https://doi.org/10.1038/s41564-024-01656-3

MicrobioRaman: an open-access web repository for microbiological Raman spectroscopy data

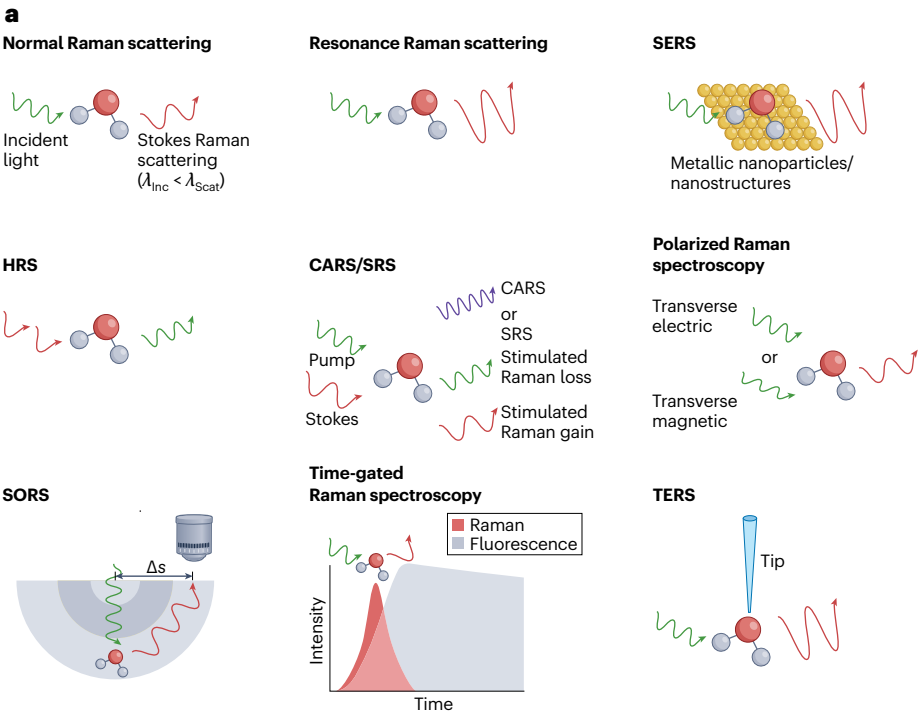
Here we present the establishment of an open-access web-based repository for microbiological Raman spectroscopy data. The data collection, called 'MicrobioRaman' (https://www.ebi.ac.uk/biostudies/MicrobioRaman/studies), was inspired by the great success and usefulness of research databases such as GenBank and UniProt. This centralized repository, residing within the BioStudies database[1] – which is maintained by a public institution, the European Bioinformatics Institute – minimizes the risk of data loss or eventual abandonment, offering a long-term common reference for analysis with advantages in accessibility and transparency over commercial data analysis tools. We feel that MicrobioRaman will provide a foundation for this growing field by serving as an open-access repository for sharing microbiological Raman data and through the codification of a set of reporting standards.

