1 Analytical

Q-TOF	Instrument	maXis 4G equipped with HD collision cell (Bruker, Leiderdorp, the Netherlands)
	Ionization	ESI positive and negative
		Nebulizer gas flow: 1 Bar Dry gas flow: 8.0 l/min Dry heater: 200°C Capillary voltage: 3250V / 3500V (positive/negative) End Plate Offset: -500V
	Funnel RF	250 Vpp
	Multipole RF	250 Vpp
	Collision cell	Cell RF 400 Vpp
		Transfer time: 80 µs Pre pulse storage time: 10 µs
	Scan range	80 – 1200 m/z
	Spectra rate	4 Hz
	Mass calibration	Automatic internal calibration with 2 mM sodium formate in 1:1 water/methanol solution
	MS/MS	AutoMS/MS mode (data-dependent)
	AutoMS/MS settings	Cycle time: 1.0 sec Threshold: 7500 cts Smart exclusion: 5x Active exclusion: Exclude after 3 spectra; Release after 0.75 min; Reconsider Precursor if Current intensity / Previous intensity = 5x CID interpolated list: 100: type=Base; width=8; CE=15 500: type=Base; width=8; CE=35 1000: type=Base; width=10; CE=50 2000: type=Base; width=15; CE=75 Precursor Acquisition Control: Low: 10000 cts; scan at 2 Hz; High: 500000 cts; scan at 8 Hz
UHPLC	Instrument	Nexera (Shimadzu, Den Bosch, the Netherlands)
	UV detection	210 nm
	LC method I	Mobile phase: A: ultrapure water; B: methanol Flow: 0.25 ml/min Column: Kinetex F5 Column temperature: 40°C Injection volume: 5 µl Gradient: T=0: 20% B; T=20: 95% B; T=25: 95% B; T=26: 100% B; T=31: 100% B; T=31.5: 20%B Gradient equilibration: 10 minutes First three minutes diverted to waste Samples were diluted 20 times to 2:8 water/methanol
	LC method II	Mobile phase: A: ultrapure water + 0.1% formic acid; B: acetonitrile + 0.1% formic acid Flow: 0.25 ml/min

		Column: C18 Column temperature: 40°C Injection volume: 10 µl Gradient: T=0: 70%B; T=10: 90%B; T=12: 100%B; T=14: 100%B; T=15: 70%B Gradient equilibration: 10 minutes First three minutes diverted to waste Samples were diluted 100 times to 3:7 water/acetonitrile

Non-target analysis data processing workflow

Non-target analysis (NTA) data was processed with patron,[1] which is an open-source platform that harmonizes various commonly used software tools used for NTA. The patRoon workflow consisted of

1. Converting raw mass spectrometry data to the open mzML format by DataAnalysis[2]
2. Extracting feature data and grouping them across sample analysis by OpenMS algorithms[3]
3. Post-filtering the feature data by
    a. Removing any features with low intensity (<10000)
    b. Removing features that are not at least 10 times higher in intensity compared to solvent blank injections.
    c. Remove features not present in all triplicate injections and with intensity variations amongst triplicates of >75% RSD
    d. Remove all features present in the control samples and inactive fractions (unless their intensity in active fractions was at least 10 times higher)
    e. Remove any overlapping features from less active fractions (4-6), unless their intensity in the most active fraction (3) was at least 1.25 times higher.
    f. Only keep features in proximity (+/- 1 minute) of the chromatographic peak detected with UV detection in the most active fraction.
4. Performing automatic annotation

a. Mass spectral data (MS and MS/MS) was automatically extracted (using the mzR algorithm[4]), where a window of 0.002 Dalton was used for spectral averaging.
b. The MS/MS data was post-filtered to only retain MS/MS peaks with >= 1% intensity and being amongst the top 10 most intense.
c. Formulae were automatically calculated for all features with both the GenForm and SIRIUS algorithms (using CHNOPS as possible elements).[5-9] The results were pooled with the consensus functionality of patRoon.
d. Compound structure annotation was performed with SIRIUS and MetFrag,[10] both using the PubChem library,[11] and results were pooled afterwards. Ranking occurred by correspondence of experimentally observed and in-silico calculated MS/MS fragments and correspondence with formula rank (step c). Furthermore, for MetFrag, candidates were also ranked on (1) similarity of experimental and library MS/MS spectra from the MassBank North America[11] and (2) the number of literature references in PubMed and the number of patents, as provided by the PubChem database. Only the top 25 ranked candidates with an explained chemical formula (step c) were kept.

