analytical chemistry

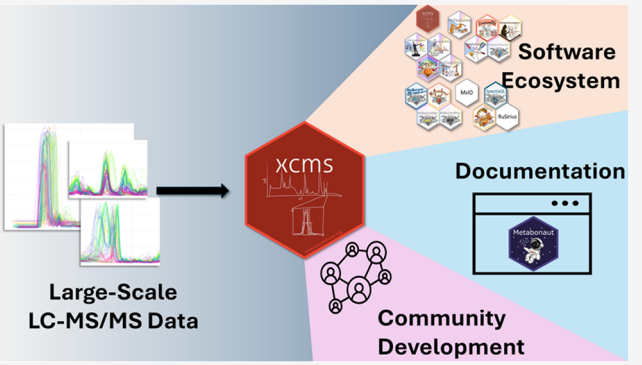
pubs.acs.org/ac
Article

xcms in Peak Form: Now Anchoring a Complete Metabolomics Data Preprocessing and Analysis Software Ecosystem

Philippine Louail, Carl Brunius, Mar Garcia-Aloy, William Kumler, Norman Storz, Jan Stanstrup, Hendrik Treutler, Pablo Vangeenderhuysen, Michael Witting, Steffen Neumann,* and Johannes Rainer

Cite This: *Anal. Chem.* 2025, 97, 27639−27645
Read Online

ABSTRACT: High-quality data preprocessing is essential for untargeted metabolomics experiments, where increasing data set scale and complexity demand adaptable, robust, and reproducible software solutions. Modern preprocessing tools must evolve to integrate seamlessly with downstream analysis platforms, ensuring efficient and streamlined workflows. Since its introduction in 2005, the *xcms* R package has become one of the most widely used tools for LC-MS data preprocessing. Developed through an open-source, community-driven approach, *xcms* maintains long-term stability while continuously expanding its capabilities and accessibility. We present recent advancements that position *xcms* as a central component of a modular and interoperable software ecosystem for metabolomics data analysis. Key improvements include enhanced scalability, enabling the processing of large-scale experiments with thousands of samples on standard computing hardware. These developments empower users to build comprehensive, customizable, and reproducible workflows tailored to diverse experimental designs and analytical needs. An expanding collection of tutorials, documentation, and teaching materials further supports both new and experienced users in leveraging broader R and Bioconductor ecosystems. These resources facilitate the integration of statistical modeling, visualization tools, and domain-specific packages, extending the reach and impact of *xcms* workflows. Together, these enhancements solidify *xcms* as a cornerstone of modern metabolomics research.



Preprocessing is the crucial first step in analyzing untargeted liquid chromatography−mass spectrometry (LC-MS) or gas chromatography−mass spectrometry (GC-MS) data. It involves quantifying ion signals in a sample, correcting potential retention time drifts within and across analytical runs, and grouping signals from the same ion species across all samples in an experiment. Preprocessing software must accommodate data from various column types, such as reversed-phase (RP) chromatography and hydrophilic interaction liquid chromatography (HILIC), which produce diverse metabolite separations and chromatographic peak structures. As high-resolution mass spectrometers like Quadrupole-Time-of-Flight (QTOF) and Orbitrap become increasingly common, and with growing demand for high-throughput metabolomics, data processing tools must keep pace with increasing data complexity and volume. Several key challenges have emerged: the need to handle large-scale data sets, ensure reproducibility and transparency, and provide flexibility for a wide range of experimental designs.

Originally introduced in 2005[1] with chromatographic peak detection and retention time alignment algorithms, *xcms* has since been cited over 3800 times and was downloaded over 150,000 times by distinct IP addresses since the introduction of download statistics in Bioconductor in 2009. Over the two decades since its initial release, *xcms* has undergone significant transformations to address the aforementioned challenges while maintaining its core principles of flexibility, external collaboration, and open-source innovation. A key factor in its adaptability and longevity was its early integration into the Bioconductor project,[2,3] making it one of the first metabolomics software packages in this ecosystem, now comprising over 2,200 packages (as of Bioconductor version 3.20). Beyond Bioconductor, *xcms* has also been incorporated into other platforms, such as *Galaxy Workflow4Metabolomics*,[4] further extending its reach and usability.

Over the past two decades, innovations in LC-MS data acquisition have significantly reshaped the field of metabolomics, and data preprocessing has evolved in parallel. A wide array of software tools is now available to support various stages of analysis.[5] In the context of preprocessing, *xcms* is

Received: July 16, 2025
Revised: November 25, 2025
Accepted: November 26, 2025
Published: December 8, 2025

