

## EuBIC-MS Annual report 2020-2021

The European Bioinformatics Community for mass spectrometry (EuBIC-MS) is the EuPA initiative for mass spectrometry (MS)-related bioinformatics. Its aim is to improve bioinformatics for MS-based research through the setup of community-driven dynamics, thereby improving collaboration, funding, publication, and training activities. Through its different activities, EuBIC tries to raise awareness for the benefits of open science and continuously supports open software. It was started in November 2015 (Vaudel, 2016) to bring together the MS and bioinformatics communities in an open, collaborative and constructive environment. The initiative is carried enthusiastically by the bioinformatics community and welcomes everyone willing to help.

EuBIC-MS organizes a dedicated yearly conference, bringing together researchers in the fields of mass spectrometry and bioinformatics, junior scientists and industry partners from across Europe. Every year, this conference alternates between a Winter School with keynote talks and workshops, and a Developers' Meeting consisting of collaborative project sessions with the participants (hackathons). Additionally, EuBIC-MS actively contributes to bioinformatics hubs, and provides various workshops and presentations during international MS (e.g. annual conference of American Society of Mass Spectrometry) and proteomics (e.g. Human Proteome Organization congress, EuPA congress) conferences.

Here is the report of our activities and achievements between September 2019 and October 2021.

### **Winter Schools and Developers' Meetings**

The EuBIC-MS Winter Schools on computational MS are sponsored by the European Proteomics Association (EuPA) and several MS companies. They bring together scientists from both academia and industry to present and discuss their research in workshops, keynote lectures, flash talks and poster presentations. We hosted two Winter Schools, the first one in Semmering, Austria in 2017 (Willems et al., 2017) and the second in Zakopane, Poland in 2019 (Kopczynski

## EuBIC-MS Annual report 2020-2021

et al., 2019). Due to the Corona pandemic, there was no Winter School organised in 2021 but we are currently planning the next one in 2022.

The Developers' Meetings is an event dedicated to computer scientists and developers, in the field of bioinformatics applied to MS, where they can discuss and work together in an open and constructive spirit. The program is split between keynote lectures and multiple hackathon sessions where the participants develop bioinformatics tools and resources that address outstanding needs in the bioinformatics community and among biologists using MS data. Our first Developers' Meeting took place in Ghent, Belgium in 2018 (Willems et al., 2018).

The EuBIC Developers' Meeting 2020 took place in Nyborg (Denmark) January 13<sup>th</sup> to 20<sup>th</sup>. The first day was educational. Then, six keynote speakers gave lectures that were followed by 3.5 days of collaborative work. The keynote speakers were: Eric Deutsch (Institute for Systems Biology, Seattle, WA, USA), Ole N. Jensen (University of Southern Denmark, Odense, DK), Andy Jones (Institute of Integrative Biology, University of Liverpool, Liverpool, UK), Lydie Lane (Swiss Institute of Bioinformatics, CMU, Geneva, CH), Alexander Peltzer (Quantitative Biology Center, University of Tübingen, Tübingen, GE), Olga Vitek (Northeastern University, Boston, MA, USA). These lectures were recorded and are available on our website and on youtube ([youtube.com/playlist?list=PL\\_6fafgzU1nHd13qqjm3uNvyHh62JpPcB](https://youtube.com/playlist?list=PL_6fafgzU1nHd13qqjm3uNvyHh62JpPcB)).

Six workshop projects were selected by the community (abstracts available at [github.com/EuBIC/EuBIC2020/issues](https://github.com/EuBIC/EuBIC2020/issues)). These resulted in concrete outputs for the community:

- Updated version of ThermoRawFile parser (Hulstaert et al., 2020) with two additional modules: (i) an automated creation of extracted ion chromatograms (XICs) from raw files, and (ii) retrieving individual spectra from a raw file and reporting these in PROXI format.
- Development of automated workflows to facilitate more reproducible proteomics data analysis in Cytoscape (Shannon et al., 2003).
- Development of MegaGO (Verschaffelt et al., 2021), a tool to assess functional (gene ontology) similarity across meta-omics data sets.
- Development of Universal Spectrum Explorer (USE) (Schmidt et al., 2021), an online tool to compare predicted spectra with experimental spectra from proteomics repositories,

## EuBIC-MS Annual report 2020-2021

including the possibility to extract these spectra via the Universal Spectrum Identifier (USI) (Deutsch et al., 2020).

- Development of PhosFake, a tool to simulate the full experimental data acquisition of a phosphoproteomics experiment ([github.com/veitveit/PhosFake](https://github.com/veitveit/PhosFake)).
- Development of a workflow to benchmark different algorithms for assembling consensus spectra ([github.com/statisticalbiotechnology/representative-spectra-benchmark](https://github.com/statisticalbiotechnology/representative-spectra-benchmark)).

