Electronic Supplementary Material (ESI) for Natural Product Reports.

The Year 2020 in Natural Product Bioinformatics: an overview of the latest tools and databases

Marnix H. Medema[1]

1. Bioinformatics Group, Wageningen University, Wageningen, The Netherlands

Table S1: Overview of natural product-related bioinformatic tools and databases published in 2020

Category / Name	*Brief description*	*URL of web server or code*	*Reference*	*Open license?*
Chemical structure databases				
NP Atlas	database of bacterial and fungal natural product structures	https://www.npatlas.org/	[1]	X
Streptome-DB 3.0	database of streptomycete natural product structures	http://www.pharmbioinf.uni-freiburg.de/streptomedb	[2]	X
CyanoMetDB	database of cyanobacterial natural product structures	N/A	[3]	X
MacrolactoneDB	database of macrolactones and their bioactivity data	https://macrolact.collaborationspharma.com/	[4]	
NORINE (updated)	database of nonribosomal peptides	https://bioinfo.lifl.fr/norine/	[5]	X
COCONUT	aggegrated collection of natural products from open databases	https://coconut.naturalproducts.net/	[7]	X
Cheminformatic tools				
NPClassifier	neural-network-based structural classification of natural products	https://npclassifier.ucsd.edu/	[8]	X
NC-MFP	natural compound molecular fingerprint	-	[9]	X
MAP4	MinHashed Atom Pair fingerprint	https://tm.gdb.tools/map4/npatlas_map_tmap/	[10]	X
DECIMER	deep learning algorithm for chemical image recognition	https://github.com/Kohulan/DECIMER-Java	[11]	X
Biosynthetic gene cluster identification & analysis				
EvoMining	evolution-based natural product genome mining	https://github.com/nselem/evomining	[14]	X
CO-OCCUR	gene cluster identification based on gene co-occurrence	https://github.com/egluckthaler/co-occur	[15]	X
SeMPI 2.0	prediction and annotation of PKS and NRPS biosynthetic gene clusters	http://sempi.pharmazie.uni-freiburg.de/	[13]	
TOUCAN	identification of fungal biosynthetic gene clusters	https://github.com/bioinfoUQAM/TOUCAN	[16]	X
DeepRiPP	identification of RiPP gene clusters and linking to mass spectrometry data	http://deepripp.magarveylab.ca/	[18]	
RRE-Finder	identification of RiPP recognition elements	https://github.com/Alexamk/RREFinder	[19]	X
decRiPPter	identification of candidate novel RiPP classes	https://decrippter.bioinformatics.nl/	[20]	X
RiPPER	genome mining of RiPP biosynthetic pathways	https://github.com/streptomyces/ripper	[21]	X
RODEO (update)	genome mining of RiPP biosynthetic pathways	http://ripp.rodeo/	[22]	X
Biosynthetic diversity analysis				
BiG-SCAPE	sequence similarity network analysis of biosynthetic gene clusters	https://bigscape-corason.secondarymetabolites.org/	[23]	X
CORASON	multi-locus phylogenies of biosynthetic genes and gene clusters	https://bigscape-corason.secondarymetabolites.org/	[23]	X

