## *Title*

Interactive MS/MS Visualization with the Metabolomics Spectrum Resolver Web Service

## *Authors*

Mingxun Wang (1, *) Simon Rogers (2, *) Wout Bittremieux (1+7) Christopher Chen (1) Pieter C. Dorrestein (1) Emma L. Schymanski (3) Tobias Schulze (6) Steffen Neumann (4+5) Rene Meier (4)

* These authors contributed equally

## *Affiliations*

1. Skaggs School of Pharmacy and Pharmaceutical Sciences, UC San Diego, 9500 Gilman Dr. La Jolla CA 92093, USA.
2. School of Computing Science, University of Glasgow, Glasgow G12 8QQ, United Kingdom.
3. Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg, 6 avenue du Swing, 4367 Belvaux, Luxembourg.
4. Leibniz Institute of Plant Biochemistry, Bioinformatics and Scientific Data, Weinberg 3, 06120 Halle, Germany.
5. German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Deutscher Platz 5e, 04103 Leipzig, Germany
6. Helmholtz Centre for Environmental Research – UFZ, Department of Effect Directed Analysis, Permoserstrasse 15, 04318 Leipzig, Germany.
7. Department of Mathematics and Computer Science, University of Antwerp, Antwerp, Belgium

## *Abstract*

The growth of online mass spectrometry metabolomics resources, including data repositories, spectral library databases, and online analysis platforms has created an environment of online/web accessibility. Here, we introduce the Metabolomics Spectrum Resolver (https://metabolomics-usi.ucsd.edu/), a tool that builds upon these exciting developments to allow for consistent data export (in human and machine-readable forms) and publication-ready visualisations for tandem mass spectrometry spectra. This tool supports the draft Human Proteome Organizations Proteomics Standards Initiative's USI specification, which has been extended to deal with the metabolomics use cases. To date, this resource already supports data formats from GNPS, MassBank, MS2LDA, MassIVE, MetaboLights, and Metabolomics Workbench and is integrated into several of these resources, providing a valuable open source community contribution (https://github.com/mwang87/MetabolomicsSpectrumResolver).

## *Introduction*

The effective exchange and visualization of tandem mass spectrometry information across a variety of resources is important in communicating and consequently building confidence in both data quality and molecule identification throughout the research and publication process. The inclusion of MS/MS spectra in scientific posters, presentations, and manuscripts in a consistent manner is often challenging and labor intensive due to the variety of formats and resources available. It is also often difficult to balance both the ease of figure generation and the level of customization possible. On one hand, existing software such as the PDV[1], IPSA[2] and Lorikeet[3] proteomics data viewers as well as vendor software are able to draw MS/MS spectra, but are aimed at interactive visualization. Three key features are often missing: vector graphics (to ensure high resolution export), customization of spectral visualization, and high-throughput automated figure generation. On the other end of the spectrum are generic vector graphics editors such as Adobe Illustrator or Inkscape. While generic editors are powerful, they lack MS/MS specific features and require high levels of customization to achieve an image ready for publication.

We present here the Metabolomics Spectrum Resolver (https://metabolomics-usi.ucsd.edu/), a web tool that enables: 1) high-quality vector graphics drawing, 2) specialised mass spectrometry formatting, 3) integration with major metabolomics data repositories, 4) integration with common online analysis tools and 5) a programmatic web interface (API).

## ***Results and Discussion***

### *MS/MS Data Sources - Integration with Community Resources*

This web tool integrates several technologies: spectrum_utils[4] for highly customizable spectrum processing/drawing, draft universal spectrum identifiers adapted for Metabolomics inspired by the Human Proteome Organization - Proteome Standards Initiative working group Universal Standards Identifier (USI)[5] (http://www.psidev.info/usi), and web APIs supported by online metabolomics data services. The current resource is able to resolve the following metabolomics online services (**Table 1**):

1) Data repositories: Metabolomics Workbench[6], MetaboLights[7], MassIVE[8]
2) Reference spectral library resources: MassBank[9], MoNA (https://mona.fiehnlab.ucdavis.edu/), GNPS[10], and MS2LDA.org MOTIFDB[11],
3) online informatics pipelines: GNPS Molecular Networking/Library Search/MASST/FBMN[12], and MS2LDA.org.[13]

Additionally, several resources (MassBank, GNPS, and MS2LDA.org) have already integrated links to Metabolomics Spectrum Resolver to facilitate the creation of images ready for publication directly from web resources at the time of analysis.

