

1 **Title**

2

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

4

5 **Authors**

6

7 Mingxun Wang (1, \*) Simon Rogers (2, \*) Wout Bittremieux (1+7) Christopher Chen (1) Pieter C.  
8 Dorrestein (1) Emma L. Schymanski (3) Tobias Schulze (6) Steffen Neumann (4+5) Rene Meier  
9 (4)

10

11 \* These authors contributed equally

12

13 **Affiliations**

14

- 15 1. Skaggs School of Pharmacy and Pharmaceutical Sciences, UC San Diego, 9500  
16 Gilman Dr. La Jolla CA 92093, USA.
- 17 2. School of Computing Science, University of Glasgow, Glasgow G12 8QQ, United  
18 Kingdom.
- 19 3. Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg, 6  
20 avenue du Swing, 4367 Belvaux, Luxembourg.
- 21 4. Leibniz Institute of Plant Biochemistry, Bioinformatics and Scientific Data, Weinberg 3,  
22 06120 Halle, Germany.
- 23 5. German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig,  
24 Deutscher Platz 5e, 04103 Leipzig, Germany
- 25 6. Helmholtz Centre for Environmental Research – UFZ, Department of Effect Directed  
26 Analysis, Permoserstrasse 15, 04318 Leipzig, Germany.
- 27 7. Department of Mathematics and Computer Science, University of Antwerp, Antwerp,  
28 Belgium

29

30 **Abstract**

31

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

43

44 **Introduction**

45  
46 The effective exchange and visualization of tandem mass spectrometry information across a  
47 variety of resources is important in communicating and consequently building confidence in both  
48 data quality and molecule identification throughout the research and publication process. The  
49 inclusion of MS/MS spectra in scientific posters, presentations, and manuscripts in a consistent  
50 manner is often challenging and labor intensive due to the variety of formats and resources  
51 available. It is also often difficult to balance both the ease of figure generation and the level of  
52 customization possible. On one hand, existing software such as the PDV<sup>1</sup>, IPISA<sup>2</sup> and Lorikeet<sup>3</sup>  
53 proteomics data viewers as well as vendor software are able to draw MS/MS spectra, but are  
54 aimed at interactive visualization. Three key features are often missing: vector graphics (to ensure  
55 high resolution export), customization of spectral visualization, and high-throughput automated  
56 figure generation. On the other end of the spectrum are generic vector graphics editors such as  
57 Adobe Illustrator or Inkscape. While generic editors are powerful, they lack MS/MS specific  
58 features and require high levels of customization to achieve an image ready for publication.

59  
60 We present here the Metabolomics Spectrum Resolver (<https://metabolomics-usi.ucsd.edu/>), a  
61 web tool that enables: 1) high-quality vector graphics drawing, 2) specialised mass spectrometry  
62 formatting, 3) integration with major metabolomics data repositories, 4) integration with common  
63 online analysis tools and 5) a programmatic web interface (API).

