nature protocols

PROTOCOL

https://doi.org/10.1038/s41596-020-00427-8

Check for updates

Optofluidic Raman-activated cell sorting for targeted genome retrieval or cultivation of microbial cells with specific functions

Kang Soo Lee[1]✉, Fátima C. Pereira[2], Márton Palatinszky[2], Lars Behrendt[3], Uria Alcolombri[1], David Berry[2,4], Michael Wagner[2,5]✉ and Roman Stocker[1]✉

Stable isotope labeling of microbial taxa of interest and their sorting provide an efficient and direct way to answer the question "who does what?" in complex microbial communities when coupled with fluorescence in situ hybridization or downstream 'omics' analyses. We have developed a platform for automated Raman-based sorting in which optical tweezers and microfluidics are used to sort individual cells of interest from microbial communities on the basis of their Raman spectra. This sorting of cells and their downstream DNA analysis, such as by mini-metagenomics or single-cell genomics, or cultivation permits a direct link to be made between the metabolic roles and the genomes of microbial cells within complex microbial communities, as well as targeted isolation of novel microbes with a specific physiology of interest. We describe a protocol from sample preparation through Raman-activated live cell sorting. Subsequent cultivation of sorted cells is described, whereas downstream DNA analysis involves well-established approaches with abundant methods available in the literature. Compared with manual sorting, this technique provides a substantially higher throughput (up to 500 cells per h). Furthermore, the platform has very high sorting accuracy (98.3 ± 1.7%) and is fully automated, thus avoiding user biases that might accompany manual sorting. We anticipate that this protocol will empower in particular environmental and host-associated microbiome research with a versatile tool to elucidate the metabolic contributions of microbial taxa within their complex communities. After a 1-d preparation of cells, sorting takes on the order of 4 h, depending on the number of cells required.

Introduction

Enormous advances in sequencing technology and mass spectrometry, in conjunction with powerful bioinformatic tools (i.e., systems biology), have led to the adoption of 'omics' (e.g., genomics, transcriptomics, proteomics, metabolomics) as the preferred means to obtain comprehensive genetic and biomolecular profiles of microbial communities such as those that play fundamental roles in aquatic and soil ecosystems and the mammalian gastrointestinal tract[1–6]. In particular, metagenomics (genome sequencing of microbial communities) has been used to investigate the identity and potential functions (who does what?) of microbes of interest within complex microbiomes. For example, the public genome database managed by the US Department of Energy Joint Genome Institute provides up-to-date information about ~313,000 organisms[7]. Single-cell genomics provides a complementary and powerful approach with which to decipher the heterogeneities in genetic composition within microbiomes—for example, within guilds of closely related species and among strains or individuals within species—that cannot be captured via metagenomic approaches[8]. Yet these methods have not reached their full potential because of several limitations. First, detecting the presence of a gene does not necessarily indicate that the cells are actively carrying out the associated function, because metagenomics data do not provide information concerning expression, and thus complementary transcriptomic and/or proteomic analyses are required. Second, these approaches

[1]Institute for Environmental Engineering, Department of Civil, Environmental and Geomatic Engineering, ETH Zurich, Zurich, Switzerland. [2]Centre for Microbiology and Environmental Systems Science, Division of Microbial Ecology, Department of Microbiology and Ecosystem Science, University of Vienna, Vienna, Austria. [3]Science for Life Laboratory, Department of Environmental Toxicology, Uppsala University, Uppsala, Sweden. [4]Joint Microbiome Facility of the Medical University of Vienna and the University of Vienna, Vienna, Austria. [5]Center for Microbial Communities, Department of Chemistry and Bioscience, Aalborg University, Aalborg, Denmark. ✉e-mail: leeka@ethz.ch; wagner@mirobial-ecology.net; romanstocker@ethz.ch

require substantial investment of resources and time and rely on correct gene annotation (which is often impossible to achieve in silico) to identify microbes performing a function of interest[8].

Stable isotope probing (SIP) techniques using deuterium, carbon-13 (^{13}C), or nitrogen-15 (^{15}N) can overcome the first issue, because only the cells that are metabolically active under the conditions of interest become labeled. SIP methods thus provide direct functional insights that can reach the resolution of single cells and thereby provide an efficient route for the discovery of novel ecological pathways and metabolisms[9–15]. Spontaneous Raman microspectroscopy provides a powerful approach with which to interrogate the chemistry of individual cells in situ on the basis of inelastic scattering generated as a result of a laser interacting with the chemical bonds within the cell[16]. By combining SIP methods with the detection of labeled cells on the basis of their Raman spectra, metabolically active taxa of interest can be sorted from within microbial communities for downstream analysis by, for example, DNA sequencing (i.e., genome-resolved, targeted 'mini-metagenomics' or single-cell genomics), thereby addressing the second issue and directly linking the functions of taxa (phenotypes) within the community to their genomes (genotypes)[17,18]. As in the contrast between metagenomics and single-cell genomics described above, this strategy can reveal microbial diversity of functional roles at the single-cell level that is not possible via bulk approaches (e.g., DNA SIP[19]). However, sorting approaches are still in their infancy, and their low throughput (1–2 cells per h by manual sorting; see 'Comparison with other methods' section) and destructive nature (cells lose their viability during sorting as a result of laser-induced photophoretic damage) have hindered broader adoption[17,20–23]. In this protocol, we describe a pipeline to make a direct link between cell phenotypes and genotypes in a high-throughput and non-destructive manner using a novel Raman-activated cell sorting (RACS) platform that enables downstream DNA analysis of sorted cells of interest[24,25] or cultivation for further phenotypic characterization.

Overview of the procedure

The platform relies on optofluidic techniques to enable automated sorting of live cells at high throughput (up to 500 cells per h; two orders of magnitude higher than manual sorting; see 'Comparison with other methods' section), using optical tweezers and microfluidics to manipulate cells to measure their Raman spectra individually and collect the cells of interest for downstream analysis or cultivation. The protocol comprises three main stages: (i) sample preparation (Step 1 in Fig. 1), (ii) sorting of target cells (Steps 2–52), and (iii) downstream analysis (Steps 53–58).

In this protocol, we mainly focus on two key components. First, we describe the general procedures for sample preparation, complemented with alternatives adapted to different sample types, for example, intestinal samples, which must be treated under anaerobic conditions, or a marine sediment sample containing cytochrome *c*–expressing microbes, which requires an enrichment step to increase the initially low relative abundance of cells of interest within the community. To identify taxa of interest within complex microbial communities for sorting, we outline two approaches: short incubation of samples in a stable isotope-containing medium to mark target cells with the stable isotope (e.g., deuterium, ^{13}C) or use of samples directly when the taxa of interest contain specific substances that are easily detectable in their Raman spectra (Fig. 1).

Second, we describe the automated RACS procedures. We provide a description of the peripheral (simple) engineering required to integrate two hardware components, a dual camera setup and laser shutters, that are not available in commercial Raman systems, yet are necessary to work with our protocol (see 'Equipment setup': 'Lower microscope' and 'Laser shutter'). To enable flexible adoption of the method for various applications, we describe the versatility and tunability of the pipeline for samples that benefit from being sorted with respect to other molecular fingerprints in lieu of the carbon–deuterium (C–D) bond (used in most of this protocol), for example, ^{13}C (a stable isotope that is commonly used to track the carbon metabolism of microbes[15]) or carotenoids (as a marker of photoautotrophic microorganisms[26], although these pigments are also produced by some non-photoautotrophic bacteria[27]) (Fig. 1).

In addition, we describe a new software that implements a machine-learning algorithm (specifically *K*-means clustering[28,29]) that enables automation of the choice of selection criteria used during sorting of taxa of interest. There is an initial step of sample assessment in which the software learns sample characteristics before the actual sorting, thereby taking into account variation in the chemical composition of the cells that can change their Raman spectra.

Finally, we discuss possible downstream methods for the subsequent cultivation of the collected cells for further phenotypic characterization ('Experimental design': 'Downstream analysis').