Further, the Metabolomics Spectrum Resolver supports the linking back to the original data resources to allow users to explore the original context of the MS/MS spectra. Due to the web-integrated nature of Metabolomics Spectrum Resolver, after MS/MS publication in manuscripts (generally meant for human consumption), it is straightforward to hyperlink back to the Metabolomics Spectrum Resolver, enabling programmatic access to the underlying MS/MS

spectra in a machine readable fashion (text or JSON). This enables re-inspection by both humans and automated tools to re-evaluate identifications with newly available computational tools. Finally, with a programmatic web API interface, software engineers are able to integrate high-throughput figure creation in a language-agnostic fashion. Already, users in the community are using this resource to programmatically pull spectral data in a standardized fashion to facilitate figure generation[14].

Finally, the Metabolomics Spectrum Resolver uses spectrum identifiers inspired by the Human Proteome Organizations Proteomics Standards Initiative's USI specification[15]. The Metabolomics Spectrum Resolver supports USI formats in the formal specification. Further, the USI specifications have been extended (https://github.com/mwang87/MetabolomicsSpectrumResolver/blob/master/README.md#usi-extended-for-metabolomics---formatting-documentation) to ensure compatibility with a wide range of (at this stage still very heterogeneous) metabolomics resources.

*Visualization Capabilities*

The Metabolomics Spectrum Resolver can be accessed at https://metabolomics-usi.ucsd.edu/ and enables users to plot a single MS/MS spectrum (**Figure 1a**) and mirror matches between two MS/MS spectra to show their similarity (e.g. between measured spectra and library matches) (**Figure 1b**). MS/MS spectra drawing can be modified by specifying mass ranges, intensity ranges, customizable peak labeling, figure size, decimal points for mass values, and an optional grid. The resulting drawing can be downloaded as an SVG (vector), PNG (raster), TSV (tab separated peaks), and JSON (machine readable peaks) (**Figure 2**).