BiG-SLICE	clustering algorithm for large sets of biosynthetic gene clusters	https://github.com/medema-group/bigslice	[24]	X
BiG-FAM	database of biosynthetic gene cluster families	https://bigfam.bioinformatics.nl/	[25]	X
clinker	visualization tool for comparing biosynthetic gene clusters	https://github.com/gamcil/clinker	[27]	X
cblaster	biosynthetic gene cluster similarity search engine	https://github.com/gamcil/cblaster	[26]	X
Biosynthetic gene cluster(-related) databases				
MIBiG 2.0	database of experimentally characterized biosynthetic gene clusters	https://mibig.secondarymetabolites.org/	[28]	X
antiSMASH-DB 3.0	database of predicted bacterial, archaeal and fungal biosynthetic gene clusters	https://antismash-db.secondarymetabolites.org/	[29]	X
IMG-ABC 5.0	database of predicted bacterial and archaeal biosynthetic gene clusters	https://img.jgi.doe.gov/abc-public/	[30]	
Prospect Fungi	database of predicted fungal biosynthetic gene clusters	http://prospect-fungi.com/	[31]	
ActDES	curated database of actinobacterial genomes for genome mining	https://github.com/nselem/ActDES	[32]	X
Target-based genome mining and function prediction				
ARTS 2.0	resistance-based genome mining of natural product biosynthetic gene clusters	https://arts3.ziemertlab.com/	[33]	X
TargetMining	identification of polyketide BGCs with self-resistance genes	https://github.com/GerganaVandova/TargetMining	[34]	X
bgc_tran	function prediction based on transporter genes	https://github.com/nickbhat/bgc_tran	[35]	X
Chemprop	neural-network-based prediction of antibiotic activity	http://chemprop.csail.mit.edu/	[60]	X
Chemical structure prediction				
PRISM 4	prediction of natural product structures and activities from genome data	https://prism.adapsyn.com/	[12]	
DDAP	prediction of polyketide synthase docking domain interactions	https://tylii.github.io/ddap	[36]	X
PKSpop	prediction of polyketide synthase docking domain interactions	https://github.com/yanwang271/PKSpop	[37]	X
SBSPKS 3	prediction of macrocyclization of nonribosomal peptides and polyketides	http://www.nii.ac.in/sbspks3.html	[38]	X
AdenylPred	prediction of substrate specificities for class I adenylate-forming enzymes	https://srobinson.shinyapps.io/AdenylPred/	[39]	X
NMR data analysis				
SMART 2.0	deep learning tool to generate structure hypotheses from 13C and 1H NMR data	https://smart.ucsd.edu/	[40]	
np-pred	prediction of common natural product families from 13C NMR data	https://github.com/saulhazelius/np-pred	[41]	X
DP4-AI	automatic processing and assignment of 13C and 1H NMR data	https://github.com/KristapsE/DP4-AI	[42]	X
MixONat	dereplication of mixtures based on 13C NMR data	https://sourceforge.net/projects/mixonat	[43]	X
Mass spectrometry data analysis				
ZODIAC	molecular formula prediction from mass spectrometry data	https://bio.informatik.uni-jena.de/software/zodiac/	[44]	X
Retip	retention time prediction for liquid chromatography mass spectrometry analysis	https://www.retip.app/	[46]	X
MetFID	predicting compound fingerprints from MS/MS data	-	[47]	
NRPro	analysis of MS/MS spectra of peptidic natural products	https://bioinfo.cristal.univ-lille.fr/nrpro/	[57]	X
MASST	similarity search engine for MS/MS spectra	https://masst.ucsd.edu/	[48]	X

Spec2Vec	mass-spectral similarity scoring through learning of structural relationships	https://github.com/iomega/spec2vec	[49]	X
FBMN	feature-based molecular networking	https://ccms-ucsd.github.io/GNPSDocumentation/featurebasedmolecularnetworking/	[50]	X
GNPS GC-MS EI Data Analysis	GC-MS data analysis using molecular networking	https://ccms-ucsd.github.io/GNPSDocumentation/gcanalysis/	[51]	X
ReDU	re-analysis of public MS/MS datasets	https://redu.ucsd.edu/	[52]	X
CANOPUS	compound class annotation from mass spectrometry data	https://github.com/boecker-lab/sirius-libs	[45]	X
Qemistree	tandem mass spectrometry molecular standards database	https://qemistree.ucsd.edu/	[53]	X
BMDMS-NP	database of natural product ESI MS/MS spectra	https://github.com/chalbori/bmdms-np	[54]	X
METLIN (update)	tandem mass spectrometry molecular standards database	http://metlin.scripps.edu/	[55]	
Linking MS data to structures and gene clusters				
CycloNovo	de novo cyclopeptide analysis in tandem mass spectrometry data	https://github.com/bbehsaz/cyclonovo	[56]	X
MetaMiner	integration of genomics and metabolomics to link RiPPs to their biosynthetic genes	https://github.com/mohimanilab/MetaMiner	[58]	X
NPLinker	metabologenomic correlation tool to link biosynthetic gene clusters to molecules	https://github.com/sdrogers/nplinker	[59]	X
Synthetic chemistry				
Chematica / Synthia (update)	Computational planning of natural product synthesis	https://www.sigmaaldrich.com/chemistry/chemical-synthesis/synthesis-software.html	[61]	

