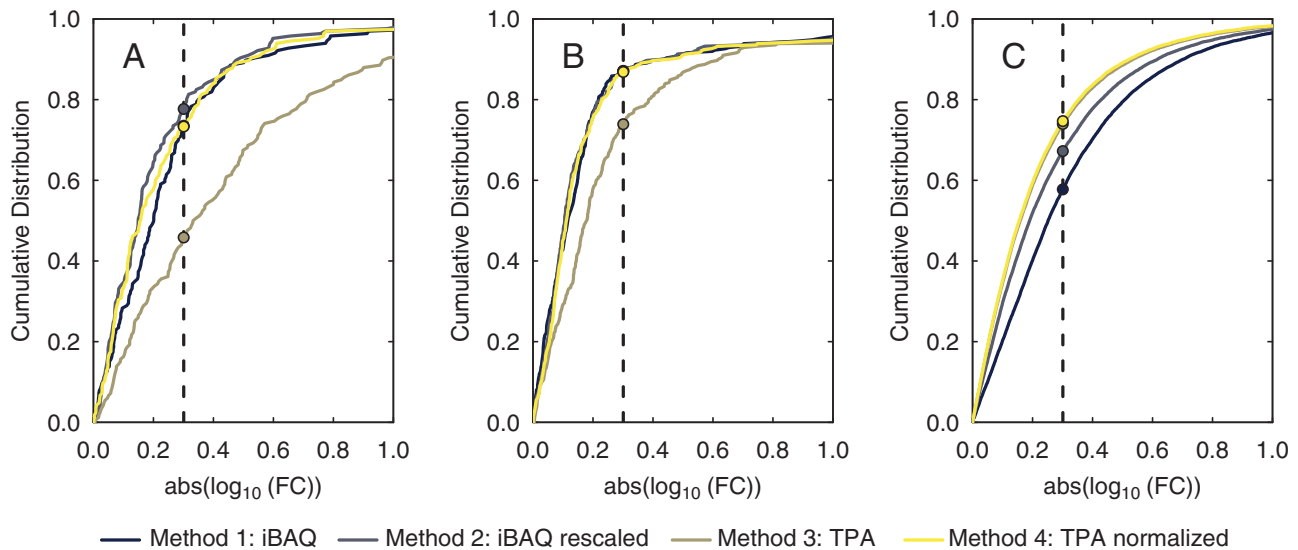


FIGURE 1 Cumulative distributions of fold changes (FC) between predictions of all four methods, with respect to (A) accuracy (test 1): predicted vs known values of the external standard ($N = 167$), (B) accuracy (test 2): predicted values versus median estimated value for all ribosomal proteins in the internal standard ($N = 312$), and (C) precision: all possible combinations between batches of the internal standard ($N = 14,182$). A fold change of 2 is indicated with a vertical dashed line.

Using the generated dataset, we evaluated accuracy and precision of predicted abundances by the four different methods. To evaluate accuracy, we computed the differences as fold changes between the predicted abundances of the external standard proteins detected by the MS ($n = 31/48$) and the known values in the UPS2 mix. Here, Methods 1, 2 and 4 performed similarly, whereas Method 3 had a significantly higher error (Figure 1A, Figure S3). Specifically, more than 50% of protein abundance predictions from Method 3 deviated from the true value by less than two-fold. We further evaluated the accuracy of each method by testing protein predictions in the ribosome, a protein complex with subunit abundance in equal stoichiometry [20]. Of these subunits, 62 out of 79 were detected in the internal standard, after accounting for paralogs, and compared to their median abundance value, with the expectation that each ribosomal subunit has the same abundance as all others in the complex [21]. Once again, Methods 1, 2 and 4 performed similarly, and outperformed Method 3 (Figure 1B, Figures S4-S5), which we found to be true in the abundance predictions of both the internal standard and the biological triplicates (Figures S6-S7).

We next proceeded to evaluate precision, by comparing protein predictions between all three batches both for the internal standard and the biological triplicates (Figures S8-S9). A cumulative distribution of all fold changes within the internal standard (Figure 1C) showed that Methods 3 and 4 significantly outperformed Methods 1 and 2 (all P -values < 0.001). In particular, by using Methods 3 or 4, protein abundance varied by less than two-fold for nearly 75% of all proteins, whereas in the case of Method 1 this was under 60%. Similar observations can also be made when looking at the biological triplicates (Figure S10). Higher inter-batch variability of Methods 1 and 2 was observed both for lowly and highly abundant proteins but especially for proteins below the detection range of the external standard curve (Figures S11-

S12), and can be explained by the bias introduced by the external standard (Figures S13-S14), which Methods 3 and 4 did not use.