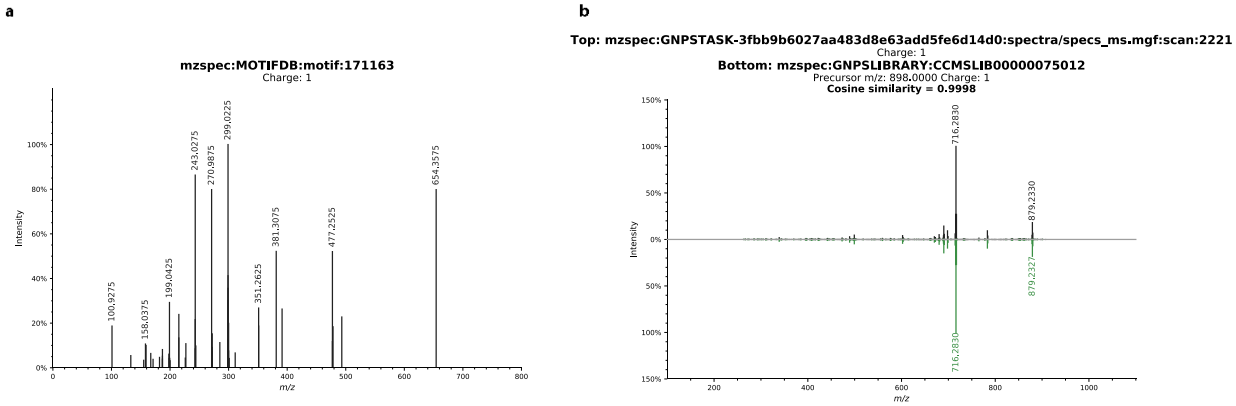


***Figure 1 - Example MS/MS*** *figure generation of (a) a single MS/MS spectrum and (b) a mirror plot.*

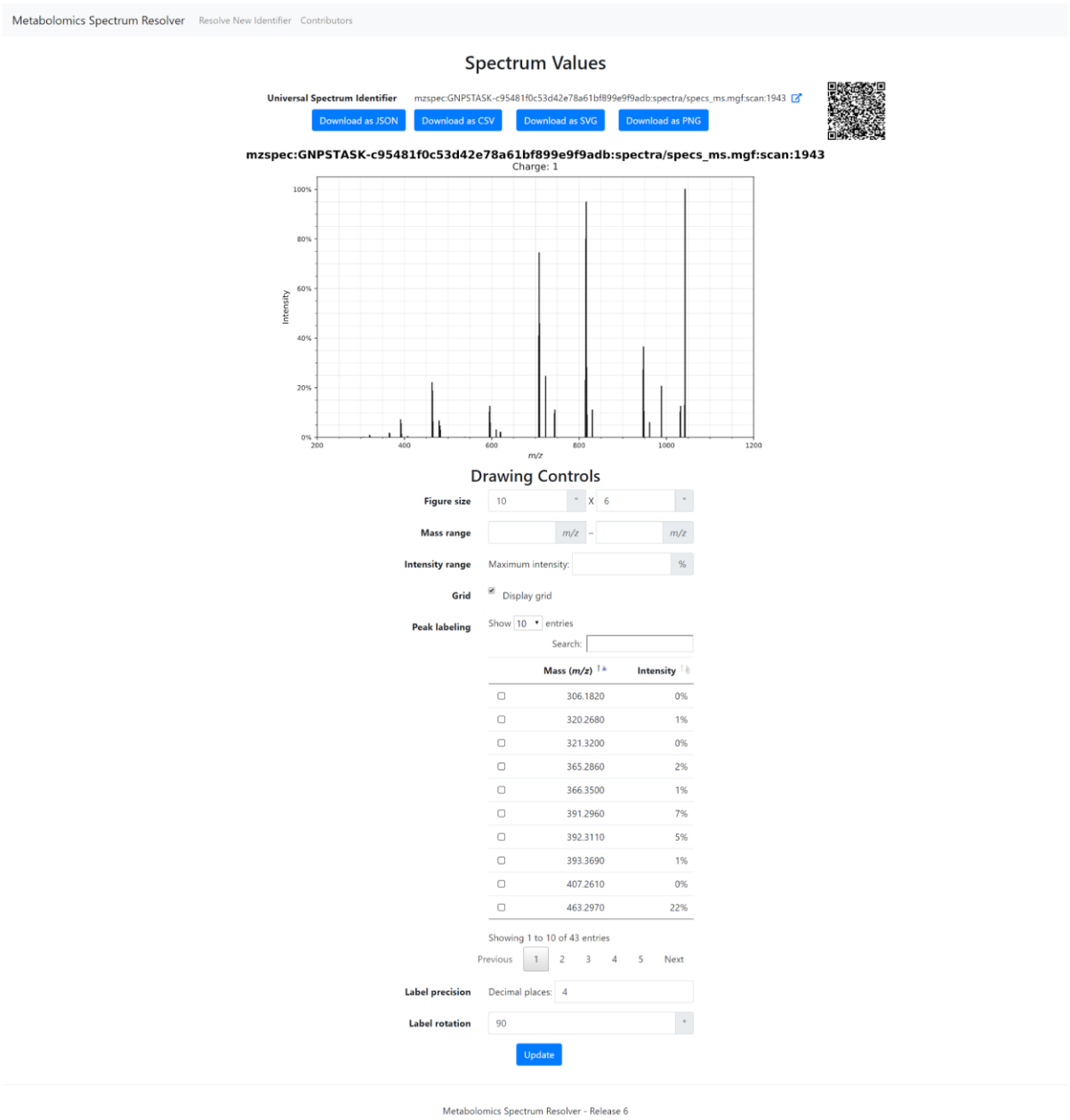


***Figure 2 - Interactive User Interface with Drawing Options*** - Enables customizing the spectrum drawing, linking back to original spectrum data via URL and a generated QR code, spectrum downloads, and programmatic data download.

**Table 1 - Examples of the Metabolomics Spectrum Resolver on various public data**

| Resource | Accession (if applicable) | Example URL |
|---|---|---|
| MassBank | SM858102 | Example |
| GNPS Spectral Library | CCMSLIB00005436077 | Example |
| GNPS Analysis Spectrum | N/A | Example |
| MS2LDA MotifDB | MOTIFDB:motif:171163 | Example |
| MS2LDA Analysis Spectrum | N/A | Example |
| MassIVE/GNPS Repository Spectrum | MSV000078547 | Example |
| Metabolights Repository Spectrum | MTBLS38 | Example |
| Metabolomics Workbench Repository Spectrum | ST000003 | Example |

## *Conclusion*

The Metabolomics Spectrum Resolver has been designed to improve the presentation and accessibility of metabolomics data, both within these supported resources and for the community. The broad support of key online metabolomics resources, programmatic API access, and high quality spectrum drawing make the resource highly accessible and of clear practical benefit to the community and we hope that many in the community will find this useful.

## *Source Code*

Source code can be found here with the MIT license:
https://github.com/mwang87/MetabolomicsSpectrumResolver