Raman spectroscopy is a type of vibrational spectroscopy that relies on inelastic scattering, in which, after interaction with molecules in a sample, the wavelength of the scattered light differs from the wavelength of the incident light, which is typically provided by a laser. This shift in wavelength differs according to the type of molecules and their vibrational modes, allowing for the analysis of complex sample chemistry in a non-destructive manner (Fig. 1a). When Raman spectroscopy is applied for the measurement of microbiological samples, specific Raman peaks at different wavenumbers indicate the presence of macromolecules such as carbohydrates, proteins, lipids, nucleic acids and pigments[2]. Recent advances in technology and data analysis now enable the investigation of molecular composition at the resolution of a single microorganism with high measurement sensitivity (Fig. 1a). By measuring the presence of peaks corresponding to specific macromolecules or differences in spectral shape, peak position and relative intensity of peaks, and often in conjunction with complementary techniques such as stable isotope probing (SIP)[3], fluorescence in situ hybridization (FISH)[4] or omics[5], Raman spectroscopy enables investigation of cell identity and phenotypes. This analytical approach is increasingly being employed to address important questions in both fundamental and applied microbiology (Fig. 1b). Notable applications include the measurement of microbial diversity in terms of cell identity, metabolic phenotype and functional role within complex microbial communities. Raman spectroscopy is also enabling researchers to untangle the complexity of microbial communities, by allowing for the tracking of molecular interactions between microorganisms or between a microorganism and its host, and the interactions between microorganisms and their environment. In comparison with other technologies offering similar capabilities (enabling analysis of molecular composition and structure of samples, for instance, Fourier-transform infrared spectroscopy, cryogenic electron microscopy, nanoscale secondary ion mass spectroscopy, nuclear magnetic resonance spectroscopy), the versatility in sample size and analysis conditions (in liquid phase or dry form) and the ability to measure live microorganisms render Raman spectroscopy applicable to diverse sample types, ranging from large nematodes (and beyond) to minuscule viruses measuring a few tens of nanometres, collected from various environments spanning oceans, soils and mammalian guts, and potentially even efforts to detect signals of life on other planets such as Mars (see refs. 3,6–10 for comprehensive reviews about Raman technologies and applications in microbiology).

Despite the potential of Raman spectroscopy in microbiology, the reporting of analytical methods and data for microbiological systems has evolved in a haphazard manner, and progress in the field is hindered by the lack of both a set of standards for data reporting and a common database to deposit microbiological Raman data. Raman data from microorganisms is relatively complex to analyse because proper interpretation is dependent on biological context, experimental conditions and data processing.

Individual Raman spectra from microbiological samples, consisting of discretized wavenumbers (typically measured in cm^{-1}) and corresponding Raman scattering signals, typically encompass many (often overlapping) peaks that represent chemical bonds of diverse types of macromolecules. Identification of the source of each peak often depends on the biological context; for example, a peak at 1,570 cm^{-1} typically corresponds to C–C stretching of nucleic acids when analysing

Fig. 1 | Overview of Raman technologies and their applications in microbiology. **a**, The working principles underlying normal Raman spectroscopy and its advanced variant systems. For normal Raman spectroscopy (where λ_{Inc} and λ_{Scat} denote the wavelength of incident light and scattering signals, respectively), a laser beam interacts with molecules within a sample, resulting in Raman scattering after the interaction. Advanced variant systems, which rely on modification of the system configuration, can be categorized into three groups depending on their specific advantages: techniques that provide a higher sensitivity for measurement (resonance Raman scattering, SERS, TERS); techniques that enable rapid measurement by virtue of the selection of specific wavenumbers (CARS, SRS); and techniques that provide other functions, such as the ability to measure peaks not detectable using normal Raman spectroscopy (HRS), among others including SORS (where Δs represents a spatial offset for detecting Raman signals from a location where a laser first hits a sample), polarized Raman spectroscopy and time-gated Raman spectroscopy. **b**, Applications of Raman spectroscopy in fundamental and applied microbiology. Raman spectroscopy is a versatile technique that enables the measurement of a broad size range of samples across diverse geographical regions and ecosystems – from large nematodes to minuscule viruses found in oceans, soils, mammalian guts and industrial plant systems – and potentially even efforts to detect signals of life on other planets such as Mars. **c**, Pipeline for acquisition and analysis of microbiological Raman data and the parameters that influence the resulting data. Experimental configurations, in addition to the samples themselves, determine the resulting Raman data. When measured Raman spectra display a different level or shape of spectral background, they require computational data processing for quantitative or qualitative analyses and comparisons between samples. A section describing these factors is a part of the reporting standard in the MicrobioRaman repository.


a
Normal Raman scattering
Incident light
Stokes Raman scattering
($\lambda_{Inc} < \lambda_{Scat}$)
Resonance Raman scattering
SERS
Metallic nanoparticles/ nanostructures
HRS
CARS/SRS
CARS or SRS
Pump
Stokes
Stimulated Raman loss
Stimulated Raman gain
Polarized Raman spectroscopy
Transverse electric
or
Transverse magnetic
SORS
Δs
Time-gated Raman spectroscopy
Raman
Fluorescence
Intensity
Time
TERS
Tip