64

## 65 **Results and Discussion**

66

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

68

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

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

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

83

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

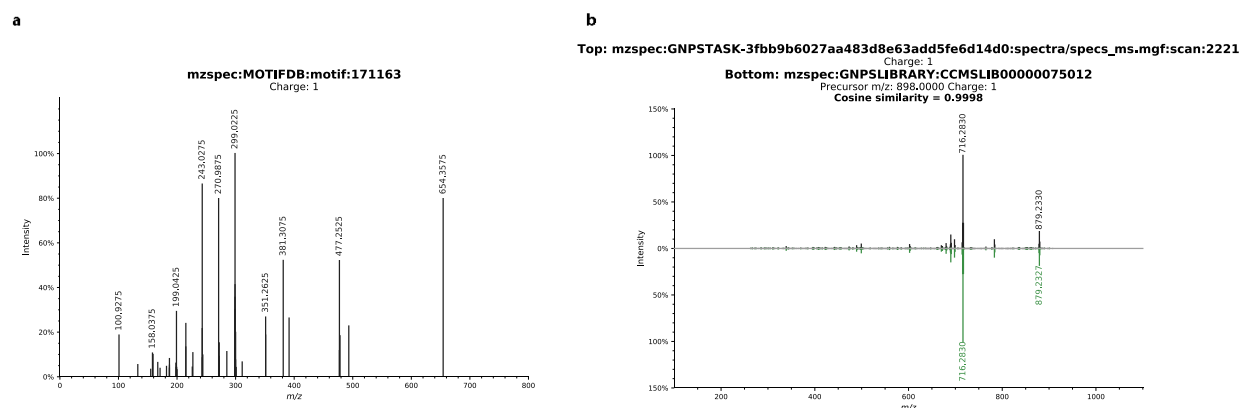
89 spectra in a machine readable fashion (text or JSON). This enables re-inspection by both humans  
90 and automated tools to re-evaluate identifications with newly available computational tools.  
91 Finally, with a programmatic web API interface, software engineers are able to integrate high-  
92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

95  
96 Finally, the Metabolomics Spectrum Resolver uses spectrum identifiers inspired by the Human  
97 Proteome Organizations Proteomics Standards Initiative's USI specification<sup>15</sup>. The Metabolomics  
98 Spectrum Resolver supports USI formats in the formal specification. Further, the USI  
99 specifications have been extended  
100 (<https://github.com/mwang87/MetabolomicsSpectrumResolver/blob/master/README.md#usi-extended-for-metabolomics---formatting-documentation>) to ensure compatibility with a wide  
101 range of (at this stage still very heterogeneous) metabolomics resources.  
102

### 103 104 *Visualization Capabilities*

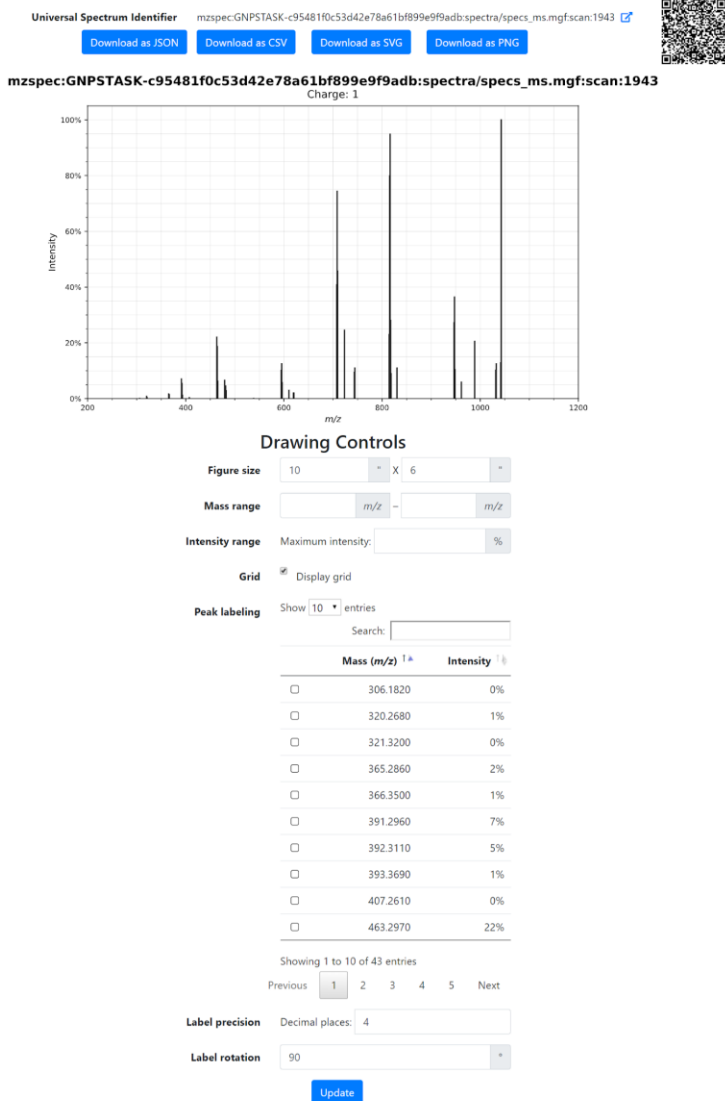
105  
106 The Metabolomics Spectrum Resolver can be accessed at <https://metabolomics-usi.ucsd.edu/>  
107 and enables users to plot a single MS/MS spectrum (**Figure 1a**) and mirror matches between two  
108 MS/MS spectra to show their similarity (e.g. between measured spectra and library matches)  
109 (**Figure 1b**). MS/MS spectra drawing can be modified by specifying mass ranges, intensity  
110 ranges, customizable peak labeling, figure size, decimal points for mass values, and an optional  
111 grid. The resulting drawing can be downloaded as an SVG (vector), PNG (raster), TSV (tab  
112 separated peaks), and JSON (machine readable peaks) (**Figure 2**).