The EuBIC Developers' meeting 2020 was a success. The next one is currently in preparation and should take place in Locarno (Switzerland) in 2023. Before that, a Winter School is in preparation for March 2022.

### Community achievements

#### *Special issue perspective in Rapid communications in Mass Spectrometry*

We published an article that provides an overview of EuBIC-MS's achievements and perspectives in a Special issue of Rapid Communications in Mass Spectrometry (Bittremieux et al., 2021).

#### *The SDRF-Proteomics format*

The EuBIC-MS community worked together with the Human Proteome Organization Proteomics Standards Initiative (HUPO-PSI) (Deutsch et al., 2017) to define a standard for metadata annotation of proteomics data: The Sample and Data Relationship Format for Proteomics file format (SDRF-Proteomics) (Dai et al., 2021). This should increase data interpretability, as well as facilitating data reuse in an automated fashion. We remain very active in the development of this format ([github.com/bigbio/proteomics-metadata-standard](https://github.com/bigbio/proteomics-metadata-standard)) and intend to promote it through presentations and workshops (for example at the Proteomics Forum 2021).

#### *EuBIC seminars 2021*

EuBIC-MS organised a series of seminars that took place at the University of Copenhagen the 6<sup>th</sup> of October 2021. It was supported by the Center for Health Data Science (HeaDS - [heads.ku.dk/](https://heads.ku.dk/))

## EuBIC-MS Annual report 2020-2021

and the Novo Nordisk Foundation Center for Protein Research (CPR). Five speakers were selected to present their work and promote EuBIC-MS to Copenhagen researchers and students. The information regarding this meeting can be found here: [eubic-ms.org/events/copenhagen-2021-seminar/](https://eubic-ms.org/events/copenhagen-2021-seminar/).

### ***Our website and resources***

In partnership with the EuPA Educational Committee (EC), the Proteomics Academy web resource (<https://www.proteomics-academy.org>) has become the central communication portal for EuBIC since its creation in 2015. To increase the visibility of EuBIC itself, we also maintain a website dedicated to the initiative at <https://eubic-ms.org>. There, people find all information regarding past and upcoming EuBIC-MS activities, including this annual report.

We recently made a short video that promotes EuBIC-MS. It is available here.

### **How to get involved?**

The EuBIC community is open to everybody working in computational proteomics and mass spectrometry. You can join us by sending an email at [info@eubic-ms.org](mailto:info@eubic-ms.org) or by using the contact form available from the eubic-ms.org website. We will then send you an invitation to join our Slack workspace (<http://eubic.slack.com>), which is our extensively used platform to interact with other EuBIC members. Additionally, you can follow and retweet our latest activities from our Twitter account ([@EuBIC\\_ms](https://twitter.com/EuBIC_ms)). By joining us, you will be part of a community effort, and thus contribute to the sustainability of these scientific activities, which we expect to foster the development of the computational MS field.

### **Acknowledgments**

The EuBIC organizing committee would like to thank the European Proteomics Association for its support and funding since the creation of the initiative. We would also like to underscore the remarkable work accomplished by all organizers of the previous Winter Schools and Developers' Meetings.