## *Acknowledgements*

We thank Madeleine Ernst, Alan Jarmusch, and Louis-Felix Nothias for their feedback on the usability of the Metabolomics Spectrum Resolver. ELS acknowledges funding support from the Luxembourg National Research Fund (FNR) for project A18/BM/12341006. MassBank Europe (https://massbank.eu/MassBank) is supported by the NORMAN Association (https://www.norman-network.net), de.NBI (FKZ 031L0107), NaToxAq (Marie Sklodowska-Curie grant agreement No. 722493), HBM4EU (European Union H2020 grant agreement No. 733032) and the Helmholtz Centre for Environmental Research - UFZ (https://www.ufz.de/index.php?en=33573)

## *Conflict of Interests*

Mingxun Wang is the founder of Ometa Labs LLC, Pieter C. Dorrestein is on the scientific advisory board of Sirenas and Cybele Microbiome

## *References*


(1) Li, K.; Vaudel, M.; Zhang, B.; Ren, Y.; Wen, B. PDV: An Integrative Proteomics Data Viewer. *Bioinformatics* **2019**, *35* (7), 1249–1251.

(2) Brademan, D. R.; Riley, N. M.; Kwiecien, N. W.; Coon, J. J. Interactive Peptide Spectral Annotator: A Versatile Web-Based Tool for Proteomic Applications. *Mol. Cell. Proteomics* **2019**, *18* (8 suppl 1), S193–S201.

(3) Lorikeet by vagisha https://uwpr.github.io/Lorikeet/ (accessed Apr 27, 2020).

(4) Bittremieux, W. Spectrum_utils: A Python Package for Mass Spectrometry Data Processing and Visualization. *Anal. Chem.* **2020**, *92* (1), 659–661.

(5) Deutsch, E. W.; Orchard, S.; Binz, P.-A.; Bittremieux, W.; Eisenacher, M.; Hermjakob, H.; Kawano, S.; Lam, H.; Mayer, G.; Menschaert, G.; Perez-Riverol, Y.; Salek, R. M.; Tabb, D. L.; Tenzer, S.; Vizcaíno, J. A.; Walzer, M.; Jones, A. R. Proteomics Standards Initiative: Fifteen Years of Progress and Future Work. *J. Proteome Res.* **2017**, *16* (12), 4288–4298.