https://doi.org/10.1021/acs.analchem.5c04338
Anal. Chem. 2025, 97, 27639−27645

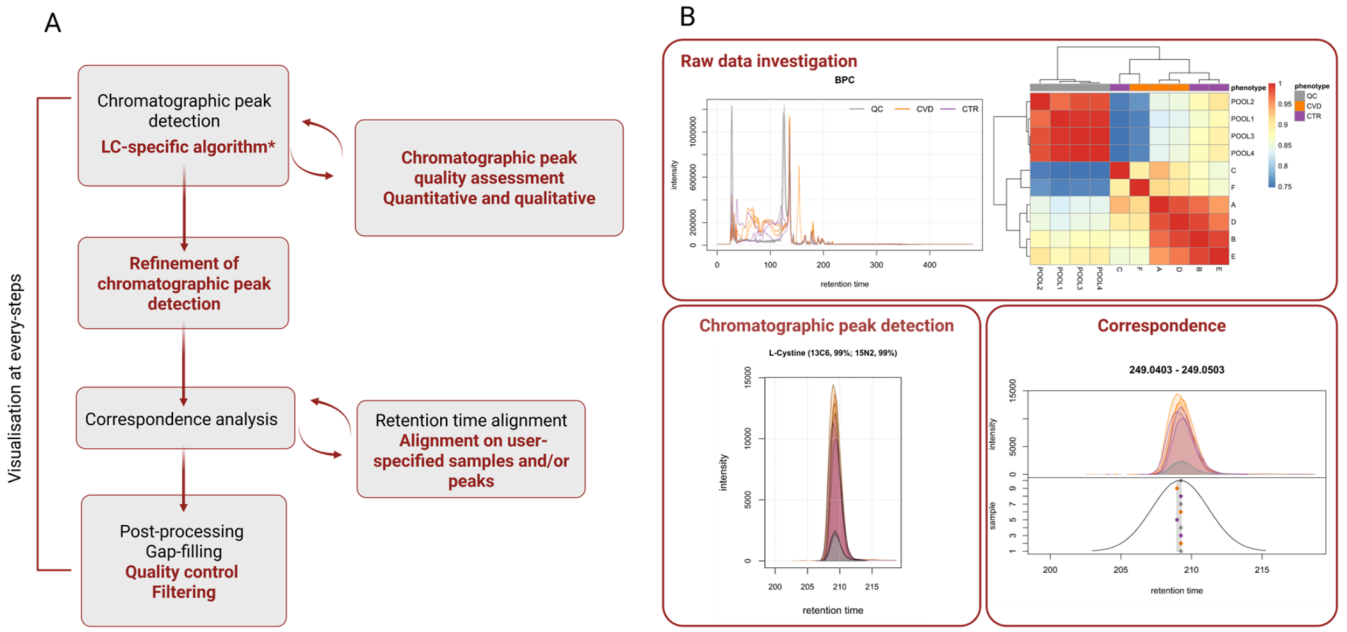


Figure 1. Graphical representation of key technical improvements in *xcms* over the past 20 years. (A) Preprocessing workflow in *xcms*: original steps from the 2006 paper are shown in black, while newly implemented improvements are highlighted in red, reflecting the continued development of *xcms* that preserves its core methodology while expanding its functionality and performance. * The new LC-specific algorithms are centWave[12] and MassifQuant[13]. (B) Examples of possible visualizations of raw data (top) and data at different preprocessing stages (lower row), enabling evaluation at each step. The source code and data to produce these images are available in the tutorials on the *Metabonaut* Web site,[14] which can be found: [Exploring and Analyzing LC-MS Data • Metabonaut]. Created in BioRender. Fuchsberger, C. (2025) https://BioRender.com/cuokujz.

accompanied by other well-established tools such as *mzmine*[6] and *MS-DIAL*,[7] both of which have seen substantial development and feature expansion in recent years. Comparative descriptions of their respective strengths and limitations have been done before.[8−10] *xcms* offers strengths in reproducibility, scalability, and expandability (with custom functions), although it requires some R/coding skills. Unlike vendor or Graphic User Interface (GUI)-based software, it therefore addresses a different target audience seeking flexibility and integrative workflows. Moreover, these tools operate outside the R and Bioconductor ecosystem, highlighting the growing importance of interoperability, an area that *xcms* developers have actively addressed. Recent efforts include the development of common export formats and supporting infrastructure designed to improve integration with external tools and programming languages beyond R.

It is important to clarify that the developments discussed in this article pertain specifically to the *xcms* R package and its evolution within the R/Bioconductor ecosystem. These updates do not apply to *XCMS Online*[11] or other web-based tools that use a different or older *xcms* code base and remain separate entities with distinct functionalities and development trajectories. This article summarizes the major technical and infrastructural advancements of the *xcms* R package over the past two decades. We present developments that have solidified *xcms* as one of the main tools for LC-MS data preprocessing. We discuss key advancements, including methodological improvements, expanded support for diverse mass spectrometry (MS) data formats, and integration with an evolving ecosystem of R packages for MS data analysis. Improvements to the underlying data structures and increased flexibility in the processing steps allow memory-saving data analysis even for very large data sets. Reproducible case studies are discussed throughout the article, offering concrete validation of the improvements described, providing practical guidance on workflow design, parameter selection, and downstream analysis.

As part of the Bioconductor ecosystem, *xcms* continues to foster a collaborative and adaptable environment for metabolomics research. Over the past 20 years, these innovations have transformed *xcms* into a comprehensive and scalable toolbox within the R environment, facilitating the development of versatile end-to-end workflows for LC-MS and LC-MS/MS data analysis.

DATA ANALYSIS CAPABILITIES

xcms has continuously evolved to meet the needs of scientists and data analysts, with several key implementations introduced to enhance its functionality. Central to these updates is the introduction of parameter classes, which form the foundation of a modular and flexible programming interface. Each major preprocessing step, such as peak detection, retention time alignment, and chromatographic peak grouping, is now handled by a dedicated function, with multiple algorithms available for each step. These algorithms are configured via dedicated parameter classes, allowing users to easily define and customize settings based on the specific requirements of their experiment. This structured approach simplifies workflow construction, enhances flexibility, and promotes reproducibility by embedding parameter settings directly into the processing history. As a result, *xcms* now provides an intuitive yet powerful interface that supports transparency and traceability throughout the data analysis process.