113  
114  
115



116  
117 **Figure 1 - Example MS/MS figure generation of (a) a single MS/MS [spectrum](#) and (b) a [mirror](#)**  
118 **[plot](#).**  
119

## Spectrum Values



120  
121  
122  
123  
124

**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.

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

126

Resource	Accession (if applicable)	Example URL
MassBank	SM858102	<a href="#">Example</a>
GNPS Spectral Library	CCMSLIB00005436077	<a href="#">Example</a>
GNPS Analysis Spectrum	N/A	<a href="#">Example</a>
MS2LDA MotifDB	MOTIFDB:motif:171163	<a href="#">Example</a>
MS2LDA Analysis Spectrum	N/A	<a href="#">Example</a>
MassIVE/GNPS Repository Spectrum	MSV000078547	<a href="#">Example</a>
Metabolights Repository Spectrum	MTBLS38	<a href="#">Example</a>
Metabolomics Workbench Repository Spectrum	ST000003	<a href="#">Example</a>

127

### 128 **Conclusion**

129

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

135

### 136 **Source Code**

137

138 Source code can be found here with the MIT license:

139 <https://github.com/mwang87/MetabolomicsSpectrumResolver>

140

### 141 **Acknowledgements**

142

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

151

### 152 **Conflict of Interests**

153

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

156

## 157 **References**

- 158 (1) Li, K.; Vaudel, M.; Zhang, B.; Ren, Y.; Wen, B. PDV: An Integrative Proteomics Data  
159 Viewer. *Bioinformatics* **2019**, *35* (7), 1249–1251.
- 160 (2) Brademan, D. R.; Riley, N. M.; Kwiecien, N. W.; Coon, J. J. Interactive Peptide Spectral  
161 Annotator: A Versatile Web-Based Tool for Proteomic Applications. *Mol. Cell. Proteomics*  
162 **2019**, *18* (8 suppl 1), S193–S201.
- 163 (3) Lorikeet by vagisha <https://uwpr.github.io/Lorikeet/> (accessed Apr 27, 2020).
- 164 (4) Bittremieux, W. Spectrum\_utils: A Python Package for Mass Spectrometry Data Processing  
165 and Visualization. *Anal. Chem.* **2020**, *92* (1), 659–661.
- 166 (5) Deutsch, E. W.; Orchard, S.; Binz, P.-A.; Bittremieux, W.; Eisenacher, M.; Hermjakob, H.;  
167 Kawano, S.; Lam, H.; Mayer, G.; Menschaert, G.; Perez-Riverol, Y.; Salek, R. M.; Tabb, D.  
168 L.; Tenzer, S.; Vizcaíno, J. A.; Walzer, M.; Jones, A. R. Proteomics Standards Initiative:  
169 Fifteen Years of Progress and Future Work. *J. Proteome Res.* **2017**, *16* (12), 4288–4298.
- 170 (6) Sud, M.; Fahy, E.; Cotter, D.; Azam, K.; Vadivelu, I.; Burant, C.; Edison, A.; Fiehn, O.;  
171 Higashi, R.; Nair, K. S.; Sumner, S.; Subramaniam, S. Metabolomics Workbench: An  
172 International Repository for Metabolomics Data and Metadata, Metabolite Standards,  