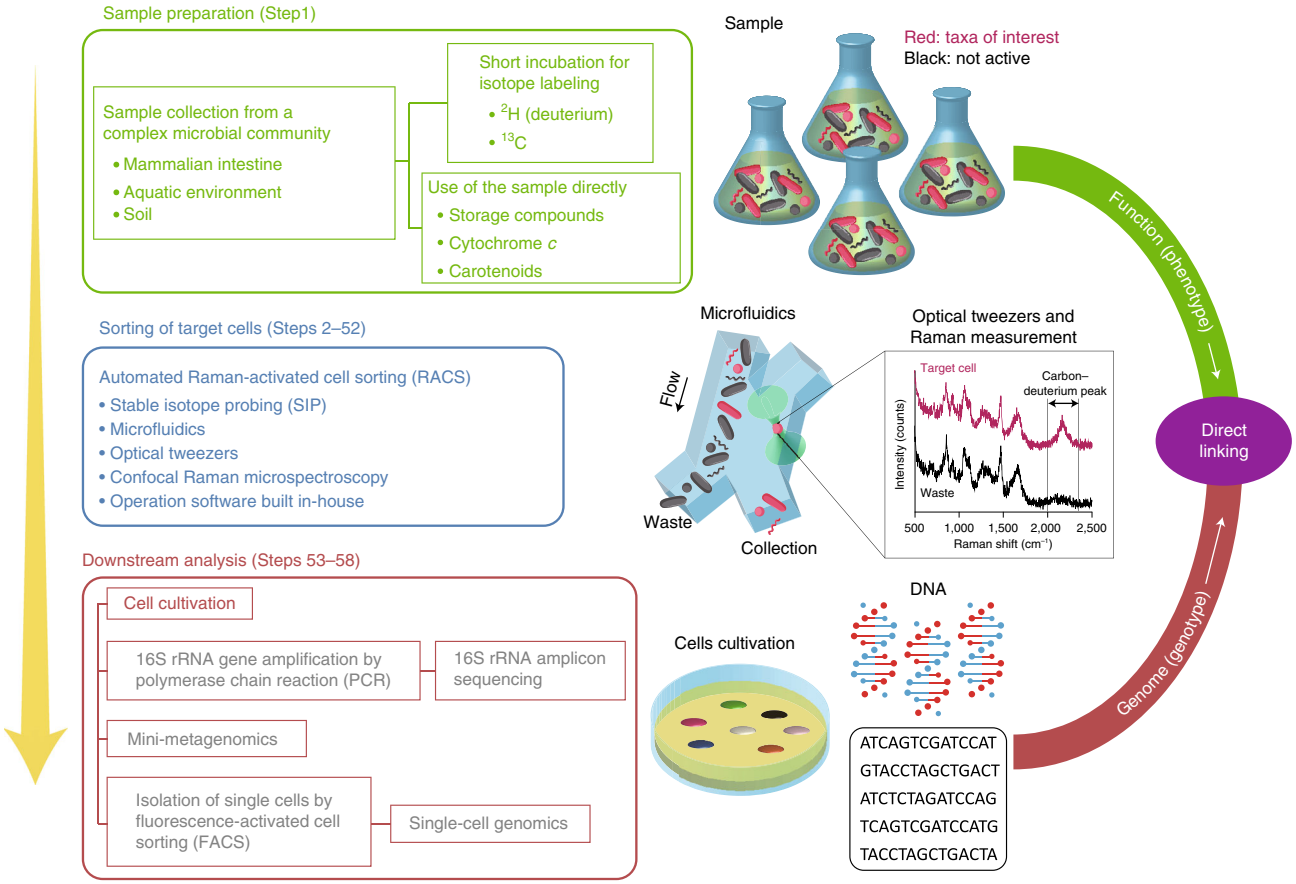


Fig. 1 | Workflow describing the direct linking between phenotypes and genomes of microbial cells from complex communities. Sample preparation (Step 1): A natural complex community sample (e.g., from mammalian gut, aquatic or soil ecosystem) is briefly incubated for stable isotope labeling (e.g., using deuterated water or ^{13}C-labeled compound) under conditions of interest (e.g., in minimal medium supplemented with a compound of interest) so that only active cells (i.e., those that metabolize that compound) become labeled. Alternatively, if the taxa of interest contain a molecule that generates a strong Raman signal (e.g., storage compounds or resonance Raman scattering in the presence of carotenoids, cytochrome *c*), the sample can be used directly without stable isotope labeling. Sorting of target cells (Steps 2–52): The cells that play a specific functional role in their community are identified by interrogating the corresponding chemical fingerprints in their Raman spectra (e.g., carbon–deuterium peak) and are sorted using an automated RACS platform that couples stable isotope probing (SIP), microfluidics, optical tweezers, confocal Raman microspectroscopy, and operation software built in-house. Downstream analysis (Steps 53–58): The collected cells after sorting are used for further characterization (via cell cultivation) or for genome sequencing, thereby directly linking the functional roles of the specific taxa within the communities to their genes. In this protocol, aspects in colored text are covered, whereas methods in gray text have been described in detail elsewhere[30].

Genomic characterization of the sorted cells involves well-established techniques for DNA analysis that are available in many research facilities (16S rRNA gene amplicon sequencing and mini-metagenomics; Fig. 1) and that are described in the literature (e.g., single-cell genomics after the isolation of individual cells by fluorescence-activated cell sorting (FACS); ref. [30]; Fig. 1); thus, we outline the possibilities but do not describe these aspects in detail.

Development of the protocol

Our pipeline leverages the integration of five major technologies to enable sorting of cells on the basis of their Raman spectra ('Sorting of target cells' in Fig. 1): (i) SIP, (ii) microfluidics, (iii) optical tweezers, (iv) confocal Raman microspectroscopy, and (v) a software platform built in-house to automate the sorting procedure. Within a microfluidic device, cells are manipulated with optical tweezers to enable their Raman spectra to be measured individually and cells of interest to be collected (Fig. 2). We initially considered using a single-file stream of cells in the microfluidic device and performing real-time measurement without optical tweezing, followed by downstream sorting using an external force (e.g., an electric force; a strategy similar to that used in FACS machines[31]). However, it proved challenging to tightly focus 1-μm cells with conventional microfluidic focusing techniques (e.g., via hydrodynamic or electric forces[32,33]) with sufficient precision to measure their Raman signal reliably and reproducibly. We thus used optical tweezers to capture individual cells

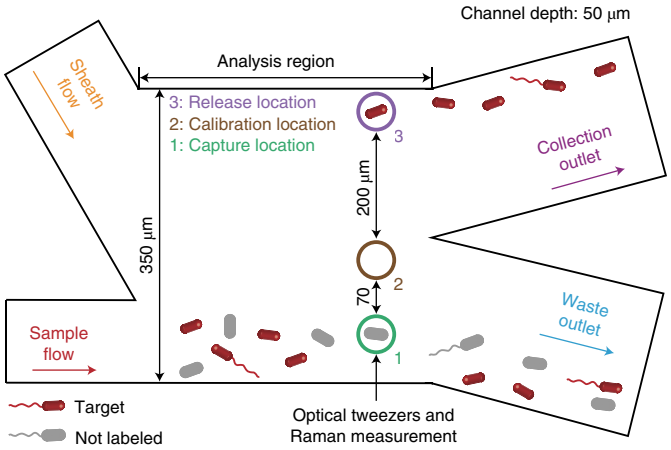


Fig. 2 | Design of the microfluidic device to sort cells of interest via RACS. The sample (containing suspended single cells) is introduced and focused onto one side of the channel by sheath fluid, so that it traverses an analysis region and exits through the waste outlet by default. Within the sample flow, lasers for optical tweezers and Raman measurement are focused using the same objective. The operation software detects the capture of a single cell in the optical tweezers (by measuring P_C; see main text) and quasi-simultaneously calculates its deuterium-labeling status (by calculating P_L). If the cell is identified as not labeled, it is immediately released and the flow carries it into the waste outlet. If the cell is identified as labeled, on the other hand, the stage moves from the 'capture' (1) to the 'release' (3) location, the cell is released and is carried by flow through the collection outlet, and the stage moves back to the capture location. The software can automatically repeat this process. The calibration location (2) is used for the initial Raman measurement of the working fluid in the absence of cells.

from the sample stream and briefly immobilize each captured cell within a Raman interrogation volume for measurement with subsecond laser exposure (0.3 s; ref. [24]).

To automate the sorting procedures, we used two working parameters, the cell and labeling indices, P_C and P_L, respectively, which are based on initial test measurements. Upon capture of a single cell in the optical tweezers, the overall Raman intensity increases: P_C measures this, thereby enabling automated detection of cell capture. Taxa of interest, for example, cells labeled with deuterium (used throughout most of this protocol), can be identified by a characteristic fingerprint in their Raman spectra, in this case the 'carbon–deuterium' (C–D) peak (spectral region of 2,040–2,300 cm^{-1}): P_L monitors this.

Selection of a suitable working fluid for the sorting is critical to achieving high throughput (Fig. 3). The medium used to isotopically label the taxa of interest during a brief incubation (D_2O) also generates a Raman signal (Fig. 3a,b) that can (depending on the medium composition) strongly interfere with the calculation of P_C and P_L values, so before analysis the sample is resuspended in a D_2O-free, non-fluorescent medium. We initially used Milli-Q water for this purpose; however, this caused osmotic stress that made the cells vulnerable to damage upon exposure to the 532-nm Raman laser and, as a result, the maximum laser power that could be used during RACS was 15 mW (see Box 1 for a guide to the selection of the laser power). To enable the use of higher laser power for sorting (the laser power determines the laser exposure time necessary for the reliable calculation of P_C and P_L, and thus the sorting throughput), we then used isotonic liquids to match the osmolarity of the original incubating medium, for example, glycerol at a certain concentration (balanced with Milli-Q water) for intestinal and freshwater samples, and artificial seawater (ASW) for marine samples (see Reagent setup) (inset in Fig. 3b). This minimized the osmotic stress to the cells and enabled the use of a Raman laser power (quantified at the level of the sample) of 80–300 mW (depending on the sample type; see Box 1) without causing cell damage.

To optimize the Raman measurements, we engineered two key components into the Raman microspectroscope system. First, to monitor the RACS process in real time, we installed a second microscope built in-house (lower microscope in Fig. 4). Compared with other microscopy techniques, including fluorescence detection, the Raman system relies on an intrinsically weak scattering signal (only 1 out of ~10^6 photons undergoes Raman scattering). As a consequence, a removable beam splitter (splitting the laser in two) lies within the laser path during the 'imaging' mode (to enable the search for a region of interest at the same time as focusing of the laser occurs), but this is removed from the laser path in the 'Raman' mode to obtain a stronger signal during measurement of the Raman spectrum (Fig. 4a). The second microscope below the sample was installed to enable

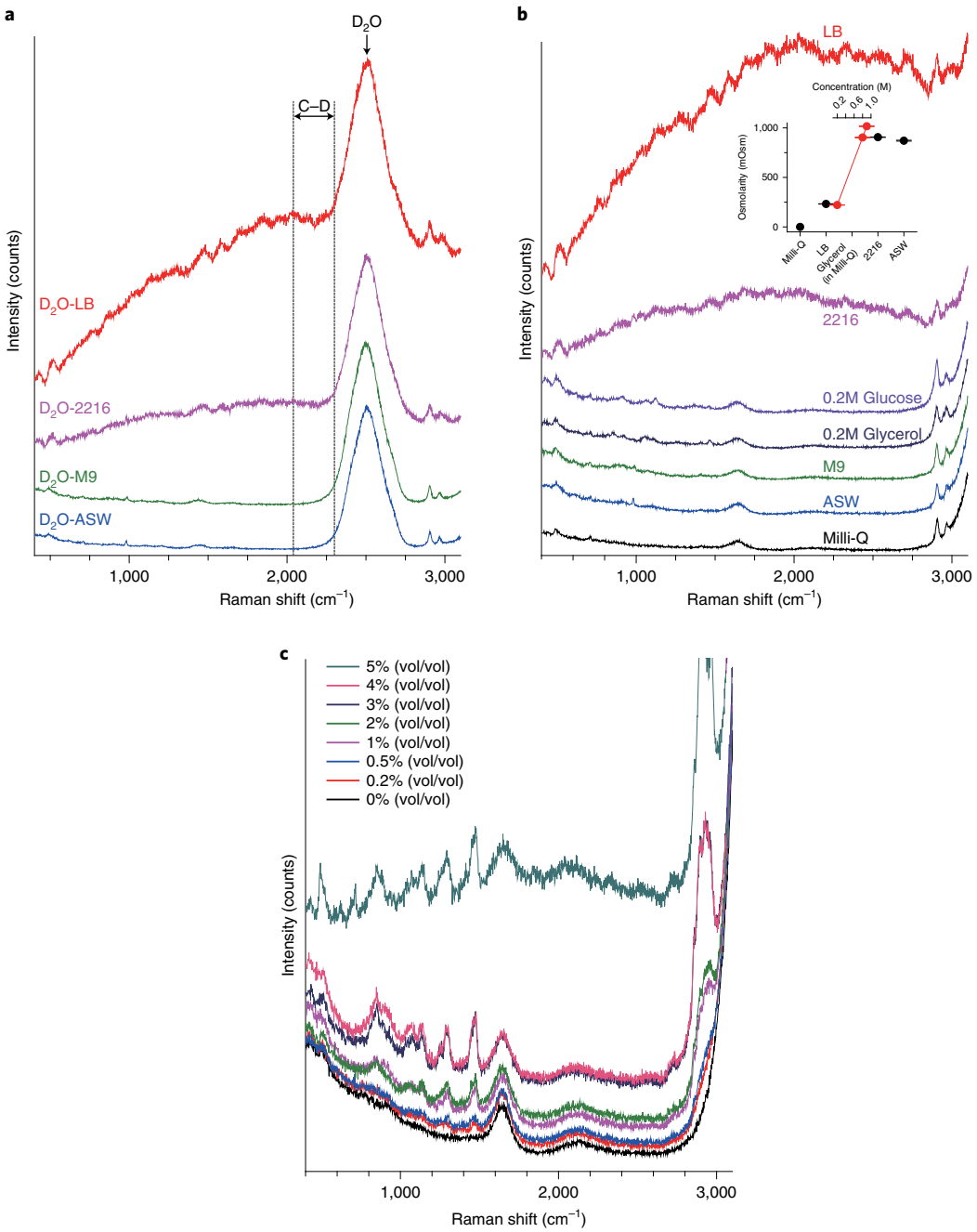
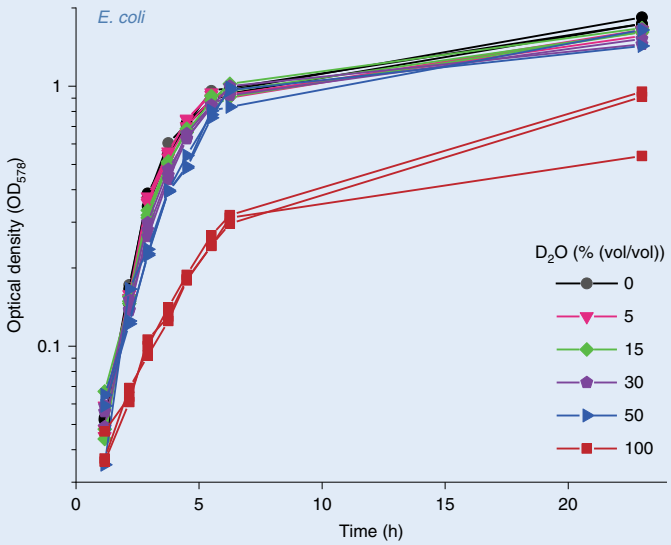