Taking into consideration results for both the accuracy and precision tests that we performed (Figure 1), we conclude that the best-performing method is Method 4, which omits the use of an external standard and instead rescales normalized MS intensities to equal the injected sample mass. Even though Methods 1 and 2 perform similarly to Method 4 in terms of accuracy, they are not as precise, while although Method 3 is as precise as Method 4, it is not as accurate. Therefore, considering that iBAQ involves significant additional costs to users (including purchasing of the external standard and additional MS running time); however, does not yield better performance, we propose that the rescaling of normalized MS intensities can be used instead. This method can also be used as a benchmark for assessing the predictive power of alternative approaches for computing absolute protein abundances from MS methods.

It is noteworthy to mention that for all methods, the variability between biological replicates in the same MS batch is considerably lower than the variability between batches of the same biological sample. We exemplify this with the biological and batch variability from Method 1 predictions (Figure 2A and B, respectively), and with a principal component analysis of the same predictions (Figure 2C), wherein samples cluster based on batches, not biological replicates. Although inter-batch variability becomes lowest when using Method 4 (Figure S8, Table S1), coming much closer to biological variability levels, still $\sim 25\%$ of predictions in the internal standard have over a two-fold of variability. This remaining variability is most likely due to the presence of stochastic and non-linear effects in shotgun proteomics [22, 23]. For instance, for each protein there were on average close to five peptides that were different between batches (Table S2, Figure S15), due to a difference in selection of the most intense (top N) precursor ions,

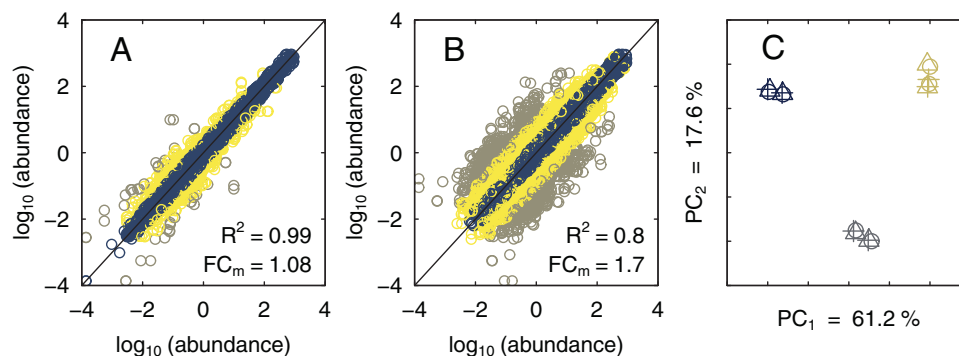


FIGURE 2 (A-B) Variability of predicted abundances [fmol/ μg protein] by Method 1, between biological replicates (A) and MS batches (B). Fold changes within a two-fold are shown in blue, between a two-fold and ten-fold in yellow, and above a ten-fold in grey. The coefficient of determination (R^2) and the median absolute fold change (FC_m) are also displayed. (C) Principal component analysis of all samples. Different colors refer to different batches, and different shapes refer to different biological replicates. The amount of variability each of the first two components explains is shown as a percentage.

ultimately affecting protein abundance predictions. Researchers working with computational methods that rely on absolute protein abundances [24] should therefore be aware of these limitations and interpret results accordingly.

In conclusion, we present a comprehensive proteomics dataset of yeast, designed for assessment of absolute protein quantification for different biological replicates and batches of samples. Furthermore, we show that a simple method of normalization and rescaling can yield superior results over more complicated and expensive methods such as iBAQ. As protein intensity is used as input, this method can be used both on pre-existing and future datasets regardless of how intensity values were generated, including labeled or unlabeled methods. We therefore expect both our dataset and method to be of benefit to users when assessing accuracy and precision of MS-based approaches in current and future proteomics studies.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Christine Räisänen and Gang Li for reviewing the manuscript, and the anonymous reviewers that contributed with valuable feedback. This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement no 686070 and no 668997, the Novo Nordisk Foundation and the Knut and Alice Wallenberg Foundation. BJS and PJJ acknowledge financial support from CONICYT (grant #6222/2014) and the Estonian Research Council (grant PUT1488P), respectively.