5. Performing 'componentization' to automatically detect features that are actually isotopes or MS adducts (using RAMClustR as algorithm).[12] The results were used to manually remove any of such features.

References

1. R. Helmus, T. T. Laak, P. d. Voogt, A. v. Wezel and E. Schymanski, Patroon: Open Source Software Platform for Environmental Mass Spectrometry Based Non-target Screening, *Journal of Cheminformatics*, Submitted.

2. L. Martens, M. Chambers, M. Sturm, D. Kessner, F. Levander, J. Shofstahl, W. H. Tang, A. Rompp, S. Neumann, A. D. Pizarro, L. Montecchi-Palazzi, N. Tasman, M. Coleman, F. Reisinger, P. Souda, H. Hermjakob, P. A. Binz and E. W. Deutsch, mzML--a community standard for mass spectrometry data, *Mol Cell Proteomics*, 2011, DOI: 10.1074/mcp.R110.000133.

3. H. L. Röst, T. Sachsenberg, S. Aiche, C. Bielow, H. Weisser, F. Aicheler, S. Andreotti, H. C. Ehrlich, P. Gutenbrunner, E. Kenar, X. Liang, S. Nahnsen, L. Nilse, J. Pfeuffer, G. Rosenberger, M. Rurik, U. Schmitt, J. Veit, M. Walzer, D. Wojnar, W. E. Wolski, O. Schilling, J. S. Choudhary, L. Malmström, R. Aebersold, K. Reinert and O. Kohlbacher, OpenMS: A flexible open-source software platform for mass spectrometry data analysis, *Nature Methods*, 2016, **13**, 741-748.

4. S. N. Bernd Fischer, Laurent Gatto, Qiang Kou, Johannes Rainer, No TimzR: parser for netCDF, mzXML, mzData and mzML and mzIdentML files (mass spectrometry data), DOI: 10.18129/B9.bioc.mzR.

5. M. Meringer, S. Reinker, J. Zhang and A. Muller, MS/MS data improves automated determination of molecular formulas by mass spectrometry, *Match*, 2011, **65**, 259-290.

6. S. Böcker, M. C. Letzel, Z. Lipták, A. Pervukhin and O. Troyanskaya, SIRIUS: decomposing isotope patterns for metabolite identification †, *Bioinformatics*, 2009, **25**, 218-224.

7. S. Böcker and K. Dührkop, Fragmentation trees reloaded, *Journal of Cheminformatics*, 2016, **8**, 5.

8. K. Dührkop, H. Shen, M. Meusel, J. Rousu and S. Böcker, Searching molecular structure databases with tandem mass spectra using CSI:FingerID, *Proceedings of the National Academy of Sciences*, 2015, **112**, 12580 - 12585.

9. K. Dührkop, M. Fleischauer, M. Ludwig, A. A. Aksenov, A. V. Melnik, M. Meusel, P. C. Dorrestein, J. Rousu and S. Böcker, SIRIUS 4: a rapid tool for turning tandem mass spectra into metabolite structure information, *Nature Methods*, 2019, **16**, 299-302.

10. C. Ruttkies, E. L. Schymanski, S. Wolf, J. Hollender and S. Neumann, MetFrag relaunched: incorporating strategies beyond in silico fragmentation., *Journal of cheminformatics*, 2016, **8**, 3.

11. S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, S. He, Q. Li, B. A. Shoemaker, P. A. Thiessen, B. Yu, L. Zaslavsky, J. Zhang and E. E. Bolton, PubChem 2019 update: Improved access to chemical data, *Nucleic Acids Research*, 2019, **47**, 1102-1109.

12. C. D. Broeckling, F. A. Afsar, S. Neumann, A. Ben-Hur and J. E. Prenni, RAMClust: A novel feature clustering method enables spectral-matching-based annotation for metabolomics data, *Analytical Chemistry*, 2014, **86**, 6812-6817.