# GNPS Dashboard: collaborative exploration of mass spectrometry data in the web browser

To the Editor — The importance of providing open and transparent solutions to use, share, inspect and reproduce mass spectrometry (MS) data analysis is becoming more apparent. To address these challenges, we developed a collaborative interactive web application, the GNPS Dashboard (https://gnps-lcms.ucsd.edu), to enable remote and synchronous collaborative research in a common analysis environment for MS data analysis (similar to collaborative text editing in browser-based word processors, such as Google Docs).

The disruptions to office and laboratory workspaces resulting from the COVID-19 pandemic, including campus closures, telework and stay-at-home orders, increased the need for online collaborative approaches to perform scientific research. Most solutions for the analysis of MS data are geared for exploration on a local workstation that lacks synchronous collaborative data exploration, requiring specific software packages, knowledge of data storage locations, familiarity with file transfer protocols and often conversion of data into compatible formats[1–4]. Commercial data exploration solutions usually lack universality and are costly, creating additional hurdles to data accessibility, preventing collaborators, reviewers and readers of scientific publications from inspecting MS data to verify interpretations. However, commercial software offers deep access to proprietary vendor-specific data formats and unique data analysis capabilities, for example, creation of private databases. More broadly, as scientific data are becoming more complex, scientific reproducibility is a challenge. Therefore, there is a growing mandate by both funding bodies and journals to make mass spectrometry data (and other data) adherent to FAIR (findable, accessible, interoperable and reusable).

To empower remote collaborations or classroom teaching, the GNPS Dashboard includes leader–follower synchronization (real-time updates from one lead user) and fully collaborative synchronization (real-time updates from multiple users). Followers can disconnect the synchronization to continue the analysis from where the leader left off without needing to reload the data. The GNPS Dashboard leader–follower paradigm has already been used in classroom teaching by at least five institutions, including undergraduate institutions (Supplementary Note 1 and 2). The fully collaborative synchronization enables multi-user visualization and data exploration similar to online synchronous collaborative
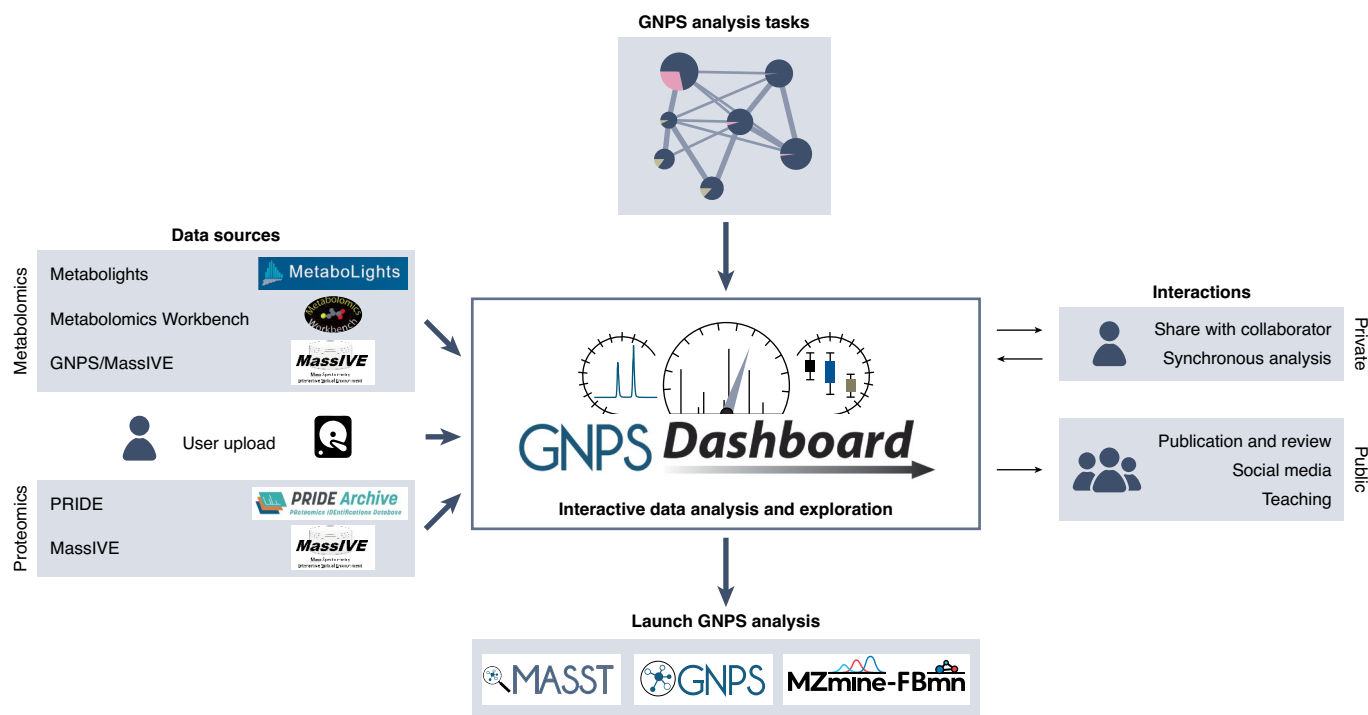