Key steps in *xcms* preprocessing have been significantly refined (Figure 1A), as detailed in the following section. For example, peak detection is now coupled with peak filtering and

https://doi.org/10.1021/acs.analchem.5c04338
Anal. Chem. 2025, 97, 27639−27645

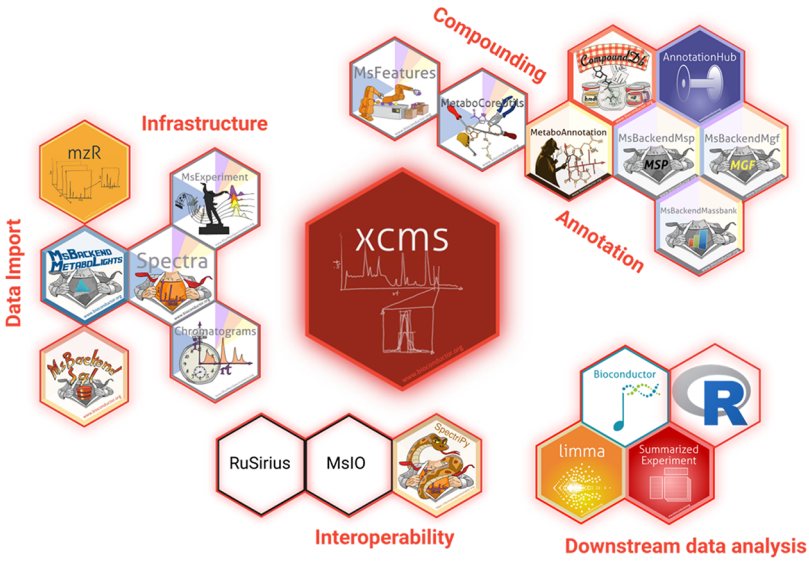


Figure 2. *xcms* is part of an ecosystem of packages, expanding the analysis possibilities. See Supporting Table S1 for a detailed listing and description of the packages in the figure and Stanstrup et al.[35] for a complete overview of metabolomics R packages. Created in BioRender. Fuchsberger, C. (2025) https://BioRender.com/jxpxdq7.

refinement, addressing artifacts that can arise in the initial steps by postprocessing detected chromatographic peaks and removing or merging them if necessary. Additionally, peak quality metrics, such as those presented in Kumler et al.,[15] can be calculated either during or after peak detection, allowing users to filter the data and retain only high-confidence signals. These improved filtering approaches help minimize the inclusion of likely false-positive peaks, a common issue in earlier studies.[16−19] The alignment process has also seen substantial improvements with the introduction of new algorithms and the option to base it on a subset of the data set or an external reference data set. The gap-filling step, designed to recover missing values for low-abundance features, has been improved by employing a more accurate estimation of the ions' expected m/z−retention time region. Lastly, several methods for quality control of features have been developed, adhering to widely accepted standards in metabolomics.[20] An end-to-end workflow using an example data set (Supporting File S1) illustrates these new preprocessing steps in detail.

The core data structures of *xcms* were completely rewritten to gain full support for LC-MS/MS data, as demonstrated in the Supporting File S2 and the *GNPS* molecular networking case study.[21] These new data structures rely on the established MS data infrastructures of *MSnbase*[22] and packages from the *RforMassSpectrometry* initiative,[23−25] in particular, the *Spectra* and *Chromatograms* R packages. Apart from a tighter integration with other R packages that will be discussed in the following section, this redesign significantly enhances *xcms*' ability to handle, subset, and filter data while also ensuring compatibility with emerging data formats. To evaluate the performance improvements of the recent developments in *xcms*, we repeated the analysis of the large LC-MS/MS data set with 1039 samples from the aforementioned *GNPS* molecular networking study using different versions of *xcms* on the same computational setup (using 4 cores of a standard notebook computer with an Intel Core i7−1370P CPU and 64 GB total RAM). The total runtime of the analysis, which includes preprocessing, selection of features' MS2 spectra, and export of the data for molecular networking with *GNPS*, took 860 min with the versions used in the original article[21] (R 3.6.3, Bioconductor 3.10, *xcms* 3.8.3). In contrast, using the current versions (R 4.5.1, Bioconductor 3.22, *xcms* 4.7.3), the processing time was reduced to only 280 min, despite an additional peak refinement step being performed. In addition to this ∼3x reduction in total runtime, memory usage was also much lower. See Supporting Files S2 and S3 for the respective analysis reports, including runtime and memory usages of the individual analysis steps.

To further reduce memory demands, a new *xcms* result object with a very low memory footprint was implemented, allowing memory-saving data analysis even for very large data sets. In a second large-scale metabolomics analysis comprising ∼4000 samples from a public data set, preprocessing could be performed on a standard notebook computer (same hardware as described above), underscoring the scalability and memory efficiency of the redesigned system (Supporting File S4). Together, these results demonstrate that the enhanced *xcms* data structures and workflows enable high-throughput LC−MS/MS data processing on standard hardware, eliminating the need for specialized computing resources.

Beyond the improved data processing algorithms, significant emphasis has been placed on enhancing data visualization within *xcms*. The aforementioned redesigned architecture optimizes chromatographic data organization and accessibility, enabling users to generate and visualize extracted ion chromatograms (EICs) from large data sets throughout the entire analysis process. Combined with the robust plotting possibilities of R, this integration ensures an intuitive and powerful visualization experience. Users can test and fine-tune settings for each preprocessing step, leading to more informed and justifiable parameter choices. Tailored plotting functions have been developed for different stages of the analysis, whether working with the full, raw MS data or examining results at various preprocessing steps. Specifically, users can generate base peak chromatograms (BPCs) from raw data, EICs with highlighted chromatographic peaks during peak detection, and density plots for feature grouping, allowing for visual insights throughout the workflow (Figure 1B).

https://doi.org/10.1021/acs.analchem.5c04338
Anal. Chem. 2025, 97, 27639−27645

INTEGRATION AND INTEROPERABILITY

The development of *xcms* and its surrounding ecosystem (Figure 2 and Supporting Table S1) is primarily driven by its integration within the Bioconductor[2,3] community, which ensures adherence to high-standard software guidelines, including rigorous unit testing and documentation. This integration promotes open-source development, encourages community contributions, and supports direct user feedback through GitHub's issue-tracking system as well as through the Bioconductor support site. Additionally, it enables *xcms* to leverage standardized data structures, such as *SummarizedExperiment*,[26] the central Bioconductor data type for representing quantitative data from biological assays. This greatly simplifies interoperability with other Bioconductor packages, strengthening the ecosystem by improving data handling within *xcms* and enhancing downstream analyses, including data normalization, visualization, statistical data analysis, or annotation. The aforementioned end-to-end workflow demonstrates this integration of *xcms* within the broader R/Bioconductor mass spectrometry ecosystem (Supporting File S1).