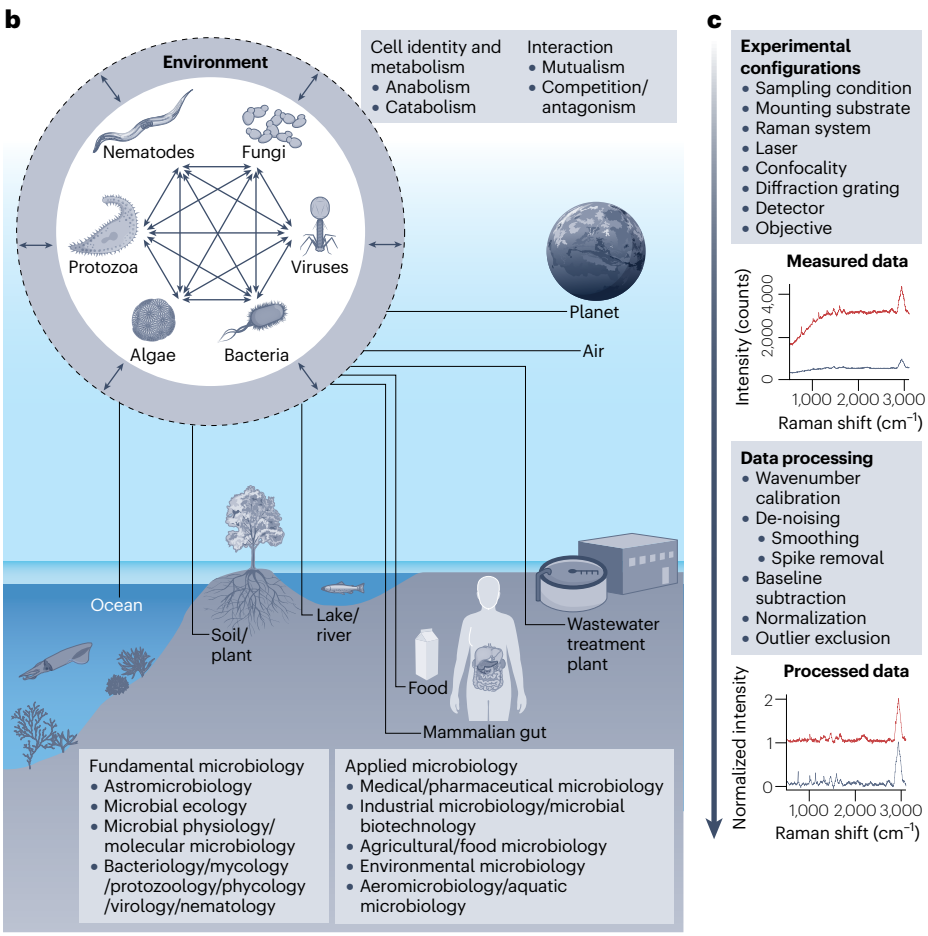

b
Environment
Nematodes
Fungi
Protozoa
Viruses
Algae
Bacteria
Cell identity and metabolism
• Anabolism
• Catabolism
Interaction
• Mutualism
• Competition/ antagonism
Planet
Air
Ocean
Soil/ plant
Lake/ river
Food
Mammalian gut
Wastewater treatment plant
Fundamental microbiology
• Astromicrobiology
• Microbial ecology
• Microbial physiology/ molecular microbiology
• Bacteriology/mycology /protozoology/phycology /virology/nematology
Applied microbiology
• Medical/pharmaceutical microbiology
• Industrial microbiology/microbial biotechnology
• Agricultural/food microbiology
• Environmental microbiology
• Aeromicrobiology/aquatic microbiology
c
Experimental configurations
• Sampling condition
• Mounting substrate
• Raman system
• Laser
• Confocality
• Diffraction grating
• Detector
• Objective
Measured data
Intensity (counts)
0
2,000 4,000
1,000 2,000 3,000
Raman shift (cm^{-1})
Data processing
• Wavenumber calibration
• De-noising
• Smoothing
• Spike removal
• Baseline subtraction
• Normalization
• Outlier exclusion
Processed data
Normalized intensity
0
1
2
1,000 2,000 3,000
Raman shift (cm^{-1})