Fig. 3 | Selection of working fluid for the RACS. a, Interference in the detection of the C–D peak caused by residual D_2O in the medium. The peaks for D_2O (2,250–2,750 cm^{-1}) and C–D (2,040–2,300 cm^{-1}) overlap, and thus the presence of D_2O in the medium during RACS inhibits the detection of the C–D peak. Curves show measurements of the different D_2O-containing media in the absence of cell. **b**, Raman spectra of several media. LB and 2216 media have strong fluorescence signals, and are thus not suitable to be used for RACS. Four other media (0.2 M glucose, 0.2 M glycerol, M9, and ASW) generate a Raman spectrum similar to that of Milli-Q water (Milli-Q) and thus can be used for RACS. All spectra are offset on the *y*-axis for visibility. Inset: Osmolarity of four media and of glycerol (balanced with Milli-Q water) at three concentrations (in red). Data points and error bars represent means ± s.d. of three replicates. **c**, Raman spectra of 0.2 M glycerol supplemented by different amounts of Tween 20 (% (vol/vol) in the legend). The spectra are not offset. The inclusion of Tween 20 above 0.5% (vol/vol) substantially increases the overall spectral background.

monitoring of the RACS process while the system is in Raman mode. Second, we used a high-power objective (specifically, a 63×, 1.2–numerical aperture (NA), chromatic aberration–corrected, water-immersion objective) for the Raman measurement and the optical tweezing (operating via a 1,064-nm laser) to satisfy three criteria.

Box 1 | Choice of parameters

Determination of optimal D_2O concentration in the medium and time for incubation ● Timing **variable depending on sample type**
The optimal concentration of D_2O in the medium for labeling is a trade-off: the higher the D_2O concentration in the medium, the greater the sensitivity of the Raman measurement, but the greater the chance of toxic effects on the microbes, to an extent that depends on the physiology of the microbes. The appropriate concentration should thus be estimated for each application, followed by selection of an appropriate incubation time for labeling. To this end, prepare cell samples in media of different D_2O concentrations (e.g., 0 (for the control), 5, 15, 30, 50, 100% (vol/vol)) and compare growth over time. Several approaches are possible for quantifying growth, for example, by glucose consumption for the mouse colon sample, by optical density at 578 nm (OD_{578}) for pure bacterial cultures (see the figure: LB media of different D_2O concentrations were used to culture the *E. coli* samples; on the basis of this evaluation, 50% D_2O-containing LB medium was used in this protocol; this figure has been adapted from our previous paper[17]; Supplementary Fig. 3A). After selection of the optimal D_2O concentration, monitor the labeling of the sample over time. For instance, incubation times of 1 h (for many pure bacterial cultures) or 6 h (for the mouse colon sample)[17] in the presence of 50% D_2O result in detectable levels of D being incorporated into metabolically active cells (as assessed by Raman microspectroscopy). To avoid cross-feeding, sample incubation must be performed for short time periods. By contrast, if the objective is to decipher microbial food webs, longer incubation times will be necessary to ensure labeling of secondary and tertiary consumers.

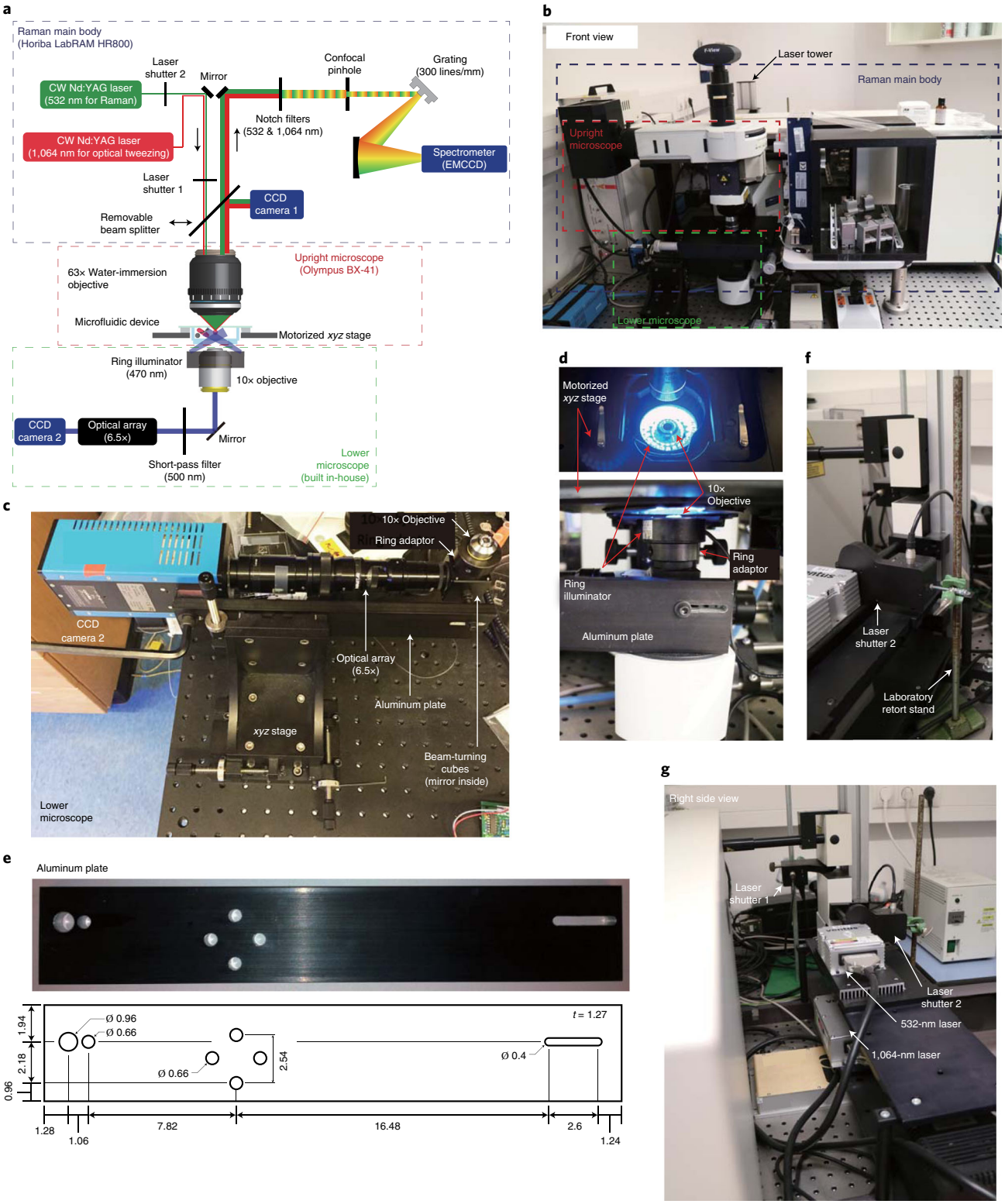


Optimal Raman laser power for the RACS ● Timing **10–20 min**
Susceptibility to laser exposure (for the 532-nm Raman laser) varies depending on sample type, thus the optimal power of the Raman laser should be determined for each application. To this end, in Step 50 gradually increase the power from 80 mW and observe with the lower microscope whether a cell captured in the optical tweezers disappears within a few seconds (because of laser-induced cell damage). The maximum power at which the cell maintains its shape can be used for RACS, although this criterion does not exclude damage that hampers subsequent cultivation of the sorted cells[77]. Our demonstrations using a pure bacterial culture (*M. adhaerens*[24]) and the mouse colon microbiota (Fig. 7e) suggest that this selection criterion is compatible with the recovery of viable cells after the RACS procedures.

1 The number of Raman-scattered photons is proportional to the laser intensity. Considering this, tightly focusing the laser beam to a limited area (comparable to the size of a typical microbe, ~1 μm) and maximizing the ability to gather light from the cell (using a high magnification and high-NA objective) is key to obtaining an optimal Raman signal for a given laser power and exposure time. We therefore used the highest magnification and highest-NA water-immersion objective that we found to be commercially available.

2 The ability to capture a single cell from within the sample stream (i.e., the optical tweezing efficiency) is determined by the relative magnitudes of the optical tweezing force and fluid drag force. In general, the smaller the laser spot size (achieved using a high magnification and high-NA objective), the greater the optical tweezing (or 'optical gradient') force.