(6) Sud, M.; Fahy, E.; Cotter, D.; Azam, K.; Vadivelu, I.; Burant, C.; Edison, A.; Fiehn, O.; Higashi, R.; Nair, K. S.; Sumner, S.; Subramaniam, S. Metabolomics Workbench: An International Repository for Metabolomics Data and Metadata, Metabolite Standards, Protocols, Tutorials and Training, and Analysis Tools. *Nucleic Acids Res.* **2016**, *44* (D1), D463–D470.

(7) Kale, N. S.; Haug, K.; Conesa, P.; Jayseelan, K.; Moreno, P.; Rocca-Serra, P.; Nainala, V. C.; Spicer, R. A.; Williams, M.; Li, X.; Salek, R. M.; Griffin, J. L.; Steinbeck, C. MetaboLights: An Open-Access Database Repository for Metabolomics Data. *Curr. Protoc. Bioinformatics* **2016**, *53*, 14.13.1–14.13.18.

(8) Wang, M.; Wang, J.; Carver, J.; Pullman, B. S.; Cha, S. W.; Bandeira, N. Assembling the Community-Scale Discoverable Human Proteome. *Cell Syst* **2018**, *7* (4), 412–421.e5.

(9) Horai, H.; Arita, M.; Kanaya, S.; Nihei, Y.; Ikeda, T.; Suwa, K.; Ojima, Y.; Tanaka, K.; Tanaka, S.; Aoshima, K.; Oda, Y.; Kakazu, Y.; Kusano, M.; Tohge, T.; Matsuda, F.; Sawada, Y.; Hirai, M. Y.; Nakanishi, H.; Ikeda, K.; Akimoto, N.; Maoka, T.; Takahashi, H.; Ara, T.; Sakurai, N.; Suzuki, H.; Shibata, D.; Neumann, S.; Iida, T.; Tanaka, K.; Funatsu, K.; Matsuura, F.; Soga, T.; Taguchi, R.; Saito, K.; Nishioka, T. MassBank: A Public Repository for Sharing Mass Spectral Data for Life Sciences. *J. Mass Spectrom.* **2010**, *45* (7), 703–714.

(10) Wang, M.; Carver, J. J.; Phelan, V. V.; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D. D.; Watrous, J.; Kapono, C. A.; Luzzatto-Knaan, T.; Porto, C.; Bouslimani, A.; Melnik, A. V.; Meehan, M. J.; Liu, W.-T.; Crüsemann, M.; Boudreau, P. D.; Esquenazi, E.; Sandoval-Calderón, M.; Kersten, R. D.; Pace, L. A.; Quinn, R. A.; Duncan, K. R.; Hsu, C.-C.; Floros, D. J.; Gavilan, R. G.; Kleigrewe, K.; Northen, T.; Dutton, R. J.; Parrot, D.; Carlson, E. E.; Aigle, B.; Michelsen, C. F.; Jelsbak, L.; Sohlenkamp, C.; Pevzner, P.; Edlund, A.; McLean, J.; Piel, J.; Murphy, B. T.; Gerwick, L.; Liaw, C.-C.; Yang, Y.-L.; Humpf, H.-U.; Maansson, M.; Keyzers, R. A.; Sims, A. C.; Johnson, A. R.; Sidebottom, A. M.; Sedio, B. E.; Klitgaard, A.; Larson, C. B.; P, C. A. B.; Torres-Mendoza, D.; Gonzalez, D. J.; Silva, D. B.; Marques, L. M.; Demarque, D. P.; Pociute, E.; O'Neill, E. C.; Briand, E.; Helfrich, E. J. N.; Granatosky, E. A.; Glukhov, E.; Ryffel, F.; Houson, H.; Mohimani, H.; Kharbush, J. J.; Zeng, Y.; Vorholt, J. A.; Kurita, K. L.; Charusanti, P.; McPhail, K. L.; Nielsen, K. F.; Vuong, L.; Elfeki, M.; Traxler, M. F.; Engene, N.; Koyama, N.; Vining, O. B.; Baric, R.; Silva, R. R.; Mascuch, S. J.; Tomasi, S.; Jenkins, S.; Macherla, V.; Hoffman, T.; Agarwal, V.; Williams,