Table 1 | Reporting standards for microbiological Raman data

Section	Parameter	Description
General	Title (M)	Project title
	Release date (M)	Desired release date, for example, to ensure compliance with a publication embargo
	Description (M)	Brief description of the project; details if isotopes (for example, SIP) or fluorescent probes (for example, FISH probes) were used
	Contacts (M)	Contact details for data authors
	Raw data files (M)	Unprocessed raw Raman data composed of wavenumbers and corresponding Raman intensities
	Peak identification (M)	Identification of peaks and methods used for identification (for example, literature, public or commercial databases); peak shifts if isotopes (for example, SIP) were used
	Software (IA)	If a commercial Raman system was used for measurement, the name and version of software
	Publications (IA)	Information about associated publications (authors, title, journal name, year)
Sample	Name of cell or compound (M)	Sample names
	Source (M)	Source of a sample, such as a strain collection, a chemical supplier, or the environment or tissue from which a sample was obtained
	Composition (M)	Entities contained in the sample, including not just the cells of interest, but also the medium, as well as any extraneous materials such as tissue, debris, biofilm matrix or soil
	Preparation (M)	For example, preservation after sampling, culturing condition (for example, sample volume, medium, light condition/diel cycle, pH, temperature, antibiotics, oxic/anoxic), sample age, fixed or unfixed, whether the cells were dry or wet for Raman measurement, whether isotopes (for example, SIP) or fluorescent probes (for example, FISH probes) were used
	Mounting substrate (M)	For example, glass coverslip, aluminium slide, CaF_2 slide, quartz slide
	Image files (IA)	Image files from Raman imaging, image files from light or electron microscopy (for example, bright-field or scanning electron microscope images of a sample, fluorescence images for FISH)
Set-up	Raman system (M)	Manufacturer and model of the scope
	Measurement type (M)	For example, normal Raman scattering, resonance Raman scattering, CARS, SRS, HRS, SERS, SORS, polarized Raman spectroscopy, TERS, time-gated Raman spectroscopy
	Lasers and connected components/parameters (M)	Wavelength and power of lasers; continuous wave or pulsed (if pulsed, pulse duration and repetition rate); if nonlinear Raman spectroscopy (for example, CARS/SRS), excitation intensity; if polarized Raman spectroscopy, polarization state; laser spot shape (for example, circular, elliptical, torus, square, rectangular); laser illumination spot size (for example, diameter for a circular shape, lengths of major and minor axes for an elliptical shape, inner and outer diameters for torus, length and breadth for a square or rectangular shape); neutral density filter (100% if not used); grating (0 if not used); acquisition time and accumulation number (for averaging) for measurement; spectral window and resolution for measurement
	Lasers and connected components/parameters (IA)	Manufacturer and model of lasers
Treated spectrum	Processed data files (IA)	Processed Raman data and data analysis (for example, principal component analysis, hierarchical cluster analysis, linear discriminant analysis)
	Data treatments (IA)	List of computational algorithms and their parameters and sources used for data processing and analysis
	Software (IA)	If manufacturer's software was used for data processing and analysis, the name and version of software
Instrument metadata	Annotations (R)	For example, the type of spike filter, detector specifications, details of a microscope objective or focusing lens, confocality, spectral binning

The 'general' section describes general information about the submission; the 'sample' section provides the biological context and treatment; the 'set-up' section provides experimental conditions; the 'treated spectrum' section describes data processing; and the 'instrument metadata' section provides additional instrument information that could help users to reproduce the measurements. The level of recommended reporting for the parameters is indicated: mandatory (M), if applicable (IA) and recommended (R). See also the help page of BioStudies (https://www.ebi.ac.uk/biostudies/submissions/help) for general instructions for submission of novel data.

microorganisms more generally, but predominantly to calcium dipicolinic acids when measuring endospore-forming bacteria[6]. Moreover, SIP or FISH, often coupled to Raman measurements to track metabolic exchange or identify microorganisms of interest, induce a red shift of Raman peaks (that is, peak positions move to lower wavenumbers) or potentially a change in overall spectral shape (due to interference between some fluorescent dyes used for FISH and certain Raman lasers), respectively, adding further complexity to the interpretation of microbiological Raman data.