3 Minimizing chromatic aberration is critical to achieving reliable Raman measurement. The Raman and optical tweezers lasers are focused using the same objective, and although the objective is designed to focus the two lasers at the same location (i.e., chromatic aberration is corrected), in practice there is a small focus difference between them. We tested five commercial objectives from different manufacturers, and on the basis of this testing, chose a Zeiss objective (C-Apochromat 63×/1.20-W Korr UV-VIS-IR M27; Fig. 5 and Table 1). We chose a water-immersion (rather than oil-immersion) objective to minimize chromatic aberration due to the discrepancy in the refractive index between the aqueous medium in the microfluidic device and the immersion liquid.



Flow stability within the microfluidic device determines the reproducibility of RACS because fluctuations in flow may carry the unsorted sample stream directly into the collection outlet (Fig. 2), thereby generating false positives. The RACS platform uses very low flow velocity (500 μm/s) to ensure that the optical tweezing force is sufficient to capture and hold cells within the sample stream. The microfluidic device consists of two inlets (for the sample and the sheath flow used to focus the

◀ **Fig. 4 | Raman microspectroscope system. a**,**b**, Schematic (**a**) and photograph (**b**) of the Raman microscope system used in this protocol. Two lasers (532 nm and 1,064 nm for Raman measurement and optical tweezing, respectively) are focused within the microfluidic device (Fig. 2) using a 63× water-immersion objective and CCD camera 1. After placing the laser spot on the cell stream, the removable beam splitter is withdrawn (CCD camera 1 is deactivated) and the RACS process begins. Rayleigh scattering (of the same wavelengths as the incident lasers, i.e., 532 and 1,064 nm) and Raman signal coming from out of focus regions are removed using notch filters (cutoffs: 532 and 1,064 nm) and confocal pinhole, respectively. Raman signal generated from the confocal volume (i.e., the cell) is then divided with respect to the wavelength by the grating (300 lines/mm, which can cover the 400–3,300 cm^{-1} spectral region with a single Raman measurement) and measured by spectrometer. Because CCD camera 1 is withdrawn during RACS, the process can be monitored from below using the lower microscope (green dashed box), which consists of a 10× objective, a mirror (to bend the light by 90°), a short-pass filter (cutoff: 500 nm, to avoid damage to the CCD camera by laser exposure), an optical array (magnification 6.5×) and CCD camera 2. Laser shutters 1 and 2 block both lasers and the 532 nm laser alone, respectively. **c**, Photograph of the lower microscope, composed of CCD camera 2, optical array, beam-turning cubes (the mirror in panel **a** is inside), and 10× objective, all assembled on the *xyz* stage using an aluminum plate fabricated in-house (see panel **e**). A ring adaptor is equipped around the 10× objective to install a low-angle ring illuminator (see panel **d**). **d**, Top (upper panel) and front (lower panel) views of the low-angle ring illuminator placed on the ring adaptor. **e**, Aluminum plate with which to assemble the lower microscope assembly onto the *xyz* stage. The holes fabricated using a milling machine on the left (two), center (four), and right (one elongated) are to enable fixing of CCD camera 2, the *xyz* stage, and the beam-turning cubes, respectively, using screws (each of the three components has a screw thread). All dimensions in centimeters. **f**, Photograph of laser shutter 2 mounted on a laboratory retort stand. **g**, Configuration of the laser tower for the two lasers and the two laser shutters.

Table 1 | Three important specifications of the objective related to the ability to conduct this protocol

Objective	Difference between two laser foci (μm)	Raman intensity[a]	Optical tweezing[b]
Olympus (UPLSAPO60XW)	1.5–2.0	1.0	Yes
Olympus (UPLAPO60XW)	1.5–2.0	1.0	No
Leica (HC PL Apo 63×/1.2W CORR CS2)	2.5–3.7	1.0	No
Leica (HCX Apo L 63×/0.90W U-V-I)	3.5–4.0	1.0	No
Zeiss (C-Apochromat 63×/1.2 W Korr UV-VIS-IR M27)	<0.5	1.0	Yes

The specifications ((i) distance between the foci of the two lasers; (ii) performance for Raman measurement (estimated by quantifying the Raman intensity at 400 cm^{-1}); and (iii) ability to perform optical tweezing under the experimental conditions) were measured using five commercial objectives. [a]To evaluate the light transmittance of the objective, Raman intensity at 400 cm^{-1} (normalized by that of the UPLSAPO60XW objective) is used as a criterion. [b]Optical tweezing performance is estimated based on the ability of the optical tweezers to capture a single cell under the experimental conditions: 500-μm/s flow velocity and 500-mW laser power.

sample stream) and two outlets (for the collection and waste streams) (Fig. 2). Maintaining a stable pressure difference between the outlets was especially difficult to achieve at this flow speed (susceptible to small changes in flow over time) when the end of the microfluidic tubing that connected to each outlet was open (i.e., open to the liquid reservoir or air). We therefore use three syringes to control the flow at the two inlets (syringe pumps in 'injection' mode) and the collection outlet (a syringe pump in 'withdrawal' mode), whereas the flow at the waste outlet is passively controlled by the flow rates of the three syringes (Fig. 6).

We characterized and optimized the platform using four pure bacterial cultures, including Gram-negative and Gram-positive bacteria (*Escherichia coli*, *Salmonella typhimurium*, *Bacillus subtilis*, and *Marinobacter adhaerens*), showing that it achieves high throughput (200–500 cells/h) and high sorting accuracy (98.3 ± 1.7%; ref. [24]). We then demonstrated the functionality of the complete pipeline—from sample preparation to downstream DNA analysis—to sort and characterize mucin degraders within a complex microbiome sample from the mouse colon. Our approach revealed that members of the Muribaculaceae family, an abundant but under-characterized family present in the guts of homoeothermic animals, are key players in the process of mucin degradation. Finally, we demonstrated the tunability of the system by sorting a marine sediment sample on the basis of cytochrome *c* signals.

We have further optimized the RACS platform since our previous publication describing the approach[24]. In this protocol, we describe the latest version with the following improvements:

- *Calculation of* P_C *and* P_L. in the original version, when the software detected the capture of a single cell in the optical tweezers on the basis of P_C value (2-s measurement time), the cell was then translocated to the sample-free region and P_L was calculated on the basis of Raman spectra obtained with a longer

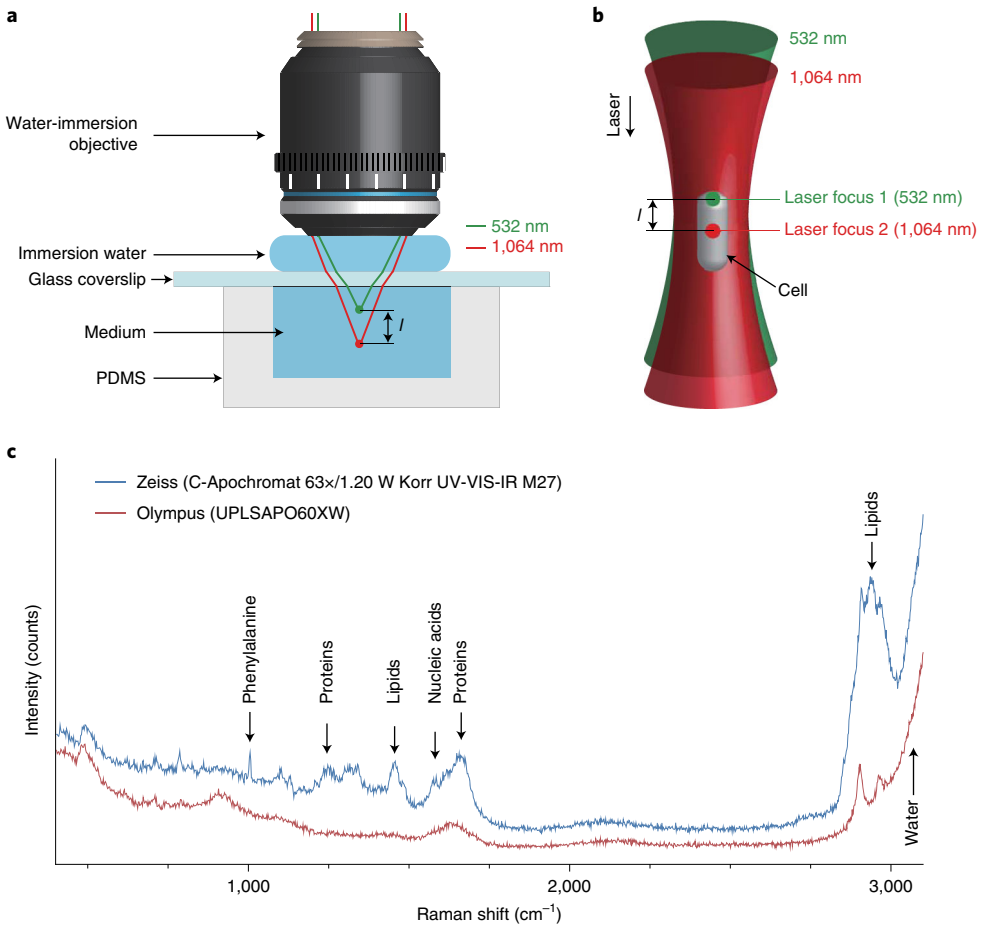


Fig. 5 | Considerations in the choice of an objective for the Raman measurement and optical tweezing. a, Schematic showing the focus difference due to chromatic aberration. The 532- and 1,064-nm lasers travel with different refraction angles at interfaces (objective–water–glass–water, by Snell's law) and thus their focus locations differ. **b**, When the (rod-shaped) cell is trapped in the focus of the optical tweezers (red dot) and the Raman measurement takes place elsewhere (green dot), the signal sensitivity is reduced. **c**, Raman spectra of a single *E. coli* cell when trapped with optical tweezers and measured using two objectives: Zeiss (l < 0.5 μm) and Olympus ($l \approx$ 1.5–2.0).

measurement time (5 s) to determine the deuterium-labeling status. Together with the time for the translocation of the cell (~1 s), ~8 s was needed in total to evaluate each cell. In the recent version, both P_C and P_L are simultaneously calculated with a 0.3-s measurement time, considerably increasing the sorting throughput (from 200 cells/h with the original version to 500 cells/h with the recent version). This is achieved by using isotonic liquid for cell resuspension (see above) and by using a second laser shutter (see 'Hardware' below), which together allow a higher laser power to be used, enabling more rapid Raman measurement.

