Want to publish in JPR? This is what you need to know!

Yates, J.R.; Cristea, I.M.; Dong, M.Q.; Eyers, C.E.; LaBaer, J.; Li, J.V.; ... ; Slavov, N.

Journal of proteome research

pubs.acs.org/jpr

Editorial

Want to Publish in *JPR*? This Is What You Need to Know!

Cite This: *J. Proteome Res.* 2022, 21, 2837−2839

Read Online

ACCESS | Metrics & More | Article Recommendations

Proteomics has long been defined as the study of ensembles of proteins from cells, complexes, and subcellular spaces, thus distinguishing it from biochemistry, where rigorous analyses of single proteins are the norm. Since its inception, the *Journal of Proteome Research* has focused on the study of ensembles of both proteins and metabolites and has chronicled the growth of the fields of proteomics and metabolomics. As these fields have grown, expectations of the manuscripts submitted to *JPR* have evolved and it is useful to periodically explain what the current expectations are.

So, what are we looking for at *JPR*? First and foremost, we continue to prioritize exciting, groundbreaking science. Studies are welcomed from a wide swath of proteomic and metabolomic science, not just the discovery of new biology or biological mechanisms. We are excited to see the development of proteomic and metabolomic methods and application to new areas, including studies in paleobiology, archeology, rare species (e.g., Komodo dragon), and forensic science, to name a few. The variety of studies appropriate for *JPR* is a testament to the versatility of proteomic and metabolomic methods to create new scientific directions.

Of course, many studies of physiological processes and biological systems employ proteomic and metabolomic technologies. Comparisons between a normal state and a perturbed state can yield important insights or hypotheses about the mechanisms behind the perturbation, but these hypotheses should be validated or tested. Common follow-up experiments include knockout, knock-down, overexpression, or chemical modulation (small molecule or drug treatment) of a protein(s) to duplicate a phenotype and/or other experiments that identify the proteins or metabolites associated with a perturbation or disease state. We do not require that manuscripts provide a cure for a disease, but we expect them to advance our understanding of the mechanisms behind the disease or biology. Simple presentation of several quantitative data sets followed by standard GO and STRING analyses are potentially interesting, but they do not validate the connection between protein or metabolite changes and the phenotype. We are increasingly seeing proteomic or metabolomic studies in nonmodel organisms where a genome may be sequenced, but molecular biology tools do not yet exist for robust follow-up experiments. For these studies, cell lines or closely related organisms can be used for follow-up experiments. Several model organisms (mouse, fruit fly, worm, and yeast) have been used as biological surrogates to study the functions of the related human genes when experiments with humans are not possible.

We also expect to see proper statistical design of quantitative experiments. A review by Oberg and Vitek from 2009 provides guidelines for experimental design.[1] Replicates are critical for determining experimental and biological variation, and studies should include the proper statistical, analytical, and biological validation necessary for the study. Similarly, biomarker studies should include enough samples to power the study and provide good statistics. This is especially important for proteomics and metabolomics where the number of variables is often high relative to the sample number and the data is at risk of being overfitted. Once a discovery set has yielded putative biomarkers, they should be *validated* in a second set of independent patient samples, and Receiver Operator Characteristic curves should be used to determine the specificity and selectivity of the candidate biomarkers. We recognize that testing potential biomarkers in a separate patient population can be time-consuming and expensive and that it may not be feasible to test biomarker results in a model system. However, for a molecule to be a true biomarker, rigorous validation is required. The literature is awash with "potential" biomarkers, and we want to avoid adding to the noise. There is a difference between the association of a protein or metabolite with a disease state and its utility as a biomarker. Researchers sometimes note that they are working with a rare disease and samples are hard to obtain. Indeed, more samples would improve the likelihood of finding rare markers, but it is often the *control samples* that are lacking, and these are much easier to obtain. Controls are critical to rule out the false positives that result from overfitting. When working with rare diseases, using twice as many controls as cases can improve the specificity of the markers.

Although we encourage follow up of data generated through -omics technologies, there are data sets that can be extremely valuable on their own. The generation of data resources are certainly of interest to *JPR*, but as you contemplate such studies you should keep in mind Sydney Brenner's CAP criteria:

- **C − Complete**, i.e. do a thorough job so that someone does not have to come along later to redo it.
- **A − Accurate**, i.e. take care to ensure the data is well collected and properly vetted.
- **P − Permanent**, i.e. the data should stand forever. It is understood that -omics data are dynamic and change

Published: November 14, 2022

https://doi.org/10.1021/acs.jproteome.2c00704
J. Proteome Res. 2022, 21, 2837−2839

based on conditions, but the data collected under the criteria used should be rock solid.

Protein lists are also being generated in new and unusual studies like paleoproteomics, forensics, and archeology. Proteins from old skeletons have been used to settle long debated taxonomic questions for extinct animals. Obviously, follow-up studies in these organisms are not possible, but they are not necessary because the studies are not investigating biological mechanisms. Nevertheless, the data used to support key findings should be substantiated by annotation of the tandem mass spectra used to make any claims and conclusions in the study.[2] When unusual or extraordinary claims are made, this data is best substantiated by synthesizing the relevant peptides or small molecules to compare with the raw data.