Experimental conditions further complicate the analysis of microbiological Raman data. Compared with Raman measurements in research fields in which samples are in the solid state (often the case in material science or electrical engineering) or target cells are relatively large (a few tens of micrometres, as in biomedical engineering), samples in microbiology often contain a diversity of molecules at relatively low concentrations (diverse cell components, with the majority in liquid phase) and target cells are rather small (for example, bacteria or archaea ranging down to a few hundreds of nanometres).

Microbiological measurements are thus substantially influenced by sampling conditions and the biotic and abiotic environment of cells at the time of the analysis (Fig. 1c).

For both quantitative and qualitative analyses of large datasets, Raman data are often processed with computational algorithms[6] (Fig. 1c). Because interpretation can often depend on the presence of peak shoulders or small changes in peak locations of the order of a few tens of wavenumbers, as is the case in isotope labelling, any computational treatment can potentially affect the interpretation of microbiological Raman data.

Considering these three aspects, microbiological Raman data share similarities with other types of microbiological data, namely, those derived from omics approaches. While fields that rely on the use of these data types have greatly benefitted from the availability of organized central and public repositories for published data with reporting standards, the lack of an actively maintained, open-access data repository for microbiological Raman data has been an obstacle to the wider adoption of Raman spectroscopy in microbiology. Currently, published data are scattered across various sources (for example, deposited on a journal publication webpage or an author's personal or institutional repository) in the absence of rational and clear reporting standards, making it challenging for researchers to access and use the data. There are several databases commercially available, for example, KnowItAll (https://sciencesolutions.wiley.com/solutions/technique/raman/knowitall-raman-collection/) and one by S.T. Japan Inc. (https://www.stjapan.de/spectra-databases/raman-spectra-databases/). These databases aim to cover the broad range of organic and inorganic materials, and are not specific to microbiological Raman data. As such, considering the peculiarities of microbiological Raman data described above, a database tailored to microbiological Raman data would be highly beneficial to promote sharing and reuse of such data across diverse users within the community. Moreover, in light of how useful research databases such as GenBank (https://www.ncbi.nlm.nih.gov/genbank/) and UniProt (https://www.uniprot.org/) have proved to be, we are witnessing the unique power of 'collective intelligence', where each user plays an important role in data accumulation over time and the amassed data are used for further analyses from different perspectives by other users. In accordance with this, a bottom-up, open-access data repository would substantially reinforce the power and usefulness of Raman spectroscopy in microbiology.

The MicrobioRaman platform is now open for current and future Raman users — covering data from normal Raman spectroscopy to its advanced variant systems[6] such as, but not limited to, resonance Raman spectroscopy, stimulated Raman spectroscopy (SRS), coherent anti-Stokes Raman spectroscopy (CARS), surface-enhanced Raman spectroscopy (SERS), tip-enhanced Raman spectroscopy (TERS), hyper Raman spectroscopy (HRS), spatially offset Raman spectroscopy (SORS), polarized Raman spectroscopy and time-gated Raman spectroscopy. Step-by-step, recipe-style instructions for deposition of novel datasets are provided on the help page (https://www.ebi.ac.uk/biostudies/submissions/help).

MicrobioRaman aims to provide a comprehensive repository of Raman data acquired from fundamental and applied microbiology research (Fig. 1b). The platform was collaboratively developed among the authors of this Correspondence, and it establishes a set of standards for data reporting to ensure reproducible Raman measurements across different users.