### References

- Bittremieux W, Bouyssié D, Dorfer V, Locard-Paulet M, Perez-Riverol Y, Schwämmle V, Uszkoreit J, Van Den Bossche T. The European Bioinformatics Community for Mass Spectrometry (EuBIC-MS): an open community for bioinformatics training and research. *Rapid Commun Mass Spectrom*. 2021 Apr 16:e9087. doi: 10.1002/rcm.9087. Epub ahead of print. PMID: 33861485.
- Dai C, Füllgrabe A, Pfeuffer J, Solovyeva EM, Deng J, Moreno P, Kamatchinathan S, Kundu DJ, George N, Fexova S, Grüning B, Föll MC, Griss J, Vaudel M, Audain E, Locard-Paulet M, Turewicz M, Eisenacher M, Uszkoreit J, Van Den Bossche T, Schwämmle V, Webel H, Schulze S, Bouyssié D, Jayaram S, Duggineni VK, Samaras P, Wilhelm M, Choi M, Wang M, Kohlbacher O, Brazma A, Papatheodorou I, Bandeira N, Deutsch EW, Vizcaíno JA, Bai M, Sachsenberg T, Levitsky LI, Perez-Riverol Y. A proteomics sample metadata representation for multiomics integration and big data analysis. *Nat Commun*. 2021 Oct 6;12(1):5854. doi: 10.1038/s41467-021-26111-3. PMID: 34615866.
- Deutsch EW, Orchard S, Binz PA, Bittremieux W, Eisenacher M, Hermjakob H, Kawano S, Lam H, Mayer G, Menschaert G, Perez-Riverol Y, Salek RM, Tabb DL, Tenzer S, Vizcaíno JA, Walzer M, Jones AR. Proteomics Standards Initiative: Fifteen Years of Progress and Future Work. *J Proteome Res*. 2017 Dec 1;16(12):4288-4298. doi: 10.1021/acs.jproteome.7b00370. Epub 2017 Sep 15. PMID: 28849660; PMCID: PMC5715286.
- Deutsch, E.W., Bandeira, N., Sharma, V., Perez-Riverol, Y., Carver, J.J., Kundu, D.J., García-Seisdedos, D., Jarnuczak, A.F., Hewapathirana, S., Pullman, B.S., Wertz, J., Sun, Z., Kawano, S., Okuda, S., Watanabe, Y., Hermjakob, H., MacLean, B., MacCoss, M.J., Zhu, Y., Ishihama, Y., Vizcaíno, J.A., 2019. The ProteomeXchange consortium in 2020: enabling 'big data' approaches in proteomics. *Nucleic Acids Res*. <https://doi.org/10.1093/nar/gkz984>
- Deutsch, E.W., Perez-Riverol, Y., Carver, J., et al. Universal Spectrum Identifier for mass spectra. *bioRxiv*.2020. <https://doi.org/10.1101/2020.12.07.415539>
- Hulstaert N, Shofstahl J, Sachsenberg T, Walzer M, Barsnes H, Martens L, Perez-Riverol Y. ThermoRawFileParser: Modular, Scalable, and Cross-Platform RAW File Conversion. *J Proteome Res*. 2020 Jan 3;19(1):537-542. doi: 10.1021/acs.jproteome.9b00328. Epub 2019 Dec 6. PMID: 31755270; PMCID: PMC7116465.
- Kopczynski, D., Bittremieux, W., Bouyssié, D., Dorfer, V., Locard-Paulet, M., Van Puyvelde, B., Schwämmle, V., Soggiu, A., Willems, S., Uszkoreit, J., 2019. Proceedings of the EuBIC Winter School 2019. *EuPA Open Proteomics*. <https://doi.org/10.1016/j.euprot.2019.07.002>
- Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S., Kundu, D.J., Inuganti, A., Griss, J., Mayer, G., Eisenacher, M., Pérez, E., Uszkoreit, J., Pfeuffer, J., Sachsenberg, T., Yilmaz, Ş., Tiwary, S., Cox, J., Audain, E., Walzer, M., Jarnuczak, A.F., Ternent, T., Brazma, A., Vizcaíno, J.A., 2019. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res*. 47, D442–D450. <https://doi.org/10.1093/nar/gky1106>
- Schmidt T, Samaras P, Dorfer V, Panse C, Kockmann T, Bichmann L, van Puyvelde B, Perez-Riverol Y, Deutsch EW, Kuster B, Wilhelm M. Universal Spectrum Explorer: A

## EuBIC-MS Annual report 2020-2021

- Standalone (Web-)Application for Cross-Resource Spectrum Comparison. *J Proteome Res.* 2021 Jun 4;20(6):3388-3394. doi: 10.1021/acs.jproteome.1c00096. Epub 2021 May 10. PMID: 33970638.
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 2003 Nov;13(11):2498-504. doi: 10.1101/gr.1239303. PMID: 14597658; PMCID: PMC403769.
- Vaudel, M., 2016. The EuPA Bioinformatics Community (EuBIC) initiative. *EuPA Open Proteomics* 11, 34. <https://doi.org/10.1016/j.euprot.2016.03.009>
- Verschaffelt P, Van Den Bossche T, Gabriel W, Burdukiewicz M, Soggiu A, Martens L, Renard BY, Schiebenhoefer H, Mesuere B. MegaGO: A Fast Yet Powerful Approach to Assess Functional Gene Ontology Similarity across Meta-Omics Data Sets. *J Proteome Res.* 2021 Apr 2;20(4):2083-2088. doi: 10.1021/acs.jproteome.0c00926. Epub 2021 Mar 4. PMID: 33661648.
- Willems, S., Bouyssié, D., David, M., Locard-Paulet, M., Mechtler, K., Schwämmle, V., Uszkoreit, J., Vaudel, M., Dorfer, V., 2017. Proceedings of the EuBIC Winter School 2017. *J. Proteomics* 161, 78–80. <https://doi.org/10.1016/j.jprot.2017.04.001>
- Willems, S., Bouyssié, D., Deforce, D., Dorfer, V., Gorshkov, V., Kopczynski, D., Laukens, K., Locard-Paulet, M., Schwämmle, V., Uszkoreit, J., Valkenborg, D., Vaudel, M., Bittremieux, W., 2018. Proceedings of the EuBIC developer's meeting 2018. *J. Proteomics* 187, 25–27. <https://doi.org/10.1016/j.jprot.2018.05.015>