Within the *RforMassSpectrometry*[23,24] initiative, various packages have been developed to serve as the basis for, or to complement, *xcms'* preprocessing possibilities. In this context, the previously mentioned *Spectra* and *Chromatograms* packages stand out as they efficiently manage the complexity of LC-MS data. The *MsFeatures* package adds yet another layer of versatility in postanalysis by grouping LC-MS *features* that represent the signal of ions and isotopes of the same compound. Meanwhile, the *MetaboCoreUtils* package provides researchers with a set of basic tools to extend and refine their analysis. The *MsExperiment* package provides a structured and user-oriented infrastructure that organizes experimental metadata and data files in a way that is directly compatible with *xcms*. Additionally, other packages in this initiative extend the utility of *xcms* beyond preprocessing, supporting postanalysis tasks. For instance, MS/MS spectra extracted from *xcms* objects can be used with the *MetaboAnnotation*[27] package for matching against reference fragment spectra. *xcms* integration within a rich ecosystem enables incorporating functionalities from diverse infrastructures into *xcms*-centered R workflows, reducing dependency on any single tool and thus enhancing both flexibility and scalability.

xcms has expanded its compatibility with external software workflows, extending its utility beyond R-based analyses. The objective is to foster a more inclusive and collaborative approach, enabling a broader user and developer base to collectively address the complex challenges inherent in LC-MS analysis. To achieve this level of interoperability, the *MsIO* package was introduced, supporting the import and export of *xcms* results in multiple file formats compatible with leading platforms. Examples include the support for the mzTab-M[28] format, which will serve as the foundation for future integration with *mzmine*[6] and *MS-DIAL*[7] as well as JSON-based and HDF5-based formats for enhanced embedding within the Galaxy platform.[29] Additionally, *xcms* now supports direct data exchange with community repositories like *MetaboLights*,[30] facilitating streamlined data sharing and reuse. Further expanding its reach, *xcms* supports integration with tools and libraries outside the R ecosystem, such as Python libraries *matchms*[31] and *spectrum_utils*[32] (via *SpectriPy*[33]) and applications like *Sirius*[34] (through ongoing development of *RuSirius*). These integrations establish *xcms* as a cornerstone in a broader ecosystem of software tools, enabling collaboration and analysis across diverse platforms and technologies.

USER EXPERIENCE AND ENABLING REPRODUCIBILITY

While the scripting use of *xcms* might be challenging for new users, it enables one to create reproducible and customizable analysis workflows, which can be difficult to achieve in GUI-based software solutions. In particular, using *xcms* with R Markdown, or the recently introduced R interface of the Quarto system,[36] allows one to document and describe the analysis and its results transparently. The improved integration of *xcms* with R and Bioconductor greatly simplifies the implementation of such workflows by avoiding tedious conversion and copying of the data. As the flexibility, scalability, and functionalities of *xcms* have grown, so has its complexity. Consequently, significant emphasis has been placed on enhancing documentation by writing reference manuals for all available functions and providing a range of tutorials. These resources (see Supporting Table S1 for links) provide examples to tune and derive settings for *xcms* preprocessing[37] as well as small use case analyses to exemplify how the functionalities of *xcms* and other packages from the *RforMassSpectrometry* initiative can be combined for in-depth data inspection[38] and small compound annotation.[27]

The most recent and significant addition to the documentation is the creation of an end-to-end data analysis workflow that takes the user from importing raw data from *MetaboLights*,[30] over LC-MS data preprocessing, quality assessment, data normalization, statistical data analysis, and finally the annotation of the identified significant features. This workflow is shared, as part of the *Metabonaut* resource, publicly on GitHub, where it is accompanied by targeted tutorials covering specific parts of analysis such as raw data investigation, integration with external tools such as Python libraries, *notame*,[39] all within an R session. This workflow, along with accompanying vignettes, represents the culmination of nearly two decades of *xcms* evolution, transforming it from a standalone preprocessing tool into a component of a much more extensive, easily scalable, and customizable toolkit for LC-MS data analysis. It also signals the future direction of *xcms* and the related software ecosystem: greater integration, enhanced interoperability, and a stronger emphasis on community collaboration and education in the field.

CONCLUSION

xcms has evolved far beyond its initial role as a preprocessing tool, becoming a core component of a flexible, scalable, and interoperable LC-MS/MS data analysis ecosystem. Its deep integration with Bioconductor has enabled advanced users to construct workflows that link preprocessing, quality control, statistical modeling, and annotation. Through GitHub discussions, community workshops, and initiatives such as *Metabonaut*, *xcms* has fostered an inclusive and collaborative user base, supporting novice users and experts. A key factor in *xcms'* long-term success is its adaptability to emerging analytical challenges. The implementation of dynamic parameter classes has introduced a more transparent and customizable workflow structure. In parallel, algorithmic refinements and enhanced visualization functions have increased both the accuracy and the interpretability of results.

https://doi.org/10.1021/acs.analchem.5c04338
Anal. Chem. 2025, 97, 27639−27645

Nevertheless, some limitations remain. For example, while support for ion mobility data exists at the backend level, current preprocessing algorithms do not yet exploit the ion mobility dimension. Also, although powerful, *xcms* can present a steep learning curve for new users due to its programmatic interface. However, ongoing efforts are underway to address these concerns. Graphical user interfaces such as *Metaboseek*,[40] *patRoon*,[41] and *Galaxy* workflows provide GUI and coding-free experience for interactive data exploration and preprocessing, making the ecosystem more accessible.