- *Hardware*. A second laser shutter (laser shutter 2; Fig. 4a,f,g) blocks the Raman laser while the cell is being translocated from the capture location to the release location for collection (Fig. 2). This reduces the laser exposure to the cell by ~0.5 s and thus reduces the risk of laser-induced photophoretic damage; cells are exposed to the Raman laser for only 0.3 s during the sorting.
- *Software*. We have developed a new sorting software based on machine learning (that determines a sorting criterion in place of P_L), using the '*K*-means clustering' algorithm[28,29] to classify cells within the sample into mutually exclusive clusters before the actual sorting and, based on this, cells of interest can be sorted by discerning the cluster of cells measured in real time; cells assigned to the cluster(s) of interest are sorted.

Applications of the method

Our protocol is applicable across many fields of microbiology and environmental systems science, especially in environmental and host-associated microbiome research. By briefly incubating a

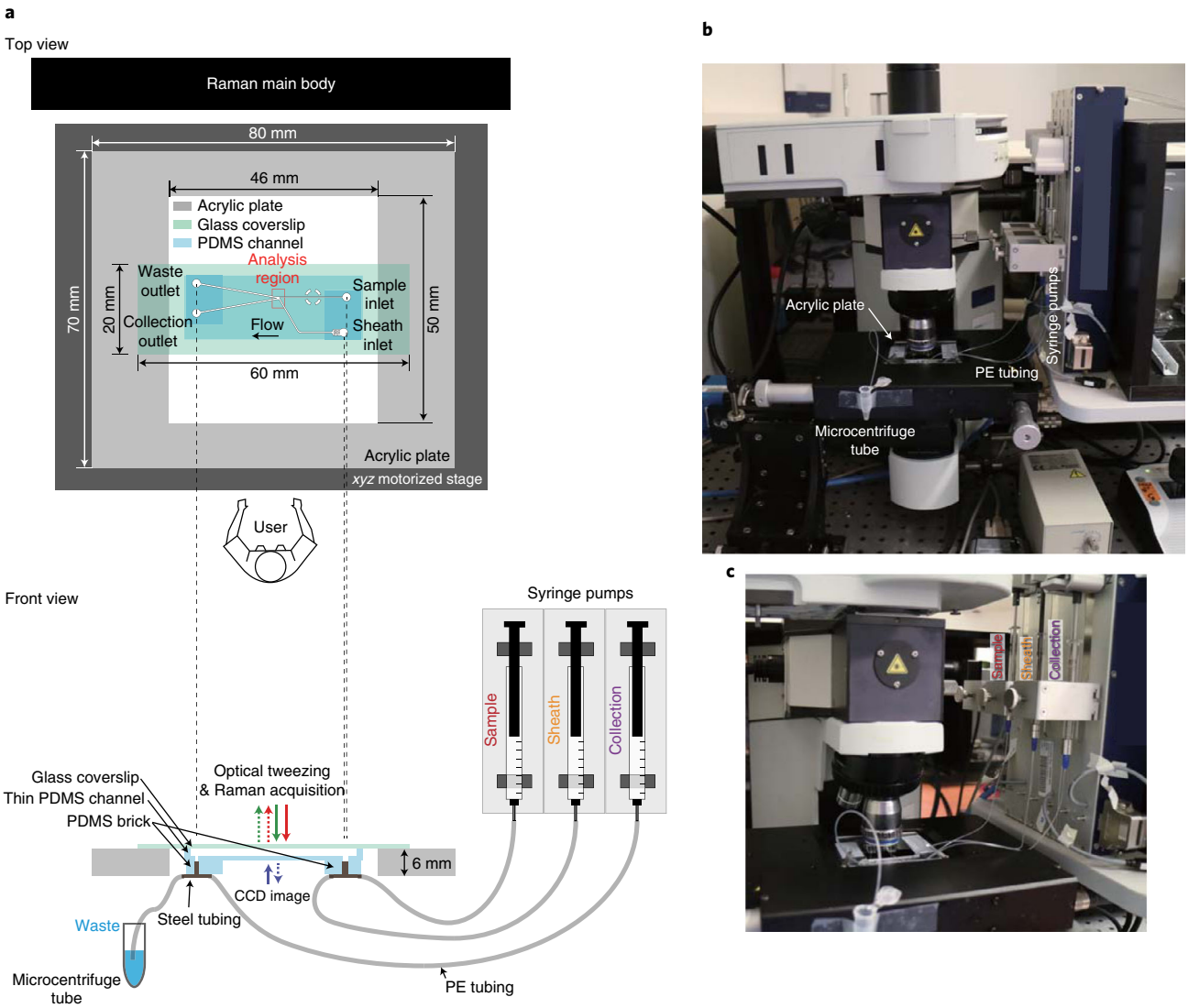


Fig. 6 | Installation of the microfluidic device and syringe pumps on the Raman microscope system. a– **c**, Schematic (**a**) and photographs of the microfluidic device on the acrylic support plate (**b**) and the tubing connections to the microfluidic device (**c**). The glass coverslip faces upward, toward the source of the lasers for both the Raman measurement and the optical tweezers. The thin (~1 mm) PDMS layer faces downward, where the Raman sorting is monitored using the lower microscope (Fig. 4). While installing the device, care should be taken to ensure that the lengths of tubing do not interfere with each other to avoid causing instability in the flow (because very low flow rates are used), and that the device is placed with the collection outlet closest to the user, to facilitate removal of cells once RACS is complete.

microbial community that contains taxa with a function of interest in D_2O-containing minimal medium (i.e., a medium that lacks nutrients for general cell growth, such as phosphate-buffered saline medium) supplemented with a specific chemical compound that can stimulate those functions, only the cells performing the function of interest will be labeled by deuterium and can then be sorted using the RACS platform for downstream analysis. This approach is particularly versatile because it does not require the substrate of interest itself to be labeled. Furthermore, the influence of other environmental parameters, such as temperature, pH, and salinity, on the specific activity of microbes can be evaluated. Potential applications, in addition to mucus degradation in the mammalian gut as we demonstrated[24], would include, for example, the sorting of microbes underlying nitrite oxidation and sulfur catabolism in the ocean[34,35], phosphate accumulation in wastewater treatment plants[36,37], and methane oxidation and sulfate reduction in marine seeps[38,39]. The sorted cells are available for DNA analysis and can be used for cultivation for further ecological evaluation, such as an interrogation of antibiotic resistance[40].

Our RACS platform is easily tunable, so that sorting microbes with respect to another stable isotope signal (in lieu of deuterium) is feasible by changing the spectral window considered. For

instance, ^{13}C (widely used in SIP approaches) enables tracking of carbon metabolic pathways by measuring the shift of the phenylalanine peak from 1,007 to 971 cm^{-1} as a result of the replacement of ^{12}C by ^{13}C (ref. [15]). This approach would enable, for example, the sorting of genetically engineered autotrophic *E. coli* from heterotrophic *E. coli* cells to study further evolutionary processes[41]. The RACS platform is also applicable to track predator–prey interactions[42,43]. The prey are labeled with ^{13}C before being exposed to a predator community, and then sorting of cells in terms of the ^{13}C–Raman signal in conjunction with downstream DNA analysis can reveal the identity and genomic potential of the predators.

Microbes possessing a specific molecule that generates a strong Raman signal can be directly sorted without relying on stable isotope labeling. For example, cytochromes[24,44], carotenoids, chlorophylls, vitamin B_{12}, heme, and rhodopsin generate sufficient Raman signal to be rapidly identified with the short laser exposure by virtue of the 'resonance' effect (which arises when the laser wavelength overlaps an electronic transition of the compound) that makes Raman measurement of these molecules much more sensitive (~10^3 times) than regular spontaneous Raman scattering[45]. Furthermore, sorting of cells containing storage compounds such as polyphosphate[36,37], polyhydroxybutyrate[36,37], glycogen[36,37], and sulfur[46] will be possible without stable isotope labeling.

The RACS platform can also be used as a counter for cells of interest without sorting, i.e., the platform measures and calculates the chemical fingerprint of interest upon cell capture in the optical tweezers and then immediately releases the cell back into the sample flow. This enables rapid characterization of novel microbiota, for example, to estimate the relative abundance of taxa of interest. In this operation mode, the throughput (the number of cells analyzed) can reach up to 1,000 cells per h. In addition to these diverse applications for the functional analysis of microbial communities, the platform is also applicable to investigating abiotic samples, for instance, the compositional identification of micro/nano-plastics in seawater and household sewage treatment systems[47–49]. In this protocol, we describe methods for implementing the RACS platform as a counter as well as a sorter.

Comparison with other methods

Throughput (how rapidly microbes are analyzed and sorted) determines the number of taxa of interest that can be retrieved from a sampled community. Considering the huge diversity of microbes and the variation in relative abundance among taxa within communities, obtaining a realistic estimate of community composition depends on analyzing large numbers of individuals. Previous work[17] was dependent on manually sorting cells one by one in a glass capillary (of rectangular cross-section, with inner dimensions of 0.1 × 1.0 mm). This approach allowed the sorting of only 1–2 cells per h (Supplementary Fig. 1). In this setup, (i) the capillary is filled with the sterile buffer, the sample is placed at one end of the capillary, and both ends are sealed with epoxy to prevent evaporation during sorting; (ii) single cells are searched using optical tweezers and, once a cell is captured, its Raman spectrum is measured; (iii) if the cell is identified as not labeled, it is immediately released by blocking the optical tweezers laser; (iv) if the cell is identified as labeled, it is translocated to the other end of the capillary (a distance of ~10 cm, which takes 7 min to traverse because more rapid movement causes a cell loss from the optical tweezers due to fluid drag) and released; (v) the objective is returned to the sample region and the steps are repeated for multiple cells; (vi) when the number of cells required for downstream analysis have been collected, the end of the capillary containing the collected cells is broken off, and the cells are recovered by brief centrifuging (<10 s) in a sterile tube. This method is extremely laborious, and user bias in applying the sorting criteria may occur. Moreover, during translocation of the target cells to the sterile end of the capillary, the cells are often lost because of attachment to the capillary wall.

Our RACS protocol achieves throughput that is two orders of magnitude greater than this manual sorting method, thereby making feasible function-based sorting of microbial taxa of interest in many environmental and host-associated microbiome ecosystems. To the best of our knowledge, our protocol is the first that provides a means to sort microbes down to a size of micrometers from within real, natural community samples on the basis of their Raman spectra. In the literature, current methods that rely on the use of Raman microspectroscopy in conjunction with automated microfluidic sorting platforms have been evaluated only with pure bacterial cultures, larger-sized eukaryotic cells, or cells containing target molecules that generate a strong resonance Raman signal (e.g., carotenoids; see above)[50,51]. The major advantage of our system is the integration of the optical tweezers, which enables a single cell to be held precisely within the Raman interrogation volume

during measurement. In other published Raman sorting methods, albeit feasible for specific samples, the use of a single-file stream of cells and their in-line measurement without optical tweezing[51] or the use of an electric force field to hold a cell during measurement (i.e., dielectrophoretic cell trapping)[50] does not provide sufficient precision to position the cell within the Raman interrogation volume for measurement; this reduces the sensitivity, reliability, and reproducibility of the measurements. Moreover, in another previously published method that does not use flow, cells are analyzed in dry conditions (which renders them more susceptible to damage by laser-induced heat in comparison with cells in suspension), and this limits the usefulness of cells collected for downstream analyses[20,21].