References

1 J. A. van Santen, G. Jacob, A. L. Singh, V. Aniebok, M. J. Balunas, D. Bunsko, F. C. Neto, L. Castaño-Espriu, C. Chang, T. N. Clark, J. L. Cleary Little, D. A. Delgadillo, P. C. Dorrestein, K. R. Duncan, J. M. Egan, M. M. Galey, F. P. J. Haeckl, A. Hua, A. H. Hughes, D. Iskakova, A. Khadilkar, J.-H. Lee, S. Lee, N. LeGrow, D. Y. Liu, J. M. Macho, C. S. McCaughey, M. H. Medema, R. P. Neupane, T. J. O'Donnell, J. S. Paula, L. M. Sanchez, A. F. Shaikh, S. Soldatou, B. R. Terlouw, T. A. Tran, M. Valentine, J. J. J. van der Hooft, D. A. Vo, M. Wang, D. Wilson, K. E. Zink and R. G. Linington, *ACS Cent Sci*, 2019, **5**, 1824–1833.

2 A. F. A. Moumbock, M. Gao, A. Qaseem, J. Li, P. A. Kirchner, B. Ndingkokhar, B. D. Bekono, C. V. Simoben, S. B. Babiaka, Y. I. Malange, F. Sauter, P. Zierep, F. Ntie-Kang and S. Günther, *Nucleic Acids Res*, 2020, **49**, D600-D604.

3 M. R. Jones, E. Pinto, M. A. Torres, F. Dörr, H. Mazur-Marzec, K. Szubert, L. Tartaglione, C. Dell'Aversano, C. O. Miles, D. G. Beach, P. McCarron, K. Sivonen, D. P. Fewer, J. Jokela and E. M. -L. Janssen, *BioRxiv*, 2020, doi:10.1101/2020.04.16.038703.

4 P. P. K. Zin, G. J. Williams and S. Ekins, *Sci. Rep.*, 2020, **10**, 6284.

5 A. Flissi, E. Ricart, C. Campart, M. Chevalier, Y. Dufresne, J. Michalik, P. Jacques, C. Flahaut, F. Lisacek, V. Leclère and M. Pupin, *Nucleic Acids Res.*, 2020, **48**, D465–D469.

6 E. Ricart, V. Leclère, A. Flissi, M. Mueller, M. Pupin and F. Lisacek, *J. Cheminform.*, 2019, **11**, 13.

7 M. Sorokina and C. Steinbeck, *Journal of Cheminformatics*, 2020, **12**, 20.

8 H. Kim, M. Wang, C. Leber, L.-F. Nothias, R. Reher, K. B. Kang, J. J. J. van der Hooft, P. Dorrestein, W. Gerwick and G. Cottrell, 2020, *ChemRxiv*, doi:10.26434/chemrxiv.12885494.v1.