AUTHOR CONTRIBUTIONS

Conceptualization: B.J.S., P.J.L., J.N.; Data generation: P.J.L., S.K.; Data Analysis: B.J.S., K.C., S.K., R.Y., I.D., A.Z.; Project Supervision: J.N.; Writing – Original Draft: B.J.S.; Writing – Review and Editing: B.J.S., P.J.L., K.C., S.K., R.Y., I.D., A.Z., J.N.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

All MS data used in this study have been deposited to the ProteomeXchange Consortium via the PRIDE [25] partner repository with the dataset identifier PXD011725. Output tables from MaxQuant, together with all necessary scripts to reproduce the results presented in this study are available at <https://github.com/SysBioChalmers/reproduce> and have been archived in Zenodo [26].

ORCID

Benjamin J. Sánchez  <https://orcid.org/0000-0001-6093-4110>
 Petri-Jaan Lahtvee  <https://orcid.org/0000-0002-3327-3190>
 Kate Campbell  <https://orcid.org/0000-0002-4173-5260>
 Sergo Kasvandik  <https://orcid.org/0000-0003-0218-2410>
 Rosemary Yu  <https://orcid.org/0000-0001-9901-4055>
 Iván Domenzain  <https://orcid.org/0000-0002-5322-2040>
 Aleksej Zelezniak  <https://orcid.org/0000-0002-3098-9441>
 Jens Nielsen  <https://orcid.org/0000-0002-9955-6003>

REFERENCES

- Mann, M., Kulak, N. A., Nagaraj, N., & Cox, J. (2013). The coming age of complete, accurate, and ubiquitous proteomes. *Molecular Cell*, 49, 583–590.
- Calderón-Celis, F., Encinar, J. R., & Sanz-Medel, A. (2017). Standardization approaches in absolute quantitative proteomics with mass spectrometry. *Mass Spectrometry Reviews*, 37, 715–737.
- Schwahnhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., & Selbach, M. (2011). Global quantification of mammalian gene expression control. *Nature*, 473, 337–342.
- Ning, K., Fermin, D., & Nesvizhskii, A. I. (2012). Comparative analysis of different label-free mass spectrometry based protein abundance estimates and their correlation with RNA-Seq gene expression data. *Journal of Proteome Research*, 11, 2261–2271.
- Wiśniewski, J. R., Ostasiewicz, P., Duś, K., Zielińska, D. F., Gnad, F., & Mann, M. (2012). Extensive quantitative remodeling of the proteome between normal colon tissue and adenocarcinoma. *Molecular Systems Biology*, 8, 611.
- Wiśniewski, J. R., Hein, M. Y., Cox, J., & Mann, M. (2014). A “Proteomic Ruler” for protein copy number and concentration estimation without spike-in standards. *Molecular Cell Proteomics*, 13, 3497–3506.