Interoperability remains a major strength of *xcms*. It aims to connect with external platforms like *GNPS*,[21,42] *Sirius*, and *MetaboLights*,[30] making it a valuable part of the global metabolomics infrastructure. Extensive tutorials, teaching resources, and community-contributed workflows have significantly improved the user experience and supported adoption across disciplines.

Through rigorous software development, open-source collaboration, and a strong community foundation, *xcms* has positioned itself as a modern, adaptable tool that continues to grow with the evolving needs of metabolomics research. As data sets increase in scale and complexity, *xcms* remains well equipped to support the next generation of scientific discoveries.

■ ASSOCIATED CONTENT

Data Availability Statement

The *xcms* software is publicly available through Bioconductor at 10.18129/B9.bioc.xcms. The latest development version can be found on GitHub at https://github.com/sneumann/xcms. The software is licensed under the GPL (≥2) license. The full license information can be accessed here. *Metabonaut* v1.2.0 is publicly available at https://rformassspectrometry.github.io/Metabonaut/, where the complete end-to-end workflow, alignment based on external references, raw data investigation, a *SpectriPy* tutorial, and a demonstration of large-scale data set preprocessing can be found. The images used in Figure 1 are also part of the end-to-end workflow. The files for the reanalysis of the large LC-MS/MS data set from Nothias et al.[21] are available on GitHub: https://github.com/jorainer/xcms-gnps-large-scale and Zenodo 10.5281/zenodo.17293665.

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.analchem.5c04338.

File with a complete *xcms*-based end-to-end metabolomics data analysis workflow (Supporting File S2) (PDF)

Analysis report file (PDF format) of the preprocessing of a large LC-MS/MS data set for molecular networking with GNPS. Analysis performed with R version 3.6.3, Bioconductor 3.10, and *xcms* 3.8.3 (Supporting File S2) (PDF)

Analysis report file (PDF format) of the preprocessing of a large LC-MS/MS data set for molecular networking with GNPS. Analysis performed with R version 4.5.1, Bioconductor 3.22, and *xcms* 4.7.3 (Supporting File S3) (PDF)

Analysis report file (PDF format) for memory-saving *xcms*-based preprocessing of a large LC-MS data set (Supporting File S4) (PDF)

Table (PDF format) listing packages present in the *xcms* toolkit and their respective information (description, status of development, repository link), therefore expanding Figure 2 (Supporting Table S1) (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Steffen Neumann − *Leibniz Institute of Plant Biochemistry, MetaCom, 06120 Halle, Germany; German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, 04103 Leipzig, Germany;* orcid.org/0000-0002-7899-7192; Email: sneumann@ipb-halle.de

Authors

Philippine Louail − *Institute for Biomedicine, Eurac Research, 39100 Bolzano, Italy; Chair for Bioinformatics, Friedrich-Schiller-University Jena, 07743 Jena, Germany;* orcid.org/0009-0007-5429-6846

Carl Brunius − *Department of Life Sciences, Food and Nutrition Science, Chalmers University of Technology, SE-412 96 Göteborg, Sweden*

Mar Garcia-Aloy − *Metabolomics Unit, Research and Innovation Centre, Fondazione Edmund Mach, 38098 San Michele all'Adige (TN), Italy;* orcid.org/0000-0002-1330-6610

William Kumler − *School of Oceanography, University of Washington, Seattle, Washington 98195, United States;* orcid.org/0000-0002-5022-8009

Norman Storz − *Leibniz Institute of Plant Biochemistry, MetaCom, 06120 Halle, Germany*

Jan Stanstrup − *Department of Nutrition, Exercise and Sports, University of Copenhagen, 1958 Frederiksberg C, Denmark;* orcid.org/0000-0003-0541-7369

Hendrik Treutler − *Leibniz Institute of Plant Biochemistry, MetaCom, 06120 Halle, Germany*

Pablo Vangeenderhuysen − *Laboratory of Integrative Metabolomics (LIMET), Ghent University, 9820 Merelbeke, Belgium;* orcid.org/0000-0002-5492-6904

Michael Witting − *Metabolomics and Proteomics Core, Helmholtz Zentrum München, 85764 Neuherberg, Germany; Chair of Analytical Food Chemistry, TUM School of Life Sciences, Technical University of Munich, 85354 Freising-Weihenstephan, Germany*

Johannes Rainer − *Institute for Biomedicine, Eurac Research, 39100 Bolzano, Italy;* orcid.org/0000-0002-6977-7147

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.analchem.5c04338

Author Contributions

P.L., J.R., and S.N. conceptualized and wrote the original manuscript draft. Main software development and maintenance were done by J.R., P.L., and S.N. C.B. contributed to the methodology for external reference alignment. W.K. and P.V. contributed to peak summarization and quality assessment functionality. M.W. contributed to DIA data analysis support. J.S., N.S., M.G.A., and H.T. contributed documentation, bug fixes, functionality for manual, or isotope-guided, peak detection and integration. All authors were involved in manuscript reviewing and gave approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

https://doi.org/10.1021/acs.analchem.5c04338
Anal. Chem. 2025, 97, 27639−27645

■ ACKNOWLEDGMENTS

The authors thank the original developers and Gary Siuzdak for sharing their work as open-source software with the community. We thank also Paul Benton and Christopher Conley for their early contributions. Further acknowledgments go to Ronan Daly, Laurent Gatto, Rick Helmus, Tony Larson, Nathaniel Mahieu, Duncan Murdoch, Sergio Oller, and Egon Willighagen for smaller contributions and bug fixes over the years. The authors thank the Department of Innovation, Research University and Museums of the Autonomous Province of Bozen/Bolzano for covering the Open Access publication costs. The authors acknowledge the support of the European Union under the HORIZON-MSCA-2021 project (Grant No. 101073062: HUMAN—Harmonising and Unifying Blood Metabolic Analysis Networks). Additionally, this work was supported by grants from the Simons Foundation (LS award ID: 385428; SCOPE award ID 329108). This work was also funded by the Federal Ministry of Research, Technology and Space in the frame of de.NBI/ELIXIR-DE (W-de.NBI-11).