9 M. Seo, H. K. Shin, Y. Myung, S. Hwang and K. T. No, *Journal of Cheminformatics*, 2020, **12**, 6.

10 A. Capecchi, D. Probst and J.-L. Reymond, *Journal of Cheminformatics*, 2020, **12**, 43

11 K. Rajan, A. Zielesny and C. Steinbeck, *Journal of Cheminformatics*, 2020, **12**, 65.

12 M. A. Skinnider, C. W. Johnston, M. Gunabalasingam, N. J. Merwin, A. M. Kieliszek, R. J. MacLellan, H. Li, M. R. M. Ranieri, A. L. H. Webster, M. P. T. Cao, A. Pfeifle, N. Spencer, Q. H. To, D. P. Wallace, C. A. Dejong and N. A. Magarvey, *Nat. Commun.*, 2020, **11**, 6058.
13 P. F. Zierep, A. T. Ceci, I. Dobrusin, S. C. Rockwell-Kollmann and S. Günther, *Metabolites*, 2021, **11**, 13.
14 N. Sélem-Mojica, C. Aguilar, K. Gutiérrez-García, C. E. Martínez-Guerrero and F. Barona-Gómez, *Microb Genom*, 2020, **5**, e000260.
15 E. Gluck-Thaler, S. Haridas, M. Binder, I. V. Grigoriev, P. W. Crous, J. W. Spatafora, K. Bushley and J. C. Slot, *Mol. Biol. Evol.*, 2020, **37**, 2838–2856.
16 H. Almeida, S. Palys, A. Tsang and A. B. Diallo, *NAR Genomics and Bioinformatics*, 2020, **2**, lqaa098.
18 N. J. Merwin, W. K. Mousa, C. A. Dejong, M. A. Skinnider, M. J. Cannon, H. Li, K. Dial, M. Gunabalasingam, C. Johnston and N. A. Magarvey, *Proc. Natl. Acad. Sci. U. S. A.*, 2020, **117**, 371–380.
19 A. M. Kloosterman, K. E. Shelton, G. P. van Wezel, M. H. Medema and D. A. Mitchell, *mSystems*, 2020, **5**, e00267-20.
20 A. M. Kloosterman, P. Cimermancic, S. S. Elsayed, C. Du, M. Hadjithomas, M. S. Donia, M. A. Fischbach, G. P. van Wezel and M. H. Medema, *PLoS Biol.*, 2020, **18**, e3001026.
21 J. Santos-Aberturas, G. Chandra, L. Frattaruolo, R. Lacret, T. H. Pham, N. M. Vior, T. H. Eyles and A. W. Truman, *Nucleic Acids Res.*, 2019, **47**, 4624–4637.
22 M. A. Georgiou, S. R. Dommaraju, X. Guo and D. A. Mitchell, 2020, *ACS Chem Biol*, **15**, 2976-2985.
23 J. C. Navarro-Muñoz, N. Selem-Mojica, M. W. Mullowney, S. Kautsar, J. H. Tryon, E. I. Parkinson, E. L. C. De Los Santos, M. Yeong, P. Cruz-Morales, S. Abubucker, A. Roeters, W. Lokhorst, A. Fernandez-Guerra, L. T. D. Cappelini, R. J. Thomson, W. W. Metcalf, N. L. Kelleher, F. Barona-Gomez and M. H. Medema, *Nat Chem Biol*, 2020, **16**, 60-68.
24 S. A. Kautsar, J. J. J. van der Hooft, D. de Ridder and M. H. Medema, *GigaScience*, 2021, **10**, giaa154.
25 S. A. Kautsar, K. Blin, S. Shaw, T. Weber and M. H. Medema, *Nucleic Acids Res.*, 2020, **49**, D490-D497.
26 C. L. M. Gilchrist, T. J. Booth and Y.-H. Chooi, *BioRxiv*, 2020, doi:10.1101/2020.11.08.370601.
27 C. L. M. Gilchrist and Y.-H. H. Chooi, *Bioinformatics*, 2021, doi:10.1093/bioinformatics/btab007.
28 S. A. Kautsar, K. Blin, S. Shaw, J. C. Navarro-Muñoz, B. R. Terlouw, J. J. J. van der Hooft, J. A. van Santen, V. Tracanna, H. G. Suarez Duran, V. Pascal Andreu, N. Selem-Mojica, M. Alanjary, S. L. Robinson, G. Lund, S. C. Epstein, A. C. Sisto, L. K. Charkoudian, J. Collemare, R. G. Linington, T. Weber and M. H. Medema, *Nucleic Acids Res.*, 2020, **48**, D454–D458.
29 K. Blin, S. Shaw, S. A. Kautsar, M. H. Medema and T. Weber, *Nucleic Acids Res.*, 2020, **49**, D639-D643.
30 K. Palaniappan, I.-M. A. Chen, K. Chu, A. Ratner, R. Seshadri, N. C. Kyrpides, N. N. Ivanova and N. J. Mouncey, *Nucleic Acids Res.*, 2020, **48**, D422–D430.
31 M. T. Robey, L. K. Caesar, M. T. Drott, N. P. Keller and N. L. Kelleher, *BioRxiv*, 2020, doi:10.1101/2020.09.21.307157.
32 J. K. Schniete, N. Selem-Mojica, A. S. Birke, P. Cruz-Morales, I. S. Hunter, F. Barona-Gómez and P. A. Hoskisson, *Microb Genom*, 2021, doi:10.1099/mgen.0.000498.
33 M. D. Mungan, M. Alanjary, K. Blin, T. Weber, M. H. Medema and N. Ziemert, *Nucleic Acids Res.*, 2020, **48**, W546–W552.
34 G. A. Vandova, A. Nivina, C. Khosla, R. W. Davis, C. R. Fisher and M. E. Hillenmeyer, *BioRxiv*, 2020, doi:10.1101/2020.06.01.128595.
35 A. Crits-Christoph, N. Bhattacharya, M. R. Olm, Y. S. Song and J. F. Banfield, *Genome Res*, 2020, doi:10.1101/gr.268169.120.
36 T. Li, A. Tripathi, F. Yu, D. H. Sherman and A. Rao, *Bioinformatics*, 2020, **36**, 942–944.
37 Y. Wang, M. C. Marrero, M. H. Medema and A. D. J. van Dijk, *Bioinformatics*, 2020, **19**, 4846-4853.
38 P. Agrawal and D. Mohanty, *Bioinformatics*, 2020, doi:10.1093/bioinformatics/btaa851.
39 S. L. Robinson, B. R. Terlouw, M. D. Smith, S. J. Pidot, T. P. Stinear, M. H. Medema and L. P. Wackett, *J. Biol. Chem.*, 2020, **295**, 14826–14839.
40 R. Reher, H. W. Kim, C. Zhang, H. H. Mao, M. Wang, L.-F. Nothias, A. M. Caraballo-Rodriguez, E. Glukhov, B. Teke, T. Leao, K. L. Alexander, B. M. Duggan, E. L. Van Everbroeck, P. C. Dorrestein, G. W. Cottrell and W. H. Gerwick, *J. Am. Chem. Soc.*, 2020, **142**, 4114–4120.
41 S. H. Martínez-Treviño, V. Uc-Cetina, M. A. Fernández-Herrera and G. Merino, *Journal of*