JPR also prides itself on publishing software tools for proteomics and metabolomics. We publish a software issue every other year to showcase new software tools for mass spectrometry, array informatics, or the conversion of data to knowledge. We very much encourage the publication of software tools that demonstrate the potential of software to identify new information or improve data analysis performance. We often get papers that describe computational modeling methods to identify drugs or peptides that would be good inhibitors for proteins. The results of these studies should be validated by testing the findings in an experimental setting. In mechanism of action studies, dosing needs to be described and compounds should be well characterized.

JPR has long been a leading venue for new proteomic and metabolomic methods, and we look forward to publishing innovations that increase sensitivity or throughput, or that provide new readouts of protein function, such as conformation, localization, and interactions. Best practices would include showing proof of the method on a system where the answer is known and then demonstrating that it works on an unknown system. The methods should be described well enough for others to readily replicate them, and they should provide compelling enough advantages over current methods to warrant their adoption. Hence, to support the novelty and usefulness of any new method, it should be benchmarked against the current best method.

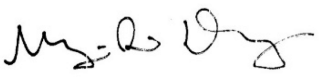
To ensure transparency and utility of the data collected from proteomic experiments, we require that data and metadata be deposited into a data repository. Proteomic data can be uploaded to a ProteomeXchange partner like PRIDE or MassIVE, or to a local national subsidiary of ProteomeXchange. Each sample reported in a paper and represented in the raw data should also be sufficiently described in the associated metadata file to support reanalysis and extended analysis of the data. Metabolomic data can be submitted to GNPS at the University of California San Diego, the Metabolomics Workbench at UCSD, or MetaboLights at EBI. RNA-Seq data needs to be submitted to repositories, e.g., NCBI Gene Expression Omnibus (GEO), following ENCODE3 RNA-Seq guidelines (https://www.encodeproject.org/about/experiment-guidelines). All FASTQ sequences should be uploaded to NCBI GenBank/Sequence Read Archive and associated metadata should be uploaded to NCBI Gene Expression Omnibus. Data workflows for processing and analysis must be based on best practices and documented for open access in GitHub. All software, ontologies, and vocabularies must be described in terms of their version numbers and dependencies and implemented in the form of containerized workflows for reusability.

We also require the submission of raw uncropped Western blot images as part of the Supporting Information, and they should be provided for review at submission. Full details of the antibodies used must be provided to ensure data reproducibility, including, species, supplier source, lot number, catalogue number, and Research Resource Identification (RRID) number.

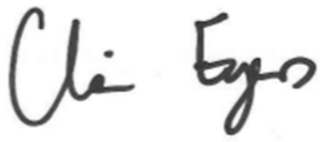
As we tell our laboratories "We are in the business of knowledge generation, so write it up!". We look forward to receiving your latest scientific advances.

John R. Yates

Ileana M. Cristea



Meng-Qui Dong



Claire E. Eyers

Joshua LaBaer

Jia V. Li

Jeremy K. Nicholson

https://doi.org/10.1021/acs.jproteome.2c00704
J. Proteome Res. 2022, 21, 2837−2839

Christopher M. Overall

Magnus Palmblad

Nikolai Slavov

Nikolai Slavov

John R. Yates, Editor-in-Chief, Scripps Research Institute, La Jolla, California, USA orcid.org/0000-0001-5267-1672

Ileana M. Cristea, Associate Editor, Princeton University, Princeton, New Jersey, USA orcid.org/0000-0002-6533-2458

Meng-Qui Dong, Associate Editor, National Institute of Biological Sciences, Beijing, China

Claire E. Eyers, Associate Editor, University of Liverpool, Liverpool, UK orcid.org/0000-0002-3223-5926

Joshua LaBaer, Associate Editor, Arizona State University, Tempe, Arizona, USA orcid.org/0000-0001-5788-9697

Jia V. Li, Topic Editor, Imperial College London, London, UK orcid.org/0000-0002-5763-6670

Jeremy K. Nicholson, Associate Editor, Murdoch University, Perth, Western Australia orcid.org/0000-0002-8123-8349

Christopher M. Overall, Associate Editor, The University of British Columbia, Vancouver, British Columbia, Canada

Magnus Palmblad, Associate Editor, Leiden University Medical Center, Leiden, Netherlands orcid.org/0000-0002-5865-8994

Nikolai Slavov, Topic Editor, Northeastern University, Boston, Massachusetts, USA orcid.org/0000-0003-2035-1820

AUTHOR INFORMATION

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jproteome.2c00704

Notes

Views expressed in this editorial are those of the authors and not necessarily the views of the ACS.

REFERENCES

(1) Oberg, A. L.; Vitek, O. Statistical design of quantitative mass spectrometry-based proteomic experiments. *J. Proteome Res.* **2009**, *8* (5), 2144−2156.

(2) Welker, F.; Collins, M. J.; Thomas, J. A.; Wadsley, M.; Brace, S.; Cappellini, E.; Turvey, S. T.; Reguero, M.; Gelfo, J. N.; Kramarz, A.; et al. Ancient proteins resolve the evolutionary history of Darwin's South American ungulates. *Nature* **2015**, *522* (7554), 81−84. Welker, F.; Ramos-Madrigal, J.; Gutenbrunner, P.; Mackie, M.; Tiwary, S.; Rakownikow Jersie-Christensen, R.; Chiva, C.; Dickinson, M. R.; Kuhlwilm, M.; de Manuel, M.; et al. The dental proteome of Homo antecessor. *Nature* **2020**, *580* (7802), 235−238.

https://doi.org/10.1021/acs.jproteome.2c00704
J. Proteome Res. 2022, 21, 2837−2839