P. G.; Dai, J.; Neupane, R.; Gurr, J.; Rodríguez, A. M. C.; Lamsa, A.; Zhang, C.; Dorrestein, K.; Duggan, B. M.; Almaliti, J.; Allard, P.-M.; Phapale, P.; Nothias, L.-F.; Alexandrov, T.; Litaudon, M.; Wolfender, J.-L.; Kyle, J. E.; Metz, T. O.; Peryea, T.; Nguyen, D.-T.; VanLeer, D.; Shinn, P.; Jadhav, A.; Müller, R.; Waters, K. M.; Shi, W.; Liu, X.; Zhang, L.; Knight, R.; Jensen, P. R.; Palsson, B. O.; Pogliano, K.; Linington, R. G.; Gutiérrez, M.; Lopes, N. P.; Gerwick, W. H.; Moore, B. S.; Dorrestein, P. C.; Bandeira, N. Sharing and Community Curation of Mass Spectrometry Data with Global Natural Products Social Molecular Networking. *Nat. Biotechnol.* **2016**, *34* (8), 828–837.

(11) Rogers, S.; Ong, C. W.; Wandy, J.; Ernst, M.; Ridder, L.; van der Hooft, J. J. J. Deciphering Complex Metabolite Mixtures by Unsupervised and Supervised Substructure Discovery and Semi-Automated Annotation from MS/MS Spectra. *Faraday Discuss.* **2019**, *218* (0), 284–302.

(12) Wang, M.; Jarmusch, A. K.; Vargas, F.; Aksenov, A. A.; Gauglitz, J. M.; Weldon, K.; Petras, D.; da Silva, R.; Quinn, R.; Melnik, A. V.; van der Hooft, J. J. J.; Caraballo-Rodríguez, A. M.; Nothias, L. F.; Aceves, C. M.; Panitchpakdi, M.; Brown, E.; Di Ottavio, F.; Sikora, N.; Elijah, E. O.; Labarta-Bajo, L.; Gentry, E. C.; Shalapour, S.; Kyle, K. E.; Puckett, S. P.; Watrous, J. D.; Carpenter, C. S.; Bouslimani, A.; Ernst, M.; Swafford, A. D.; Zúñiga, E. I.; Balunas, M. J.; Klassen, J. L.; Loomba, R.; Knight, R.; Bandeira, N.; Dorrestein, P. C. Mass Spectrometry Searches Using MASST. *Nat. Biotechnol.* **2020**, *38* (1), 23–26.

(13) Wandy, J.; Zhu, Y.; van der Hooft, J. J. J.; Daly, R.; Barrett, M. P.; Rogers, S. Ms2lda.org: Web-Based Topic Modelling for Substructure Discovery in Mass Spectrometry. *Bioinformatics* **2018**, *34* (2), 317–318.

(14) Ernst, M.; Rogers, S.; Lausten-Thomsen, U.; Bjorkbom, A.; Svane Laursen, S.; Courraud, J.; Borglum, A.; Nordentoft, M.; Werge, T.; Bo Mortensen, P.; Hougaard, D. M.; Cohen, A. S. Gestational-Age-Dependent Development of the Neonatal Metabolome. *medRxiv*, 2020. https://doi.org/10.1101/2020.03.27.20045534.

(15) Universal Spectrum Identifier | HUPO Proteomics Standards Initiative http://www.psidev.info/usi (accessed Apr 27, 2020).