Chemical Information and Modeling, 2020, **60**, 3376–3386.
42 A. Howarth, K. Ermanis and J. Goodman, *Chem Sci*, 2020, **11**, 4351-4359.
43 A. Bruguière, S. Derbré, J. Dietsch, J. Leguy, V. Rahier, Q. Pottier, D. Bréard, S. Suor-Cherer, G. Viault, A.-M. Le Ray, F. Saubion and P. Richomme, *Analytical Chemistry*, 2020, **92**, 8793–8801.
44 M. Ludwig, L.-F. Nothias, K. Dührkop, I. Koester, M. Fleischauer, M. A. Hoffmann, D. Petras, F. Vargas, M. Morsy, L. Aluwihare, P. C. Dorrestein and S. Böcker, *Nature Machine Intelligence*, 2020, **2**, 629–641.
45 K. Dührkop, L.-F. Nothias, M. Fleischauer, R. Reher, M. Ludwig, M. A. Hoffmann, D. Petras, W. H. Gerwick, J. Rousu, P. C. Dorrestein and S. Böcker, *Nat. Biotechnol.*, 2020, doi:10.1038/s41587-020-0740-8.
46 P. Bonini, T. Kind, H. Tsugawa, D. K. Barupal and O. Fiehn, *Anal. Chem.*, 2020, **92**, 7515–7522.
47 Z. Fan, A. Alley, K. Ghaffari and H. W. Ressom, *Metabolomics*, 2020, **16**, 104.
48 M. Wang, A. K. Jarmusch, F. Vargas, A. A. Aksenov, J. M. Gauglitz, K. Weldon, D. Petras, R. da Silva, R. Quinn, A. V. Melnik, J. J. J. van der Hooft, A. M. Caraballo-Rodríguez, L. F. Nothias, C. M. Aceves, M. Panitchpakdi, E. Brown, F. Di Ottavio, N. Sikora, E. O. Elijah, L. Labarta-Bajo, E. C. Gentry, S. Shalapour, K. E. Kyle, S. P. Puckett, J. D. Watrous, C. S. Carpenter, A. Bouslimani, M. Ernst, A. D. Swafford, E. I. Zúñiga, M. J. Balunas, J. L. Klassen, R. Loomba, R. Knight, N. Bandeira and P. C. Dorrestein, *Nat. Biotechnol.*, 2020, **38**, 23–26.
49 F. Huber, L. Ridder, S. Verhoeven, J. H. Spaaks, F. Diblen, S. Rogers and J. J. J. van der Hooft, *BioRxiv*, 2020, doi:10.1101/2020.08.11.245928.
50 L.-F. Nothias, D. Petras, R. Schmid, K. Dührkop, J. Rainer, A. Sarvepalli, I. Protsyuk, M. Ernst, H. Tsugawa, M. Fleischauer, F. Aicheler, A. A. Aksenov, O. Alka, P.-M. Allard, A. Barsch, X. Cachet, A. M. Caraballo-Rodriguez, R. R. Da Silva, T. Dang, N. Garg, J. M. Gauglitz, A. Gurevich, G. Isaac, A. K. Jarmusch, Z. Kameník, K. B. Kang, N. Kessler, I. Koester, A. Korf, A. Le Gouellec, M. Ludwig, C. Martin H, L.-I. McCall, J. McSayles, S. W. Meyer, H. Mohimani, M. Morsy, O. Moyne, S. Neumann, H. Neuweger, N. H. Nguyen, M. Nothias-Esposito, J. Paolini, V. V. Phelan, T. Pluskal, R. A. Quinn, S. Rogers, B. Shrestha, A. Tripathi, J. J. J. van der Hooft, F. Vargas, K. C. Weldon, M. Witting, H. Yang, Z. Zhang, F. Zubeil, O. Kohlbacher, S. Böcker, T. Alexandrov, N. Bandeira, M. Wang and P. C. Dorrestein, *Nat. Methods*, 2020, **17**, 905–908.