The standards for data reporting consist of five sections (Table 1): (1) general information about the authors and project underlying the data submitted; (2) biological context, including both general information and specific sample details; (3) experimental conditions, encompassing the set-up used for Raman measurements; (4) data processing, particularly focusing on the treatment of the spectrum and classification of a dataset into subgroups; and (5) instrument metadata, such as the type of spike filter, detector specifications, details of the microscope objective or focusing lens, confocality and spectral binning. Additionally, the platform allows data submitters to specify a public release date for newly deposited data, for example, to ensure compliance with publication embargos.

As MicrobioRaman grows, it will become a valuable resource with diverse applications. It will serve as a chemical catalogue, housing data on the distribution of compounds across taxa and ecosystems. Furthermore, it will function as a source of standardized experimental designs, inspiring novel approaches. The current wave of applications of machine learning is already beginning to impact Raman-based approaches in microbiology. The ability to collect Raman data and make them broadly accessible is timely in this regard, as the effectiveness of machine learning approaches often relies on collective intelligence — in particular, data in the repository may be reused as part of training datasets in supervised approaches[6].

In conclusion, we believe that, by establishing reporting standards and facilitating data sharing among Raman users, MicrobioRaman will play an important role in promoting the adoption of Raman spectroscopy in microbiology. This initiative represents a cornerstone for reproducible Raman measurements and will seed further developments in this field. We envision the development of new functions for MicrobioRaman as it grows with active participation from Raman users in the community and the accumulation of novel microbiological Raman data. For example, creation of an open-access library of biological molecules for peak identification within spectra and its integration with MicrobioRaman could be considered. With this Correspondence, we pledge to deposit our future data into this newly constructed infrastructure and we encourage other Raman users to contribute, further reinforcing the power and potential of reproducible Raman measurements in microbiology.

Kang Soo Lee[1] ✉, Zachary Landry[1,2], Awais Athar[3], Uria Alcolombri[4], Pratchaya Pramoj Na Ayutthaya[1], David Berry[5,6], Philippe de Bettignies[7], Ji-Xin Cheng[8], Gabor Csucs[9], Li Cui[10], Volker Deckert[11,12], Thomas Dieing[13], Jennifer Dionne[14], Ondrej Doskocil[15], Glen D'Souza[16,17], Cristina García-Timermans[18], Notburga Gierlinger[19], Keisuke Goda[20,21,22], Roland Hatzenpichler[23], Richard J. Henshaw[1], Wei E. Huang[24], Ievgeniia Iermak[13], Natalia P. Ivleva[25], Janina Kneipp[26], Patrick Kubryk[27], Kirsten Küsel[28,29], Tae Kwon Lee[30], Sung Sik Lee[9,31], Bo Ma[32], Clara Martínez-Pérez[1], Pavel Matousek[33], Rainer U. Meckenstock[34], Wei Min[35,36], Peter Mojzeš[37], Oliver Müller[1], Naresh Kumar[38], Per Halkjær Nielsen[39], Ioan Notingher[40], Márton Palatinszky[5], Fátima C. Pereira[5,41], Giuseppe Pezzotti[42,43,44], Zdenek Pilat[15], Filip Plesinger[15], Jürgen Popp[11,12], Alexander J. Probst[45], Alessandra Riva[5,46], Amr. A. E. Saleh[47], Ota Samek[15], Haley M. Sapers[48], Olga T. Schubert[16,17], Astrid K. M. Stubbusch[16,17,49], Loza F. Tadesse[50,51,52], Gordon T. Taylor[53], Michael Wagner[5,39], Jing Wang[1,54], Huabing Yin[55], Yang Yue[1,54], Renato Zenobi[38], Jacopo Zini[56,57], Ugis Sarkans[3] ✉ & Roman Stocker[1] ✉