173 Protocols, Tutorials and Training, and Analysis Tools. *Nucleic Acids Res.* **2016**, *44* (D1),  
174 D463–D470.
- 175 (7) Kale, N. S.; Haug, K.; Conesa, P.; Jayseelan, K.; Moreno, P.; Rocca-Serra, P.; Nainala, V.  
176 C.; Spicer, R. A.; Williams, M.; Li, X.; Salek, R. M.; Griffin, J. L.; Steinbeck, C.  
177 MetaboLights: An Open-Access Database Repository for Metabolomics Data. *Curr. Protoc.*  
178 *Bioinformatics* **2016**, *53*, 14.13.1–14.13.18.
- 179 (8) Wang, M.; Wang, J.; Carver, J.; Pullman, B. S.; Cha, S. W.; Bandeira, N. Assembling the  
180 Community-Scale Discoverable Human Proteome. *Cell Syst* **2018**, *7* (4), 412–421.e5.
- 181 (9) Horai, H.; Arita, M.; Kanaya, S.; Nihei, Y.; Ikeda, T.; Suwa, K.; Ojima, Y.; Tanaka, K.;  
182 Tanaka, S.; Aoshima, K.; Oda, Y.; Kakazu, Y.; Kusano, M.; Tohge, T.; Matsuda, F.;  
183 Sawada, Y.; Hirai, M. Y.; Nakanishi, H.; Ikeda, K.; Akimoto, N.; Maoka, T.; Takahashi, H.;  
184 Ara, T.; Sakurai, N.; Suzuki, H.; Shibata, D.; Neumann, S.; Iida, T.; Tanaka, K.; Funatsu, K.;  
185 Matsuura, F.; Soga, T.; Taguchi, R.; Saito, K.; Nishioka, T. MassBank: A Public Repository  
186 for Sharing Mass Spectral Data for Life Sciences. *J. Mass Spectrom.* **2010**, *45* (7), 703–  
187 714.
- 188 (10) Wang, M.; Carver, J. J.; Phelan, V. V.; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D.  
189 D.; Watrous, J.; Kaponov, C. A.; Luzzatto-Knaan, T.; Porto, C.; Bouslimani, A.; Melnik, A. V.;  
190 Meehan, M. J.; Liu, W.-T.; Crüsemann, M.; Boudreau, P. D.; Esquenazi, E.; Sandoval-  
191 Calderón, M.; Kersten, R. D.; Pace, L. A.; Quinn, R. A.; Duncan, K. R.; Hsu, C.-C.; Floros,  
192 D. J.; Gavilan, R. G.; Kleigrew, K.; Northen, T.; Dutton, R. J.; Parrot, D.; Carlson, E. E.;  
193 Aigle, B.; Michelsen, C. F.; Jelsbak, L.; Sohlenkamp, C.; Pevzner, P.; Edlund, A.; McLean,  
194 J.; Piel, J.; Murphy, B. T.; Gerwick, L.; Liaw, C.-C.; Yang, Y.-L.; Humpf, H.-U.; Maansson,  
195 M.; Keyzers, R. A.; Sims, A. C.; Johnson, A. R.; Sidebottom, A. M.; Sedio, B. E.; Klitgaard,  
196 A.; Larson, C. B.; P, C. A. B.; Torres-Mendoza, D.; Gonzalez, D. J.; Silva, D. B.; Marques,  
197 L. M.; Demarque, D. P.; Pociute, E.; O'Neill, E. C.; Briand, E.; Helfrich, E. J. N.;  
198 Granatosky, E. A.; Glukhov, E.; Ryffel, F.; Houson, H.; Mohimani, H.; Kharbush, J. J.; Zeng,  
199 Y.; Vorholt, J. A.; Kurita, K. L.; Charusanti, P.; McPhail, K. L.; Nielsen, K. F.; Vuong, L.;  
200 Elfeki, M.; Traxler, M. F.; Engene, N.; Koyama, N.; Vining, O. B.; Baric, R.; Silva, R. R.;  
201 Mascuch, S. J.; Tomasi, S.; Jenkins, S.; Macherla, V.; Hoffman, T.; Agarwal, V.; Williams,