51 A. A. Aksenov, I. Laponogov, Z. Zhang, S. L. F. Doran, I. Belluomo, D. Veselkov, W. Bittremieux, L. F. Nothias, M. Nothias-Esposito, K. N. Maloney, B. B. Misra, A. V. Melnik, A. Smirnov, X. Du, K. L. Jones, K. Dorrestein, M. Panitchpakdi, M. Ernst, J. J. J. van der Hooft, M. Gonzalez, C. Carazzone, A. Amézquita, C. Callewaert, J. T. Morton, R. A. Quinn, A. Bouslimani, A. A. Orio, D. Petras, A. M. Smania, S. P. Couvillion, M. C. Burnet, C. D. Nicora, E. Zink, T. O. Metz, V. Artaev, E. Humston-Fulmer, R. Gregor, M. M. Meijler, I. Mizrahi, S. Eyal, B. Anderson, R. Dutton, R. Lugan, P. Le Boulch, Y. Guitton, S. Prevost, A. Poirier, G. Dervilly, B. Le Bizec, A. Fait, N. S. Persi, C. Song, K. Gashu, R. Coras, M. Guma, J. Manasson, J. U. Scher, D. K. Barupal, S. Alseekh, A. R. Fernie, R. Mirnezami, V. Vasiliou, R. Schmid, R. S. Borisov, L. N. Kulikova, R. Knight, M. Wang, G. B. Hanna, P. C. Dorrestein and K. Veselkov, *Nat. Biotechnol.*, 2020, doi:10.1038/s41587-020-0700-3.
52 A. K. Jarmusch, M. Wang, C. M. Aceves, R. S. Advani, S. Aguirre, A. A. Aksenov, G. Aleti, A. T. Aron, A. Bauermeister, S. Bolleddu, A. Bouslimani, A. M. C. Rodriguez, R. Chaar, R. Coras, E. O. Elijah, M. Ernst, J. M. Gauglitz, E. C. Gentry, M. Husband, S. A. Jarmusch, K. L. Jones, Z. Kamenik, A. Le Gouellec, A. Lu, L.-I. McCall, K. L. McPhail, M. J. Meehan, A. V. Melnik, R. C. Menezes, Y. A. M. Giraldo, N. H. Nguyen, L. F. Nothias, M. Nothias-Esposito, M. Panitchpakdi, D. Petras, R. A. Quinn, N. Sikora, J. J. J. van der Hooft, F. Vargas, A. Vrbanac, K. C. Weldon, R. Knight, N. Bandeira and P. C. Dorrestein, *Nat. Methods*, 2020, **17**, 901–904.
53 A. Tripathi, Y. Vázquez-Baeza, J. M. Gauglitz, M. Wang, K. Dührkop, M. Nothias-Esposito, D. D. Acharya, M. Ernst, J. J. J. van der Hooft, Q. Zhu, D. McDonald, A. Gonzalez, J. Handelsman, M. Fleischauer, M. Ludwig, S. Böcker, L.-F. Nothias, R. Knight and P. C. Dorrestein, *Nat. Chem. Biol.*, 2020, **17**, 146-151.
54 S. Lee, S. Hwang, M. Seo, K. B. Shin, K. H. Kim, G. W. Park, J. Y. Kim, J. S. Yoo and K. T. No, *Phytochemistry*, 2020, **177**, 112427.
55 J. Xue, C. Guijas, H. Paul Benton, B. Warth and G. Siuzdak, *Nature Methods*, 2020, **17**, 953–954.
56 B. Behsaz, H. Mohimani, A. Gurevich, A. Prjibelski, M. Fisher, F. Vargas, L. Smarr, P. C. Dorrestein, J. S. Mylne and P. A. Pevzner, *Cell Syst*, 2020, **10**, 99–108.e5.
57 E. Ricart, M. Pupin, M. Müller and F. Lisacek, *Anal. Chem.*, 2020, **92**, 15862–15871.
58 L. Cao, A. Gurevich, K. L. Alexander, C. B. Naman, T. Leão, E. Glukhov, T. Luzzatto-Knaan, F. Vargas, R. Quinn, A. Bouslimani, L. F. Nothias, N. K. Singh, J. G. Sanders, R. A. S. Benitez, L. R.