Correspondence

[1]Institute for Environmental Engineering, Department of Civil, Environmental and Geomatic Engineering, ETH Zurich, Zurich, Switzerland. [2]Department of Marine and Environmental Biology, University of Southern California, Los Angeles, CA, USA. [3]European Molecular Biology Laboratory, European Bioinformatics Institute, EMBL-EBI, Cambridge, UK. [4]Department of Plant and Environmental Sciences, Institute of Life Science, The Hebrew University of Jerusalem, Jerusalem, Israel. [5]Division of Microbial Ecology, Centre for Microbiology and Environmental Systems Science, University of Vienna, Vienna, Austria. [6]Joint Microbiome Facility of the Medical University of Vienna and the University of Vienna, Vienna, Austria. [7]HORIBA France SAS, Palaiseau, France. [8]Department of Electrical and Computer Engineering and Department of Biomedical Engineering, Boston University, Boston, MA, USA. [9]Scientific Center for Optical and Electron Microscopy, ETH Zurich, Zurich, Switzerland. [10]Key Lab of Urban Environment and Health, Institute of Urban Environment, Chinese Academy of Sciences, Xiamen, China. [11]Institute of Physical Chemistry (IPC) and Abbe Center of Photonics (ACP), Friedrich Schiller University Jena, member of the Leibniz Centre for Photonics in Infection Research (LPI), Jena, Germany. [12]Leibniz Institute of Photonic Technology e.V. Jena, member of Leibniz Health Technology, member of the Leibniz Centre for Photonics in Infection Research (LPI), Jena, Germany. [13]WITec GmbH, Ulm, Germany. [14]Department of Materials Science and Engineering, and Department of Radiology, Stanford University, Stanford, CA, USA. [15]Institute of Scientific Instruments of the Czech Academy of Sciences, v.v.i, Brno, Czech Republic. [16]Department of Environmental Microbiology, Eawag, Dübendorf, Switzerland. [17]Institute of Biogeochemistry and Pollutant Dynamics, Department of Environmental Systems Science, ETH Zurich, Zurich, Switzerland. [18]CMET, Center for Microbial Technology and Ecology, Department of Biotechnology, Ghent University, Gent, Belgium. [19]Institute of Biophysics, Department of Bionanosciences, University of Natural Resources and Life Sciences (BOKU), Vienna, Austria. [20]Department of Chemistry, The University of Tokyo, Tokyo, Japan. [21]Institute of Technological Sciences, Wuhan University, Wuhan, China. [22]Department of Bioengineering, University of California, Los Angeles, CA, USA. [23]Department of Chemistry and Biochemistry, Department of Microbiology and Cell Biology, Center for Biofilm Engineering, and Thermal Biology Institute, Montana State University, Bozeman, MT, USA. [24]Department of Engineering Science, University of Oxford, Oxford, UK. [25]Chair of Analytical Chemistry and Water Chemistry, Institute of Water Chemistry, TUM School of Natural Sciences (Dep. Chemistry), Technical University of Munich, Garching, Germany. [26]Department of Chemistry, Humboldt- Universität zu Berlin, Berlin, Germany. [27]Renishaw GmbH, Pliezhausen, Germany. [28]Institute of Biodiversity, Aquatic Geomicrobiology, Friedrich Schiller University Jena, Jena, Germany. [29]Cluster of Excellence Balance of the Microverse, Friedrich Schiller University Jena, Jena, Germany. [30]Department of Environmental and Energy Engineering, Yonsei University, Wonju, Republic of Korea. [31]Institute of Biochemistry, Department of Biology, ETH Zurich, Zurich, Switzerland. [32]Single-Cell Center, CAS Key Laboratory of Biofuels and Shandong Key Laboratory of Energy Genetics, Qingdao Institute of Bioengineering and Bioprocess Technology, Chinese Academy of Sciences, Qingdao, China. [33]Central Laser Facility, Research Complex at Harwell, STFC Rutherford Appleton Laboratory, UKRI, Harwell, UK. [34]Environmental Microbiology and Biotechnology (EMB), University of Duisburg-Essen, Essen, Germany. [35]Department of Chemistry, Columbia University, New York, NY, USA. [36]Kavli Institute for Brain Science, Columbia University, New York, NY, USA. [37]Institute of Physics, Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic. [38]Department of Chemistry and Applied Biosciences, ETH Zurich, Zurich, Switzerland. [39]Center for Microbial Communities, Department of Chemistry and Bioscience, Aalborg University, Aalborg, Denmark. [40]School of Physics and Astronomy, University of Nottingham, Nottingham, UK. [41]School of Biological Sciences, University of Southampton, Southampton, UK. [42]Ceramic Physics Laboratory, Kyoto Institute of Technology, Kyoto, Japan. [43]Department of Molecular Genetics, Institute of Biomedical Science, Kansai Medical University, Osaka, Japan. [44]Department of Immunology, Graduate School of Medical Science, Kyoto Prefectural University of Medicine, Kyoto, Japan. [45]Environmental Metagenomics, Research Center One Health Ruhr of the University Alliance Ruhr, Faculty of Chemistry, University of Duisburg-Essen, Essen, Germany. [46]School of Life Sciences, Technical University of Munich, Freising-Weihenstephan, Germany. [47]Department of Engineering Mathematics and Physics, Faculty of Engineering, Cairo University, Giza, Egypt. [48]Centre for Research in Earth and Space Sciences, York University, Toronto, Ontario, Canada. [49]Geological Institute, Department of Earth Sciences, ETH Zurich, Zurich, Switzerland. [50]Department of Mechanical Engineering, MIT, Cambridge, MA, USA. [51]Ragon Institute of MGH, MIT and Harvard, Cambridge, MA, USA. [52]Jameel Clinic for AI & Healthcare, MIT, Cambridge, MA, USA. [53]School of Marine and Atmospheric Sciences, Stony Brook University, Stony Brook, NY, USA. [54]Advanced Analytical Technologies, Empa, Dübendorf, Switzerland. [55]Division of Biomedical Engineering, James Watt School of Engineering, University of Glasgow, Glasgow, UK. [56]Timegate Instruments Oy, Oulu, Finland. [57]Drug Research Program, Division of Pharmaceutical Biosciences, Faculty of Pharmacy, University of Helsinki, Helsinki, Finland.