**Fig. 1 | Overview of the GNPS Dashboard in the data analysis ecosystem.** The GNPS Dashboard has been integrated into the web-accessible MS ecosystem. The top and left panels show how the GNPS Dashboard can ingress datasets from public metabolomics resources: MetaboLights, Metabolomics Workbench and GNPS/MassIVE; and from proteomics resources: PRIDE and MassIVE, private user uploads and private GNPS analysis tasks. The bottom panel shows that users can launch downstream analysis on their data directly out of the GNPS Dashboard. Finally, the right panel shows how the GNPS Dashboard and its visualizations can interact with the wider community with reproducible URL or QR code link outs from any analysis and teaching/synchronous collaboration modes for real-time interactivity.

document editing. For example, users can initiate a collaborative session with two or more people on any web-accessible device (providing a dashboard collaboration start link, instructions, video tutorial and video application). A snapshot of the collaborative work is automatically created and can be shared with collaborators or included in publications. The final state of the analysis is saved together with every discrete action, enabling inspection of the evolution of data analysis.

The GNPS Dashboard includes tools that facilitate the exploration of liquid and gas chromatography–mass spectrometry data for collaborative examination, and hands-on teaching of MS concepts using private and publicly available MS data, including files stored in the MS data repositories GNPS/MassIVE[5], MetaboLights[6], ProteomeXchange[7] and Metabolomics Workbench[8] (Fig. 1; providing a link to instructions and video tutorial). All public data from compatible repositories can be selected using the GNPS dataset explorer (https://gnps-explorer.ucsd.edu/). Files not deposited in public MS repositories can be uploaded through a drag-and-drop option for file transfer. mzXML, mzML, CDF and raw formats are compatible with GNPS Dashboard. Through deep linking from the GNPS platform, the GNPS Dashboard serves as a data explorer and hub for further data analysis, including Molecular Networking[5], gas chromatography–mass spectrometry deconvolution[9], in silico annotation and Mass Spectrometry Search Tool (MASST[10]) (Supplementary Note 2).

The GNPS Dashboard's visualization tool enables inspection of total ion chromatograms, retention time versus m/z heat map, extracted ion chromatograms, tandem mass spectra for inspection/visualization of individual compounds and quantitative comparison of the peak abundances of two groups as boxplots — all with publication-quality figures. The dashboard aids peer review of scientific manuscripts and inspection of public quantitative MS data to validate published results. Beyond visualization and analysis of MS data, the dashboard supports the development of other bioinformatics tools that do not have their own web-enabled user interfaces. Documentation and tutorials are available in Supplementary Note 1 and Supplementary Fig. 1.

The GNPS Dashboard encodes the visualization in a shareable URL, which empowers users to share the exact same visualization with colleagues, thus reducing miscommunication and improving data transparency, for example, during remote meetings. Every visualization

and analysis result can be shared via a URL that will re-launch the original data visualization on their device, along with the history of the analysis (up to 1,000 steps per session). Users can share these links with collaborators and embed them in publications, presentations or social media posts (for example, Example Tweet). The final visualization can also be shared as a quick response (QR) code. Anyone with a link or QR code can build on the analysis and re-share their additions. Links and QR codes will remain valid for data that have been archived in a public repository.