Use of coherent Raman scattering systems, for example, coherent anti-Stokes Raman scattering (CARS) or stimulated Raman scattering (SRS) spectroscopy, have the potential to provide even higher throughput. These technologies rely on a multiphoton process wherein the energy difference between two lasers (Stokes and pump lasers) is resonant with the vibrational frequency of a specific molecular bond of interest, thereby generating a 1,000 times higher signal intensity than regular spontaneous Raman microspectroscopy[52]. They were originally developed to achieve Raman-based imaging (i.e., scanning a field of view with respect to a single molecular fingerprint) and have recently been adapted to RACS applications[53–55]. To this end, a microfluidic device has been engineered to enable in-line measurement of a single-file stream of cells without optical tweezing. The system has been applied only to large-sized eukaryotic microalgae cells, yet integration of these variant Raman technologies has the potential to provide even more rapid and sensitive sorting.

In addition to Raman-based cell sorting, a few other methods exist that enable coupling genomic analysis to a specific function of microbial cells in a complex community[22,56,57]. For example, a combination of biorthogonal non-canonical amino acid tagging and FACS (BONCAT-FACS), followed by downstream DNA analysis, is very powerful in establishing direct links between cell phenotypes and genotypes[39] but is challenging to perform in certain environments with a chemical landscape that interferes with the labeling procedure.

Limitations

While using the pipeline described in this protocol, the following considerations should be kept in mind:

- Although the RACS system provides throughput that is two orders of magnitude greater than that of manual sorting methods, the number of sorted target cells strongly depends on the relative abundance of taxa with a function of interest within the sampled microbial community. When the taxa of interest are rare in the sample (e.g., when sorting mucin-degrading microbes in the mammalian gut microbiota[24]), whole-genome amplification (WGA) procedures, such as multiple displacement amplification (MDA), might be required to amplify the DNA to a quantity sufficient for subsequent metagenomic analysis. In this case, the method does not provide information on the relative abundance of each taxon within the sorted cells (i.e., a measure of the contribution of each operational taxonomic unit (OTU) to the function of interest).
- Our platform could potentially work for diverse cell types across the tree of life: prokaryotes (e.g., bacteria, as in this protocol[24], and archaea) and eukaryotes (e.g., algae, fungi, and isolated cells from multicellular organisms, such as tumor or blood cells). To enable diverse applications, reliable and consistent Raman measurement with a short laser exposure time (0.3 s) is critical, and the optical tweezing efficiency is the key to this. It is strongly dependent on the cell size at the given experimental conditions (magnification and numerical aperture of the objective, laser power, flow speed), whereas heterogeneity in cell morphology is of lesser influence because, regardless of cell shape, the cell is trapped at the equilibrium position within the optical force field[58,59]. In general, the larger the cell, the greater the optical tweezing efficiency; this can be theoretically calculated by tracing the photons ($d > \lambda$; where d and λ are the cell size and the laser wavelength, respectively)[60] or by solving the electromagnetic field ($d \sim \lambda$; ref. [61]). There is thus no upper limit to the size of cells that can be analyzed, as long as they can be manipulated within the microfluidic device. At the lower limit of cell size, we have successfully tested the platform for cell sizes down to 1 μm. Low optical tweezing efficiency may prevent use of the RACS protocol (or require further optimization) for small environmental bacteria or archaea, as well as for viruses (typically measuring 20–300 nm).
- The choice of the laser power is a delicate trade-off. Although using higher laser power enables rapid measurement and rapid decision-making based on the proposed RACS indices (P_C and P_L), it brings the risk of photophoretic damage to the cell due to laser exposure, reducing cell viability and thus potentially preventing downstream cultivation, or even causing cell lysis. The laser at visible

wavelength (i.e., the 532-nm Raman laser) is strongly absorbed by the microbes, whereas infrared wavelengths (i.e., the 1,064-nm optical tweezers laser) are subject to much weaker absorption effects[62]. Because the applicable laser power strongly depends on the species, it is not possible to make general recommendations that are optimal for all cases. We therefore provide guidelines to choose the laser power for the user's specific application (see Step 50 and Box 1).

- The cultivation of novel microbes requires optimization in each case. Previously, we used a pure bacterial culture to examine the feasibility of cell cultivation after RACS[24], and in this protocol, we provide further confirmation by describing the cultivation of mucin-degrading microbes after RACS, either on agar plates or in liquid medium (Fig. 7e). This shows that the laser exposure (for Raman measurement and optical tweezing) during RACS does not induce substantial photophoretic damage to cells, and thus our platform provides a means to cultivate the taxa of interest within the working pipeline. However, many microbes cannot be cultured using standard media and conditions; thus, cultivation of novel microbes as part of a study will require further investigation in each case to identify optimal growth conditions[63].
- The code built in-house (Supplementary Software 1 and 2) enables the RACS platform to be run within a commercial Raman microspectroscope (LabRAM HR800), and thus several commands are specific for the Horiba system. We include a footnote to those commands (marked by 'Horiba-specific') so that users can modify them as appropriate to the system available. The two other major Raman manufacturers (Renishaw and Bruker) have confirmed that their systems are compatible with control using third-party software (e.g., MATLAB, as in this protocol, or Python). Our code can also be adopted with suitable modifications for other Raman systems (including those built in-house) if they allow an application programming interface (API).

Expertise needed to implement the protocol

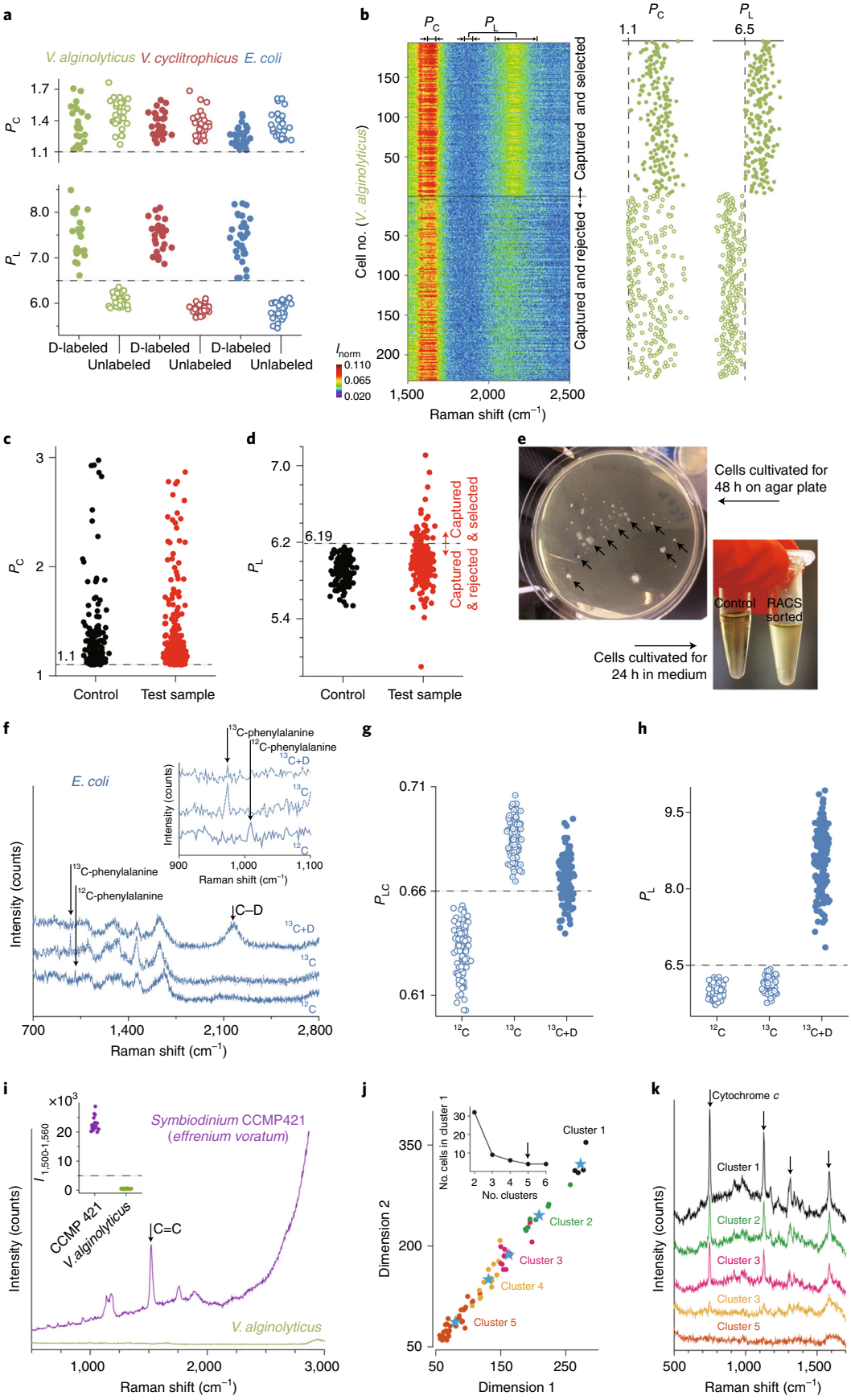
Operation of our pipeline requires little specialist knowledge. The protocol to set up the system can be replicated by any motivated researcher at a competent graduate student or postdoc level who has a basic knowledge of programming for instrument control. For broad adoption, simplicity of the system configuration and operation are key. Our RACS system includes several features that ensure a straightforward setup by a motivated user in the absence of special expertise. Raman microspectroscopes are not prohibitively expensive compared with other instruments that have been used in combination with SIP for the phenotypic identification of taxa of interest, such as nanoscale secondary ion mass spectrometry (nanoSIMS)[64], and are available in many research universities and industrial companies. The optical tweezers can be simply integrated into the Raman system by adding the tweezing laser into the same optical path as the Raman laser (Fig. 4a,g). Only basic microfluidics knowledge is required (see 'Experimental design': 'Sorting of target cells'). A detailed description of the software platform (Supplementary Software 1 and 2) is provided in this protocol (see Step 52; Fig. 8).