Thompson, M.-N. Hamid, J. T. Morton, A. Mikheenko, A. Shlemov, A. Korobeynikov, I. Friedberg, R. Knight, K. Venkateswaran, W. H. Gerwick, L. Gerwick, P. C. Dorrestein, P. A. Pevzner and H. Mohimani, *Cell Syst*, 2019, **9**, 600–608.e4.

59 G. H. Eldjárn, A. Ramsay, J. J. J. van der Hooft, K. R. Duncan, S. Soldatou, J. Rousu, R. Daly, J. Wandy and S. Rogers, *BioRxiv*, 2020, doi:10.1101/2020.06.12.148205.

60 J. M. Stokes, K. Yang, K. Swanson, W. Jin, A. Cubillos-Ruiz, N. M. Donghia, C. R. MacNair, S. French, L. A. Carfrae, Z. Bloom-Ackermann, V. M. Tran, A. Chiappino-Pepe, A. H. Badran, I. W. Andrews, E. J. Chory, G. M. Church, E. D. Brown, T. S. Jaakkola, R. Barzilay and J. J. Collins, *Cell*, 2020, **181**, 475–483.

61 B. Mikulak-Klucznik, P. Gołębiowska, A. A. Bayly, O. Popik, T. Klucznik, S. Szymkuć, E. P. Gajewska, P. Dittwald, O. Staszewska-Krajewska, W. Beker, T. Badowski, K. A. Scheidt, K. Molga, J. Młynarski, M. Mrksich and B. A. Grzybowski, *Nature*, 2020, **588**, 83-88.