Together, the capabilities of the GNPS Dashboard should help improve MS data accessibility and interoperability, lower barriers of entry for MS data analysis in the research environment and the classroom, encourage data transparency and sharing, and strengthen the reproducibility of MS data analysis.

## Code availability

While the GNPS Dashboard is accessible as a free public web service, it is possible to locally install the GNPS Dashboard to function with local data sources, making collaborative analysis and sharing possible privately within an institution when necessary. The source code is available through the GNPS web environment and GitHub, enabling quick installation on local servers. The GNPS Dashboard source code can be found on GitHub: https://github.com/mwang87/GNPS_LCMSDashboard under a modified UCSD BSD License. The GNPS Dataset Explorer source code can be found on GitHub: https://github.com/mwang87/GNPS_DatasetExplorer under an MIT License. Further implementation details can be found in Supplementary Methods. ❐

Daniel Petras [ID][1,2,3], Vanessa V. Phelan[4], Deepa Acharya[5], Andrew E. Allen [ID][2,6], Allegra T. Aron[1], Nuno Bandeira [ID][7], Benjamin P. Bowen [ID][8], Deirdre Belle-Oudry[9], Simon Boecker [ID][10], Dale A. Cummings Jr.[11,12], Jessica M. Deutsch [ID][13], Eoin Fahy [ID][14], Neha Garg[13], Rachel Gregor[15], Jo Handelsman[5], Mirtha Navarro-Hoyos[16], Alan K. Jarmusch[1,17], Scott A. Jarmusch [ID][18], Katherine Louie[8], Katherine N. Maloney[19], Michael T. Marty [ID][9], Michael M. Meijler [ID][20,21], Itzhak Mizrahi [ID][21,22], Rachel L. Neve [ID][23], Trent R. Northen [ID][8], Carlos Molina-Santiago [ID][24], Morgan Panitchpakdi[1], Benjamin Pullman [ID][7], Aaron W. Puri[11,12], Robin Schmid[1], Shankar Subramaniam[14], Monica Thukral[2,6], Felipe Vasquez-Castro[25], Pieter C. Dorrestein [ID][1,2,26,27] and Mingxun Wang [ID][1,27 ✉]

[1]Collaborative Mass Spectrometry Innovation Center, Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego,

La Jolla, CA, USA. [2]Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, USA. [3]CMFI Cluster of Excellence, Interfaculty Institute of Microbiology and Medicine, University of Tübingen, Tübingen, Germany. [4]Department of Pharmaceutical Sciences, Skaggs School of Pharmacy and Pharmaceutical Sciences, University of Colorado, Anschutz Medical Campus, Aurora, CO, USA. [5]Wisconsin Institute for Discovery and Department of Plant Pathology, University of Wisconsin-Madison, Madison, WI, USA. [6]Environmental Genomics, J. Craig Venter Institute, La Jolla, CA, USA. [7]Center for Computational Mass Spectrometry, Department of Computer Science and Engineering, University of California, San Diego, La Jolla, CA, USA. [8]DOE Joint Genome Institute and Lawrence Berkeley National Laboratory, Berkeley, CA, USA. [9]Department of Chemistry and Biochemistry, University of Arizona, Tucson, AZ, USA. [10]Analysis and Redesign of Biological Networks, Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg, Germany. [11]Department of Chemistry, University of Utah, Salt Lake City, UT, USA. [12]Henry Eyring Center for Cell & Genome Science, University of Utah, Salt Lake City, UT, USA. [13]School of Chemistry and Biochemistry, Center for Microbial Dynamics and Infection, Georgia Institute of Technology, Atlanta, GA, USA. [14]Department of Bioengineering, University of California San Diego, La Jolla, CA, USA. [15]Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA. [16]BIoactivity for Sustainable Development Group (BIODESS), Department of Chemistry, University of Costa Rica, San Jose, Costa Rica. [17]Immunity, Inflammation, and Disease Laboratory, Division of Intramural Research, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC, USA. [18]Department of Biotechnology and Biomedicine, Technical University of Denmark, Kongens Lyngby, Denmark. [19]Department of Chemistry, Point Loma Nazarene University, San Diego, CA, USA. [20]Department of Chemistry, Ben-Gurion University of the Negev, Be'er Sheva, Israel. [21]National Institute of Biotechnology in the Negev, Ben-Gurion University of the Negev, Be'er Sheva, Israel. [22]Department of Life Sciences, Ben-Gurion University of the Negev, Be'er Sheva, Israel. [23]Department of Immunology and Microbiology, University of Colorado, Anschutz Medical Campus, Aurora, CO, USA. [24]Instituto de Hortofruticultura Subtropical y Mediterránea, Universidad de Málaga-Consejo Superior de Investigaciones Científicas (IHSM-UMA-CSIC), Departamento de Microbiología, Universidad de Málaga, Málaga, Spain. [25]Centro Nacional de Innovaciones Biotecnologicas (CENIBiot), CeNAT-CONARE, 1174-1200 San Jose, Costa Rica. [26]Center for Microbiome Innovation, University of California San Diego, La Jolla, CA, USA. [27]Department of Pharmacology, School of Medicine, University of California San Diego, La Jolla, CA, USA.
✉e-mail: miw023@ucsd.edu