Experimental design

The workflow to make a direct link between cell phenotypes and genotypes via targeted isolation of taxa of interest from within microbial communities and their downstream DNA analyses or cultivation is depicted in Fig. 1. It comprises three major parts: sample preparation, sorting of target cells, and downstream analysis. We here describe how our protocol can be used for the analysis of samples from the mammalian mouse colon in which the taxa of interest (those that are involved in mucin degradation) are labeled with deuterium. In parallel, we describe how to apply our protocol to the analysis of laboratory model microorganisms (an autotrophic alga that contains carotenoids; heterotrophic bacteria labeled with deuterium, ^{13}C, or both) and a marine sediment sample containing cells that express high levels of cytochrome *c*.

Sample preparation

Samples collected from complex microbial communities (e.g., from the mammalian intestine, or from aquatic or soil environments) can be analyzed via two approaches: either by briefly incubating them to mark the taxa of interest with a stable isotope or by directly using them without labeling. The former is appropriate when the cellular functions of interest are not directly reflected in the Raman spectra of the respective microbes. Samples are briefly incubated (typically for a few hours[17]) in D_2O-containing 'minimal' medium supplemented with a specific compound that stimulates the function of interest, such as mucin, to uncover the taxa contributing to its metabolism. By comparison to a

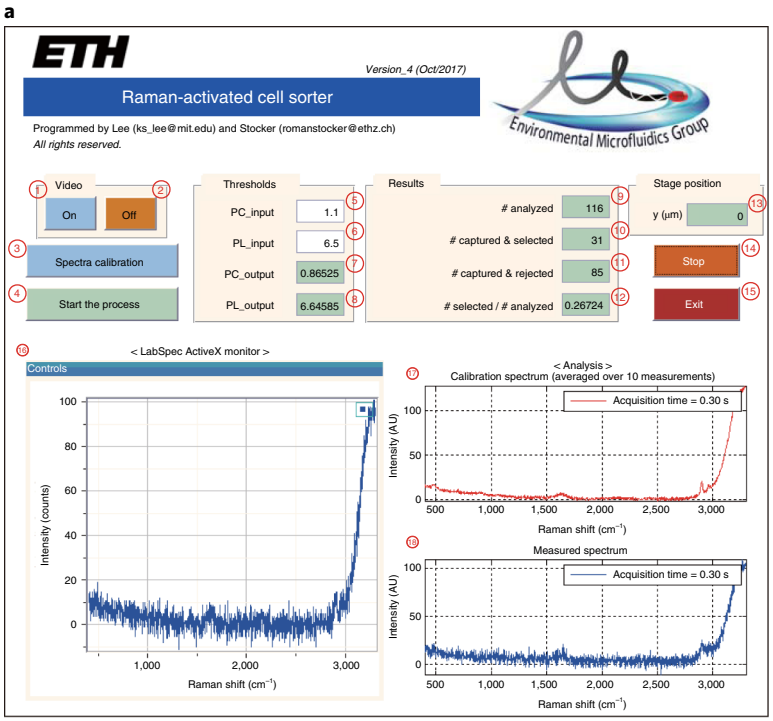

a
V. alginolyticus V. cyclitrophicus E. coli
P_C
P_L
D-labeled Unlabeled
D-labeled Unlabeled
D-labeled Unlabeled
b
P_C
P_L
Cell no. (V. alginolyticus)
Captured and selected
Captured and rejected
I_{norm}
Raman shift (cm^{-1})
c
Control
Test sample
d
6.19
Captured & rejected
Captured & selected
e
Cells cultivated for 48 h on agar plate
Cells cultivated for 24 h in medium
Control
RACS sorted
f
E. coli
^{13}C-phenylalanine
^{12}C-phenylalanine
C–D
^{13}C+D
^{13}C
^{12}C
Intensity (counts)
Raman shift (cm^{-1})
g
P_{LC}
h
P_L
i
Symbiodinium CCMP421 (effrenium voratum)
CCMP 421
V. alginolyticus
C=C
$I_{1,500\text{-}1,560}$
j
Dimension 1
Dimension 2
No. cells in cluster 1
No. clusters
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
k
Cytochrome c

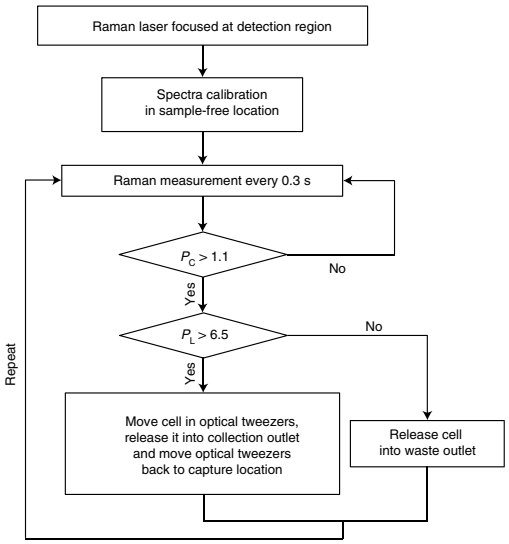
◀ **Fig. 7 | Anticipated results.** See main text for statistics: sample number, average, and standard deviation. **a**, P_C and P_L values for three pure bacterial cultures: *V. alginolyticus*, *V. cyclitrophicus*, and *E. coli*. Dashed lines represent the thresholds to detect the capture of a cell in the optical tweezers (P_C) and to distinguish between deuterium-labeled and unlabeled cells (P_L). **b**, Raman spectra of 'captured and selected' and 'captured and rejected' cells (*V. alginolyticus*; left) and their corresponding P_C and P_L values (right) using a 1:1 mixture of labeled and unlabeled cells. Values of 1.1 and 6.5 were used as thresholds for P_C and P_L, respectively, based on the measurements in **a**. **c**,**d**, P_C (**c**) and P_L (**d**) values measured in samples from a mouse colon microbial community. Based on the values measured using an unlabeled control sample (black points), $P_L = 6.19$ was chosen as a threshold for the RACS analysis of the community after deuterium labeling during incubation in a minimal medium supplemented with mucin (red points). **e**, Cells are viable and can be cultured after RACS. Cells (from the mouse colon sample) collected during RACS were cultured for 48 h or 24 h on a YCFA agar plate or in YCFA medium (shown for comparison with a control sample of YCFA medium containing no cells), respectively. Arrows in the left panel indicate colonies grown from single cells. **f**, Representative Raman spectra of *E. coli* cells that are not labeled (^{12}C) and *E. coli* cells labeled with ^{13}C and both ^{13}C and deuterium. The inset shows a magnification of the spectral region 900–1,100 cm^{-1}. **g**,**h**, P_{LC} (**g**) and P_L (**h**) values of *E. coli* that are not labeled (^{12}C) and that are labeled with ^{13}C and both ^{13}C and deuterium. Dashed lines represent the threshold values to identify the ^{13}C ($P_{LC} = 0.66$) and deuterium labeling status ($P_L = 6.5$). **i**, Representative Raman spectra of carotenoid-containing (*E. voratum*) and non-carotenoid-containing (*V. alginolyticus*) microorganisms. The inset shows values of $I_{1,500\text{-}1,560}$ (integrated intensity) of the two microorganisms and the dashed line ($I_{1,500\text{-}1,560} = 5{,}000$), indicating an example threshold value to distinguish between two populations (see main text). **j**, Classification of cells ($n = 70$) from a marine sediment sample into five clusters. The number of clusters was chosen based on iteration, by gradually increasing the number of clusters (inset) until the number of cells in the uppermost cluster (cluster 1; distinct from the others and containing cells that express the highest level of cytochrome *c*) no longer changed. The blue stars represent the centroids of the clusters. For the calculation, four spectral regions 745–755,

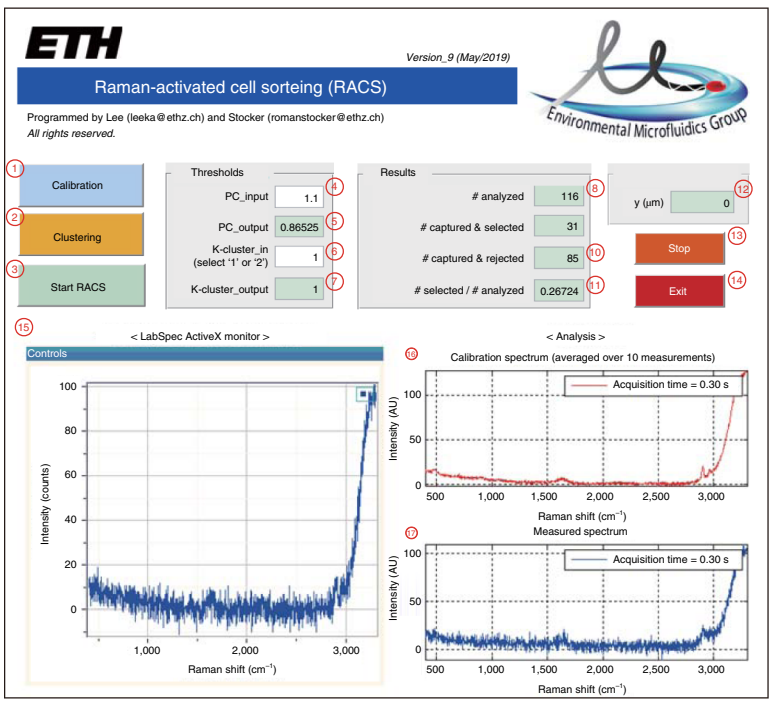
necessary control reaction performed in the absence of this compound (but in the presence of D_2O), the taxa directly or indirectly (via cross-feeding) stimulated in their activity by the compound addition can be identified by incorporation of D into their biomass (mostly within lipids). These deuterated target cells can be detected and then sorted by measuring the corresponding C–D peak (2,040–2,300 cm^{-1}) in their Raman spectra. The D_2O concentration in the medium must be chosen with consideration of a trade-off: the higher the D_2O concentration, the stronger the C–D signal but the higher the potential toxicity effect on the cells in the sample[17]. Thus, preliminary tests to optimize the D_2O concentration in the incubating medium are required (see Box 1). In lieu of deuterium, ^{13}C can be used to mark the cells of interest when tracking of autotrophs or other specific carbon metabolic pathways is of interest. The advantage of ^{13}C-labeling is that it tracks assimilation of specific carbon compounds (e.g., by a glucose utilizer), whereas D_2O labels all cells that are metabolically active under the incubation conditions and thus always requires the control experiment described above. ^{13}C-labeling has been traditionally used for a wide range of studies, although many chemical compounds are difficult to obtain in a ^{13}C-labeled derivative form. In this case, the microbial community is incubated in a minimal medium supplemented with a ^{13}C-labeled compound metabolized during the process of interest (e.g., ^{13}C-glucose or ^{13}C-dimethylsulfoniopropionate (DMSP)). The microbes that incorporate ^{13}C can be identified via Raman microspectroscopy by measuring the shift of their amino acid phenylalanine peak from 1,007 to 971 cm^{-1} as a result of the replacement of ^{12}C by ^{13}C (i.e., a red shift due to increased molecular mass[15,65]). To avoid cross-feeding, incubation experiments of microbial communities with D_2O or ^{13}C-labeled compounds should always be performed for short time periods. By contrast, if deciphering microbial food webs is of interest, longer incubation times will be necessary to ensure labeling of secondary and tertiary consumers. Ideally, the concentration of the added ^{13}C-labeled compound of interest should mimic its in situ concentration. Higher compound concentrations are sometimes necessary to achieve detectable Raman signals, but results obtained under such conditions have to be interpreted with caution.