✉e-mail: leeka@ethz.ch; ugis@ebi.ac.uk; romanstocker@ethz.ch

Published online: 7 May 2024

References

1. Sarkans, U. et al. *Nucleic Acids Res.* **46**, D1266–D1270 (2018).
2. Huang, W. E., Griffiths, R. I., Thompson, I. P., Bailey, M. J. & Whiteley, A. *Anal. Chem.* **76**, 4452–4458 (2004).
3. Alcolombri, U., Pioli, R., Stocker, R. & Berry, D. *ISME Commun.* **2**, 55 (2022).
4. Wagner, M. & Haider, S. *Curr. Opin. Biotechnol.* **23**, 96–102 (2012).
5. Pereira, F. C. et al. *Nat. Commun.* **11**, 5104 (2020).
6. Lee, K. S. et al. *Nat. Rev. Methods Primers* **1**, 80 (2021).
7. Hatzenpichler, R., Krukenberg, V., Spietz, R. L. & Jay, Z. J. *Nat. Rev. Microbiol.* **18**, 241–256 (2020).
8. Wang, Y., Huang, W. E., Cui, L. & Wagner, M. *Curr. Opin. Biotechnol.* **41**, 34–42 (2016).
9. Gala de Pablo, J., Lindley, M., Hiramatsu, K. & Goda, K. *Acc. Chem. Res.* **54**, 2132–2143 (2021).
10. Guo, S., Popp, J. & Bocklitz, T. *Nat. Protoc.* **16**, 5426–5459 (2021).

Acknowledgements

We gratefully acknowledge funding from a Gordon and Betty Moore Foundation Symbiosis in Aquatic Systems Initiative Investigator Award (GBMF9197 to R.S.; https://doi.org/10.37807/GBMF9197) and the Simons Foundation through the Principles of Microbial Ecosystems (PriME) collaboration (grant no. 542395 to R.S.). Development of BioStudies is supported by the European Molecular Biology Laboratory member states. A.J.P. was supported by the German Research Foundation (DFG)—CRC 1439/1, project number 426547801. The Stocker group thanks R. Naisbit for scientific editing.

Author contributions

K.S.L., Z.L., A.A., U.S. and R.S. wrote the paper. All authors reviewed and approved the paper.

Competing interests

The authors declare no competing interests.