7. Arike, L., Valgepea, K., Peil, L., Nahku, R., Adamberg, K., & Vilu, R. (2012). Comparison and applications of label-free absolute proteome quantification methods on *Escherichia coli*. *Journal Proteomics*, 75, 5437–5448.
8. Weisser, H., Nahnsen, S., Grossmann, J., Nilse, L., Quandt, A., Brauer, H., Sturm, M., Kenar, E., Kohlbacher, O., Aebersold, R., & Malmström, L. (2013). An automated pipeline for high-throughput label-free quantitative proteomics. *Journal of Proteome Research*, 12, 1628–1644.
9. Shuford, C. M., Walters, J. J., Holland, P. M., Sreenivasan, U., Askari, N., Ray, K., & Grant, R. P. (2017). Absolute protein quantification by mass spectrometry: Not as simple as advertised. *Analytical Chemistry*, 89, 7406–7415.
10. Lawless, C., Holman, S. W., Brownridge, P., Lanthaler, K., Harman, V. M., Watkins, R., Hammond, D. E., Miller, R. L., Sims, P. F. G., Grant, C. M., Eysers, C. E., Beynon, R. J., & Hubbard, S. J. (2016). Direct and absolute quantification of over 1800 yeast proteins via selected reaction monitoring. *Molecular Cell Proteomics*, 15, 1309–1322.
11. Ho, B., Baryshnikova, A., & Brown, G. W. (2018). Unification of protein abundance datasets yields a quantitative *Saccharomyces cerevisiae* proteome. *Cell Systems*, 6, 192–205.e3.
12. Hukelmann, J. L., Anderson, K. E., Sinclair, L. V., Grzes, K. M., Murillo, A. B., Hawkins, P. T., Stephens, L. R., Lamond, A. I., & Cantrell, D. A. (2016). The cytotoxic T cell proteome and its shaping by mammalian target of rapamycin. *Nature Immunology*, 17, 104–112.
13. Lahtvee, P. J., Sánchez, B. J., Smialowska, A., Kasvandik, S., Elsemman, I. E., Gatto, F., & Nielsen, J. (2017). Absolute quantification of protein and mRNA abundances demonstrate variability in gene-specific translation efficiency in yeast. *Cell Systems*, 4, 495–504.e5.
14. Shalit, T., Elinger, D., Savidor, A., Gabashvili, A., & Levin, Y. (2015). MS1-based label-free proteomics using a quadrupole orbitrap mass spectrometer. *Journal of Proteome Research*, 14, 1979–1986.
15. Vowinckel, J., Zelezniak, A., Bruderer, R., Müllereder, M., Reiter, L., & Ralser, M. (2018). Cost-effective generation of precise label-free quantitative proteomes in high-throughput by microLC and data-independent acquisition. *Scientific Reports*, 8, 1–10.
16. Rosenberger, G., Ludwig, C., Röst, H. L., Aebersold, R., & Malmström, L. (2014). aLFQ: An R-package for estimating absolute protein quantities from label-free LC-MS/MS proteomics data. *Bioinformatics*, 30, 2511–2513.
17. Cox, J., & Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology*, 26, 1367–1372.
18. Wiśniewski, J. R., & Rakus, D. (2014). Multi-enzyme digestion FASP and the 'Total Protein Approach'-based absolute quantification of the *Escherichia coli* proteome. *Journal Proteomics*, 109, 322–331.
19. Geiger, T., Cox, J., Ostasiewicz, P., Wisniewski, J. R., & Mann, M. (2010). Super-SILAC mix for quantitative proteomics of human tumor tissue. *Nature Methods*, 7, 383–385.
20. Jenner, L., Melnikov, S., de Loubresse, N. G., Ben-Shem, A., Iskakova, M., Urzhumtsev, A., Meskauskas, A., Dinman, J., Yusupova, G., & Yusupov, M. (2012). Crystal structure of the 80S yeast ribosome. *Current Opinion in Structural Biology*, 22, 759–767.
21. Fabre, B., Lambour, T., Bouyssié, D., Menneteau, T., Monsarrat, B., Burlet-Schiltz, O., & Bousquet-Dubouch, M. P. (2014). Comparison of label-free quantification methods for the determination of protein complexes subunits stoichiometry. *EuPA Open Proteomics*, 4, 82–86.
22. Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E., & Storey, J. D. (2012). The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics*, 28, 882–883.
23. Čuklina, J., ETH Zurich (2018).
24. Sánchez, B. J., Zhang, C., Nilsson, A., Lahtvee, P., Kerkhoven, E. J., & Nielsen, J. (2017). *Molecular Systems Biology*, 13, 935.
25. Vizcaíno, J. A., Csordas, A., Del-Toro, N., Dianes, J. A., Griss, J., Lavidas, I., Mayer, G., Perez-Riverol, Y., Reisinger, F., Ternent, T., Xu, Q. W., Wang, R., & Hermjakob, H. (2016). *Nucleic Acids Research*, 44, D447–D456.
26. Sánchez, B. J. (2020). <https://doi.org/10.5281/zenodo.4192409>.

SUPPORTING INFORMATION

Additional supporting information may be found online <https://doi.org/10.1002/pmic.202000093> in the Supporting Information section at the end of the article.

How to cite this article: Sánchez, B. J., Lahtvee, P.-J., Campbell K., et al. (2021). Benchmarking accuracy and precision of intensity-based absolute quantification of protein abundances in *Saccharomyces cerevisiae*. *Proteomics*, e2000093. <https://doi.org/10.1002/pmic.202000093>