■ REFERENCES

(1) Smith, C. A.; Want, E. J.; O'Maille, G.; Abagyan, R.; Siuzdak, G. *Anal. Chem.* **2006**, *78* (3), 779−787.
(2) Gentleman, R. C.; Carey, V. J.; Bates, D. M.; Bolstad, B.; Dettling, M.; Dudoit, S.; Ellis, B.; Gautier, L.; Ge, Y.; Gentry, J.; Hornik, K.; Hothorn, T.; Huber, W.; Iacus, S.; Irizarry, R.; Leisch, F.; Li, C.; Maechler, M.; Rossini, A. J.; Sawitzki, G.; Smith, C.; Smyth, G.; Tierney, L.; Yang, J. Y. H.; Zhang, J. *Genome Biol.* **2004**, *5* (10), No. R80.
(3) Huber, W.; Carey, V. J.; Gentleman, R.; Anders, S.; Carlson, M.; Carvalho, B. S.; Bravo, H. C.; Davis, S.; Gatto, L.; Girke, T.; Gottardo, R.; Hahne, F.; Hansen, K. D.; Irizarry, R. A.; Lawrence, M.; Love, M. I.; MacDonald, J.; Obenchain, V.; Oleś, A. K.; Pagès, H.; Reyes, A.; Shannon, P.; Smyth, G. K.; Tenenbaum, D.; Waldron, L.; Morgan, M. *Nat. Methods* **2015**, *12* (2), 115−121.
(4) Giacomoni, F.; Le Corguillé, G.; Monsoor, M.; Landi, M.; Pericard, P.; Pétéra, M.; Duperier, C.; Tremblay-Franco, M.; Martin, J.-F.; Jacob, D.; Goulitquer, S.; Thévenot, E. A.; Caron, C. *Bioinformatics* **2015**, *31* (9), 1493−1495.
(5) Misra, B. B. *Electrophoresis* **2018**, *39* (7), 909−923.
(6) Schmid, R.; Heuckeroth, S.; Korf, A.; Smirnov, A.; Myers, O.; Dyrlund, T. S.; Bushuiev, R.; Murray, K. J.; Hoffmann, N.; Lu, M.; Sarvepalli, A.; Zhang, Z.; Fleischauer, M.; Dührkop, K.; Wesner, M.; Hoogstra, S. J.; Rudt, E.; Mokshyna, O.; Brungs, C.; Ponomarov, K.; Mutabdžija, L.; Damiani, T.; Pudney, C. J.; Earll, M.; Helmer, P. O.; Fallon, T. R.; Schulze, T.; Rivas-Ubach, A.; Bilbao, A.; Richter, H.; Nothias, L.-F.; Wang, M.; Orešič, M.; Weng, J.-K.; Böcker, S.; Jeibmann, A.; Hayen, H.; Karst, U.; Dorrestein, P. C.; Petras, D.; Du, X.; Pluskal, T. *Nat. Biotechnol.* **2023**, *41* (4), 447−449.
(7) Tsugawa, H.; Cajka, T.; Kind, T.; Ma, Y.; Higgins, B.; Ikeda, K.; Kanazawa, M.; VanderGheynst, J.; Fiehn, O.; Arita, M. *Nat. Methods* **2015**, *12* (6), 523−526.
(8) Coble, J. B.; Fraga, C. G. *J. Chromatogr. A* **2014**, *1358*, 155−164.
(9) Guo, J.; Huan, T. *Anal. Chem.* **2023**, *95* (14), 5894−5902.
(10) Hohrenk, L. L.; Itzel, F.; Baetz, N.; Tuerk, J.; Vosough, M.; Schmidt, T. C. *Anal. Chem.* **2020**, *92* (2), 1898−1907.
(11) Tautenhahn, R.; Patti, G. J.; Rinehart, D.; Siuzdak, G. *Anal. Chem.* **2012**, *84* (11), 5035−5039.
(12) Tautenhahn, R.; Böttcher, C.; Neumann, S. *BMC Bioinf.* **2008**, *9* (1), No. 504.
(13) Conley, C. J.; Smith, R.; Torgrip, R. J. O.; Taylor, R. M.; Tautenhahn, R.; Prince, J. T. *Bioinformatics* **2014**, *30* (18), 2636−2643.
(14) Louail, P.; De Graeve, M.; Tagliaferri, A.; Verri, H. V.; de Sá e Silva, D. M.; Rainer, J. Rformassspectrometry/Metabonaut: V1.2.0. 2025 DOI: 10.5281/ZENODO.15062929.
(15) Kumler, W.; Hazelton, B. J.; Ingalls, A. E. *BMC Bioinf.* **2023**, *24* (1), No. 404.
(16) Gürdeniz, G.; Kristensen, M.; Skov, T.; Dragsted, L. O. *Metabolites* **2012**, *2* (1), 77−99.
(17) Aigensberger, M.; Bueschl, C.; Castillo-Lopez, E.; Ricci, S.; Rivera-Chacon, R.; Zebeli, Q.; Berthiller, F.; Schwartz-Zimmermann, H. E. *Anal. Chim. Acta* **2025**, *1336*, No. 343491.
(18) Hemmer, S.; Manier, S. K.; Fischmann, S.; Westphal, F.; Wagmann, L.; Meyer, M. R. *Metabolites* **2020**, *10* (9), No. 378.
(19) Myers, O. D.; Sumner, S. J.; Li, S.; Barnes, S.; Du, X. *Anal. Chem.* **2017**, *89* (17), 8689−8695.
(20) Broadhurst, D.; Goodacre, R.; Reinke, S. N.; Kuligowski, J.; Wilson, I. D.; Lewis, M. R.; Dunn, W. B. *Metabolomics* **2018**, *14* (6), No. 72.