- 202 P. G.; Dai, J.; Neupane, R.; Gurr, J.; Rodríguez, A. M. C.; Lamsa, A.; Zhang, C.; Dorrestein,  
203 K.; Duggan, B. M.; Almaliti, J.; Allard, P.-M.; Phapale, P.; Nothias, L.-F.; Alexandrov, T.;  
204 Litaudon, M.; Wolfender, J.-L.; Kyle, J. E.; Metz, T. O.; Peryea, T.; Nguyen, D.-T.; VanLeer,  
205 D.; Shinn, P.; Jadhav, A.; Müller, R.; Waters, K. M.; Shi, W.; Liu, X.; Zhang, L.; Knight, R.;  
206 Jensen, P. R.; Palsson, B. O.; Pogliano, K.; Linington, R. G.; Gutiérrez, M.; Lopes, N. P.;  
207 Gerwick, W. H.; Moore, B. S.; Dorrestein, P. C.; Bandeira, N. Sharing and Community  
208 Curation of Mass Spectrometry Data with Global Natural Products Social Molecular  
209 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.  
210 Deciphering Complex Metabolite Mixtures by Unsupervised and Supervised Substructure  
211 Discovery and Semi-Automated Annotation from MS/MS Spectra. *Faraday Discuss.* **2019**,  
212 *218* (0), 284–302.
- (12) Wang, M.; Jarmusch, A. K.; Vargas, F.; Aksenov, A. A.; Gauglitz, J. M.; Weldon, K.;  
213 Petras, D.; da Silva, R.; Quinn, R.; Melnik, A. V.; van der Hooft, J. J. J.; Caraballo-  
214 Rodríguez, A. M.; Nothias, L. F.; Aceves, C. M.; Panitchpakdi, M.; Brown, E.; Di Ottavio, F.;  
215 Sikora, N.; Elijah, E. O.; Labarta-Bajo, L.; Gentry, E. C.; Shalapour, S.; Kyle, K. E.; Puckett,  
216 S. P.; Watrous, J. D.; Carpenter, C. S.; Bouslimani, A.; Ernst, M.; Swafford, A. D.; Zúñiga,  
217 E. I.; Balunas, M. J.; Klassen, J. L.; Loomba, R.; Knight, R.; Bandeira, N.; Dorrestein, P. C.  
218 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.  
219 Ms2lda.org: Web-Based Topic Modelling for Substructure Discovery in Mass Spectrometry.  
220 *Bioinformatics* **2018**, *34* (2), 317–318.
- (14) Ernst, M.; Rogers, S.; Lausten-Thomsen, U.; Bjorkbom, A.; Svane Laursen, S.;  
221 Courraud, J.; Borglum, A.; Nordentoft, M.; Werge, T.; Bo Mortensen, P.; Hougaard, D. M.;  
222 Cohen, A. S. Gestational-Age-Dependent Development of the Neonatal Metabolome.  
223 *medRxiv*, 2020. <https://doi.org/10.1101/2020.03.27.20045534>.
- (15) Universal Spectrum Identifier | HUPO Proteomics Standards Initiative  
224 <http://www.psidev.info/usi> (accessed Apr 27, 2020).