## References

1. Pluskal, T. et al. *BMC Bioinform.* **11**, 395 (2010).
2. Tsugawa, H. et al. *Nat. Methods* **12**, 523–526 (2015).
3. Röst, H. L. et al. *Nat. Methods* **13**, 741–748 (2016).
4. Huang, Y.-C. et al. *J. Cheminform.* **13**, 8 (2021).
5. Wang, M. et al. *Nat. Biotechnol.* **34**, 828–837 (2016).
6. Haug, K. et al. *Nucleic Acids Res.* **48**, D440–D444 (2020).
7. Vizcaíno, J. A. et al. *Nat. Biotechnol.* **32**, 223–226 (2014).
8. Sud, M. et al. *Nucleic Acids Res.* **44**, D463–D470 (2016).
9. Aksenov, A. A. et al. *Nat. Biotechnol.* **39**, 169–173 (2021).
10. Wang, M. et al. *Nat. Biotechnol.* **38**, 23–26 (2020).

## Author contributions

M.W. and D.P. conceived the project. M.W. developed the software for GNPS Dashboard and GNPS Dataset Explorer. M.W., D.P., V.V.P. and P.C.D. provided guidance and supervision. M.W., D.P., V.V.P., K.N.M., A.T.A., A.K.J., B.P., D.A.C.J., A.W.P., C.M.S., M.T., N.G., N.B. and P.C.D. tested and offered user feedback for the GNPS Dashboard. S.S. and E.F. implemented the APIs in Metabolomics Workbench to support GNPS Dashboard. T.N., B.B. and K.L. facilitated making metabolomics data in Genome Portal available in GNPS Dashboard. M.W., D.P., V.V.P., K.N.M., A.T.A., A.K.J., S.A.J., D.A., B.P., D.A.C.J., A.W.P., C.M.S., M.T., A.E.A., R.S., N.B., S.B., M.T.M., D.B.O., M.N.H., R.G., M.M.M., I.M., J.H., F.V.C., J.M.D., N.G., N.B., R.L.N. and P.C.D. wrote use cases for the GNPS Dashboard Supplemental Information. P.C.D. provided funding for the project. P.C.D., M.W., V.V.P. and D.P. wrote the draft manuscript. All authors edited and approved the final manuscript.

## Competing interests

P.C.D. is a scientific advisor of Sirenas, Galileo and Cybele, and scientific advisor and co-founder of Ometa Labs LLC and Enveda with approval by the UC San Diego. M.W. is a founder of Ometa Labs LLC. T.R.N. is an advisor of Brightseed Bio.

## Additional information

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s41592-021-01339-5.

**Peer review information** *Nature Methods* thanks Juan Antonio Vizcaíno, Jean-Luc Wolfender and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.