(21) Nothias, L.-F.; Petras, D.; Schmid, R.; Dührkop, K.; Rainer, J.; Sarvepalli, A.; Protsyuk, I.; Ernst, M.; Tsugawa, H.; Fleischauer, M.; Aicheler, F.; Aksenov, A. A.; Alka, O.; Allard, P.-M.; Barsch, A.; Cachet, X.; Caraballo-Rodriguez, A. M.; Da Silva, R. R.; Dang, T.; Garg, N.; Gauglitz, J. M.; Gurevich, A.; Isaac, G.; Jarmusch, A. K.; Kameník, Z.; Kang, K. B.; Kessler, N.; Koester, I.; Korf, A.; Le Gouellec, A.; Ludwig, M.; Martin, H. C.; McCall, L.-I.; McSayles, J.; Meyer, S. W.; Mohimani, H.; Morsy, M.; Moyne, O.; Neumann, S.; Neuweger, H.; Nguyen, N. H.; Nothias-Esposito, M.; Paolini, J.; Phelan, V. V.; Pluskal, T.; Quinn, R. A.; Rogers, S.; Shrestha, B.; Tripathi, A.; van der Hooft, J. J. J.; Vargas, F.; Weldon, K. C.; Witting, M.; Yang, H.; Zhang, Z.; Zubeil, F.; Kohlbacher, O.; Böcker, S.; Alexandrov, T.; Bandeira, N.; Wang, M.; Dorrestein, P. C. *Nat. Methods* **2020**, *17* (9), 905−908.
(22) Gatto, L.; Gibb, S.; Rainer, J. *J. Proteome Res.* **2021**, *20* (1), 1063−1069.
(23) Gatto, L.; Vanderaa, C.; Gibb, S.; Rainer, J. An Open Software Development-Based Ecosystem of R Packages for Proteomics Data Analysis. 2024, DOI: 10.5281/zenodo.13382962.
(24) Rainer, J.; Louail, P.; Vicini, A.; Gine, R.; Badia, J. M.; Stravs, M.; Garcia-Aloy, M.; Huber, C.; Salzer, L.; Stanstrup, J.; Shahaf, N.; Panse, C.; Naake, T.; Kumler, W.; Vangeenderhuysen, P.; Brunius, C.; Hecht, H.; Neumann, S.; Witting, M.; Gibb, S.; Gatto, L. An Open Software Development-Based Ecosystem of R Packages for Metabolomics Data Analysis. 2024. DOI: 10.5281/ZENODO.7936786.
(25) Rainer, J.; De Graeve, M.; Louai, P.; Gibb, S.; Gatto, L.An Open Infrastructure for Mass Spectrometry Data in R *Open Science Framework July 22* 2025 DOI: 10.31219/osf.io/cwt2v_v2.
(26) Martin Morgan, V. O. SummarizedExperiment. 2017, DOI: 10.18129/B9.BIOC.SUMMARIZEDEXPERIMENT.
(27) Rainer, J.; Vicini, A.; Salzer, L.; Stanstrup, J.; Badia, J. M.; Neumann, S.; Stravs, M. A.; Verri Hernandes, V.; Gatto, L.; Gibb, S.; Witting, M. *Metabolites* **2022**, *12* (2), No. 173.
(28) Hoffmann, N.; Rein, J.; Sachsenberg, T.; Hartler, J.; Haug, K.; Mayer, G.; Alka, O.; Dayalan, S.; Pearce, J. T. M.; Rocca-Serra, P.; Qi, D.; Eisenacher, M.; Perez-Riverol, Y.; Vizcaíno, J. A.; Salek, R. M.; Neumann, S.; Jones, A. R. *Anal. Chem.* **2019**, *91* (5), 3302−3310.
(29) Abueg, L. A. L.; Afgan, E.; Allart, O.; Awan, A. H.; Bacon, W. A.; Baker, D.; Bassetti, M.; Batut, B.; Bernt, M.; Blankenberg, D.; Bombarely, A.; Bretaudeau, A.; Bromhead, C. J.; Burke, M. L.; Capon, P. K.; Čech, M.; Chavero-Díez, M.; Chilton, J. M.; Collins, T. J.; Coppens, F.; Coraor, N.; Cuccuru, G.; Cumbo, F.; Davis, J.; De Geest, P. F.; De Koning, W.; Demko, M.; DeSanto, A.; Begines, J. M. D.; Doyle, M. A.; Droesbeke, B.; Erxleben-Eggenhofer, A.; Föll, M. C.; Formenti, G.; Fouilloux, A.; Gangazhe, R.; Genthon, T.; Goecks, J.; Beltran, A. N. G.; Goonasekera, N. A.; Goué, N.; Griffin, T. J.; Grüning, B. A.; Guerler, A.; Gundersen, S.; Gustafsson, O. J. R.; Hall, C.; Harrop, T. W.; Hecht, H.; Heidari, A.; Heisner, T.; Heyl, F.; Hiltemann, S.; Hotz, H.-R.; Hyde, C. J.; Jagtap, P. D.; Jakiela, J.; Johnson, J. E.; Joshi, J.; Jossé, M.; Jum'ah, K.; Kalaš, M.; Kamieniecka, K.; Kayikcioglu, T.; Konkol, M.; Kostrykin, L.; Kucher, N.; Kumar, A.; Kuntz, M.; Lariviere, D.; Lazarus, R.; Bras, Y. L.; Corguillé, G. L.; Lee,

J.; Leo, S.; Liborio, L.; Libouban, R.; Tabernero, D. L.; Lopez-Delisle, L.; Los, L. S.; Mahmoud, A.; Makunin, I.; Marin, P.; Mehta, S.; Mok, W.; Moreno, P. A.; Morier-Genoud, F.; Mosher, S.; Müller, T.; Nasr, E.; Nekrutenko, A.; Nelson, T. M.; Oba, A. J.; Ostrovsky, A.; Polunina, P. V.; Poterlowicz, K.; Price, E. J.; Price, G. R.; Rasche, H.; Raubenolt, B.; Royaux, C.; Sargent, L.; Savage, M. T.; Savchenko, V.; Savchenko, D.; Schatz, M. C.; Seguineau, P.; Serrano-Solano, B.; Soranzo, N.; Srikakulam, S. K.; Suderman, K.; Syme, A. E.; Tangaro, M. A.; Tedds, J. A.; Tekman, M.; Cheng Mike Thang, W.; Thanki, A. S.; Uhl, M.; Van Den Beek, M.; Varshney, D.; Vessio, J.; Videm, P.; Von Kuster, G.; Watson, G. R.; Whitaker-Allen, N.; Winter, U.; Wolstencroft, M.; Zambelli, F.; Zierep, P.; Zoabi, R. *Nucleic Acids Res.* **2024**, *52* (W1), W83−W94.

(30) Yurekten, O.; Payne, T.; Tejera, N.; Amaladoss, F. X.; Martin, C.; Williams, M.; O'Donovan, C. *Nucleic Acids Res.* **2024**, *52* (D1), D640−D646.

(31) Huber, F.; Verhoeven, S.; Meijer, C.; Spreeuw, H.; Castilla, E. M. V.; Geng, C.; Hooft, J. J.; j van der Rogers, S.; Belloum, A.; Diblen, F.; Spaaks, J. H. *J. Open Source Software* **2020**, *5* (52), No. 2411.

(32) Bittremieux, W.; Levitsky, L.; Pilz, M.; Sachsenberg, T.; Huber, F.; Wang, M.; Dorrestein, P. C. *J. Proteome Res.* **2023**, *22* (2), 625−631.

(33) Graeve, M. D.; Bittremieux, W.; Naake, T.; Huber, C.; Anagho-Mattanovich, M.; Hoffmann, N.; Marchal, P.; Chrone, V.; Louail, P.; Hecht, H.; Witting, M.; Rainer, J. *J. Open Source Software* **2025**, *10* (109), No. 8070.

(34) Dührkop, K.; Fleischauer, M.; Ludwig, M.; Aksenov, A. A.; Melnik, A. V.; Meusel, M.; Dorrestein, P. C.; Rousu, J.; Böcker, S. *Nat. Methods* **2019**, *16* (4), 299−302.

(35) Stanstrup, J.; Broeckling, C. D.; Helmus, R.; Hoffmann, N.; Mathé, E.; Naake, T.; Nicolotti, L.; Peters, K.; Rainer, J.; Salek, R. M.; Schulze, T.; Schymanski, E. L.; Stravs, M. A.; Thévenot, E. A.; Treutler, H.; Weber, R. J. M.; Willighagen, E.; Witting, M.; Neumann, S. *Metabolites* **2019**, *9* (10), No. 200.

(36) Allaire, J. J.; Dervieux, C. Quarto: R Interface to "Quarto" Markdown Publishing System. 2025.

(37) Rainer, J.; Louail, P. jorainer/xcmsTutorials: xcmsTutorials Version 1.1.0. 2024, DOI: 10.5281/zenodo.11185521.

(38) Rainer, J.; Louail, P.; Witting, M.; Gatto, L.; Gibb, S. jorainer/SpectraTutorials: SpectraTutorial Version 1.1.0, 2024 DOI: 10.5281/zenodo.11210190.

(39) Klåvus, A.; Kokla, M.; Noerman, S.; Koistinen, V. M.; Tuomainen, M.; Zarei, I.; Meuronen, T.; Häkkinen, M. R.; Rummukainen, S.; Farizah Babu, A.; Sallinen, T.; Kärkkäinen, O.; Paananen, J.; Broadhurst, D.; Brunius, C.; Hanhineva, K. *Metabolites* **2020**, *10* (4), No. 135.

(40) Helf, M. J.; Fox, B. W.; Artyukhin, A. B.; Zhang, Y. K.; Schroeder, F. C. *Nat. Commun.* **2022**, *13* (1), No. 782.

(41) Helmus, R.; van de Velde, B.; Brunner, A. M.; ter Laak, T. L.; van Wezel, A. P.; Schymanski, E. L. *J. Open Source Software* **2022**, *7* (71), No. 4029.

(42) Schmid, R.; Petras, D.; Nothias, L.-F.; Wang, M.; Aron, A. T.; Jagels, A.; Tsugawa, H.; Rainer, J.; Garcia-Aloy, M.; Dührkop, K.; Korf, A.; Pluskal, T.; Kameník, Z.; Jarmusch, A. K.; Caraballo-Rodríguez, A. M.; Weldon, K. C.; Nothias-Esposito, M.; Aksenov, A. A.; Bauermeister, A.; Albarracin Orio, A.; Grundmann, C. O.; Vargas, F.; Koester, I.; Gauglitz, J. M.; Gentry, E. C.; Hövelmann, Y.; Kalinina, S. A.; Pendergraft, M. A.; Panitchpakdi, M.; Tehan, R.; Le Gouellec, A.; Aleti, G.; Mannochio Russo, H.; Arndt, B.; Hübner, F.; Hayen, H.; Zhi, H.; Raffatellu, M.; Prather, K. A.; Aluwihare, L. I.; Böcker, S.; McPhail, K. L.; Humpf, H.-U.; Karst, U.; Dorrestein, P. C. *Nat. Commun.* **2021**, *12* (1), No. 3832.

https://doi.org/10.1021/acs.analchem.5c04338
Anal. Chem. 2025, 97, 27639−27645