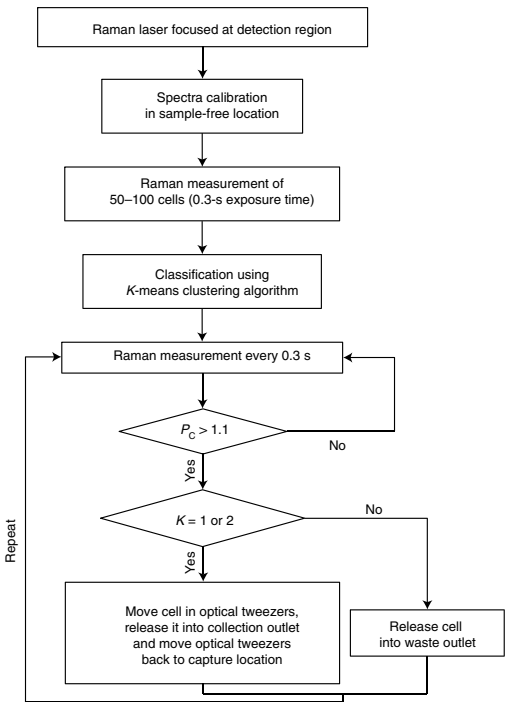
When the taxa of interest contain a specific molecule that generates a Raman signal of sufficient intensity, for example, because they have accumulated large amounts of storage compounds or contain molecules of interest that are easily detectable by virtue of the resonance effect, the sample can directly be used for RACS without prior labeling. For instance, cells expressing a high level of cytochrome *c* in marine sediment samples can be sorted by measuring the spectral regions of 750, 1,127, 1,314, and 1585 cm^{-1} (ref. [44]). Microbes containing carotenoids, a general (but not absolutely specific) indicator for photoautotrophic microorganisms such as phytoplankton (e.g., diatoms and the coral symbionts of the *Symbiodinium* genus) and cyanobacteria, can be distinguished from many co-existing heterotrophic bacteria by monitoring the characteristic Raman peak for carotenoids at 1,520 cm^{-1} (resulting from the stretching of the C=C double bond)[26].

Preparation of certain samples may require additional precautions. For example, gastrointestinal microbiota should be handled under anaerobic conditions (e.g., using an anaerobic tent). For marine sediment samples, enrichment procedures may be required to increase the initially low abundance of taxa of interest (e.g., cytochrome *c*–expressing cells) within the community. For example, culturing









the sample in minimal medium supplemented with nitrite (NO_2^-) will enable nitrite-oxidizing cells expressing high levels of cytochrome *c* to be active, whereas other species lack an energy source for growth. Nitrite consumption by the sample can be monitored using a nitrite indicator. Increasing the relative abundance of the taxa of interest will thereby speed up the sorting of target cells.

Sorting of target cells

Sorting of target cells from their communities is achieved by automated RACS in a microfluidic device, the major methodological advance in our pipeline (Fig. 1). Microfluidic devices, often

◀ **Fig. 8 | RACS software platform and operation algorithm. a**,**b**, MATLAB graphical user interface (GUI) and the workflow of the RACS. The numbered elements represent buttons to turn on (1) and off (2) the video of the Raman system (CCD camera 1 in Fig. 4); (3) the button to begin calibration, which instructs the stage to move to the sample-free region (70 μm away from the sample stream) and measure the Raman spectrum of the working fluid (averaged over 10 measurements); (4) the button to start the RACS process (see panel **b**); (5) the threshold set by the user to detect the capture of a cell by the optical tweezers (cell index, P_C); (6) the threshold set by the user to identify a cell as deuterium labeled (labeling index, P_L); displays of the P_C (7) and P_L (8) values in real time; displays of the numbers of cells analyzed (9), captured and selected (10), and captured and rejected (11); the proportion of selected cells among those analyzed (12); (13) the current position in the *y* direction of the microscope stage (perpendicular to the flow direction); buttons to stop the RACS process (14) and to close the software platform (15); displays of the Raman spectrum measured in real time (16), of the calibration spectrum (17) (used to calculate P_C), and of the latest captured and selected cell (18). **c**,**d**, GUI for the RACS platform operated using a machine-learning algorithm (*K*-means clustering) and the workflow for the operation. The numbered elements are: (1) the button to begin calibration, which instructs the stage to move to the sample-free region (70 μm away from the sample stream) and measure the Raman spectrum of the working fluid (averaged over 10 measurements); (2) the button to begin the initial determination of the clusters, which instructs the software to evaluate the Raman spectral characteristics of the sample (when a cell is captured in the optical tweezers, it is measured and immediately released without sorting). Upon clicking the 'Stop' button (13) when the number of cells analyzed (displayed in box 8) meets the user's requirements, the software classifies the measured data into several clusters using the *K*-means clustering algorithm (the number of cells to be analyzed during this initial stage depends on the relative abundance of taxa of interest within the sample); (3) button to start the RACS process (see panel **d**); (4) the threshold set by the user for P_C and (5) the P_C value measured in real time; (6) the value of the clusters from which cells are to be collected (1 and 2 mean the uppermost cluster and the uppermost and second uppermost clusters, respectively; e.g., cluster 1 alone and clusters 1 and 2 in Fig. 7j,k); (7) real-time output of the cluster value of the cell currently being analyzed ('N/A' indicates clusters other than those to be sorted, e.g., clusters 3–5 if the software is classifying the data into five clusters and clusters 1 and 2 are being collected; see Fig. 7j for the selection criterion for the number of clusters); displays of the number of cells analyzed (8), captured and selected (9), and captured and rejected (10); the proportion of selected cells among those analyzed (11); (12) the current position in the *y* direction of the microscope stage (perpendicular to the flow direction); buttons to stop the RACS process (13) and to close the software platform (14); displays of the Raman spectrum measured in real time (15), of the calibration spectrum (16) (used to calculate P_C), and of the latest captured and selected cell (17).

fabricated using soft-lithography[66], have been widely adopted in microbiology research laboratories and have matured to the stage where devices can be provided by commercial manufacturers. In this protocol, we provide designs for all necessary components to replicate our microfluidic device, regardless of the depth of the users' expertise in microfabrication. Users with access to a cleanroom facility can prepare the device from scratch (i.e., fabrication from a master mold using a photomask (provided in Supplementary Data 1) and creation of the PDMS microfluidic device), whereas those who do not have expertise in microfabrication can rely on commercial manufacturers to fabricate the master mold—e.g., Darwin Microfluidics (https://darwin-microfluidics.com), microfluidic ChipShop (https://www.microfluidic-chipshop.com/services/design-development-manufacturing) or uFluidix (https://www.ufluidix.com/services/#design_and_development)—and then fabricate the PDMS microfluidic device in their laboratory, or outsource both steps. In this protocol, we describe the steps involved in fabrication of the PDMS microfluidic device once a master mold has been obtained (Steps 2–18). The PDMS layer containing the fluidic channels is chemically bonded to a glass coverslip via corona (plasma) treatment. To enable visualization of the RACS process through the PDMS layer (while the Raman measurement is made through the glass coverslip; Figs. 4 and 6; see also below), the PDMS layer is thin (~1 mm) to minimize interference in the optical path for imaging. The use of a coverslip made of CaF_2 or quartz (instead of a regular glass coverslip) was considered on the basis of reports about the lower spectral background of these materials in the literature[67,68]; however, these specialized coverslips do not provide a substantial advantage for the detection of the peaks of interest in this protocol (C–D, phenylalanine, carotenoids, cytochrome *c*) because the Raman measurements occur 15 μm away from the coverslip: hence, the choice of coverslip material has limited effect on the Raman signal because the Raman microspectroscope is a 'confocal' system. By contrast, coverslips made of CaF_2 or quartz bring the disadvantage that they cannot be attached to PDMS via corona treatment.

The RACS process is fully automated using a platform built in-house (Fig. 8a,b and Supplementary Video 1), with which two indices are used: the 'cell index' (P_C: to detect the capture of a cell in the optical tweezers) and the 'labeling index' (P_L: to identify the deuterium-labeling status or other molecular signal of interest of the captured cell). The two indicators are computed from specific regions of a Raman spectrum as follows:

$$P_C = \frac{I_{1,620-1,670}}{I_{\text{Fluid, }1,620-1,670}}, \quad (1)$$

$$P_L = \frac{I_{2,040-2,300}}{I_{1,850-1,900}}, \quad (2)$$

where I represents the integrated intensity within the spectral region denoted by the subscripts (the subscripts are the wavenumbers bounding the spectral region considered). The two spectral regions of

1,620–1,670 cm^{-1} and 1,850–1,900 cm^{-1} were chosen because they are not susceptible to the signals coming from the microfluidic device (PDMS and glass); that is, the Raman measurement is reliable regardless of the small change (a few microns) in the vertical position of the laser spot within the microfluidic device, between the PDMS and glass surfaces. The Raman intensity in the spectral region 1,620–1,670 cm^{-1} increases upon capture of a cell (relative to the reference value of the carrier fluid, $I_{\text{fluid, 1,620–1,670}}$) and the spectral region 2,040–2,300 cm^{-1} is a strong fingerprint for the C–D bond. The P_L value has to be adapted if other molecular bonds are targeted (see below; Eq. (3)).

The sample is continuously flowed into the microfluidic device and focused on one side of the channel by sheath flow, so that the default situation is for cells to traverse the device and exit into the waste outlet (Fig. 2). Two lasers, for Raman measurement (532 nm) and optical tweezing (1,064 nm), are focused on the same point in the sample stream with a single 63×, 1.2-NA, water-immersion objective. Upon capture of a cell in the optical tweezers, the overall Raman spectrum intensity increases. To measure this change using P_C ($P_C > 1.1$ is the criterion operationally adopted to determine capture of a cell; 0.3-s exposure time; Fig. 7a–c), the platform is calibrated before each RACS. To this end, the laser spot is moved to a cell-free region (between the locations used for cell capture and release during sorting; calibration location in Fig. 2), and the Raman spectrum of the working fluid (in the absence of cells) is averaged over ten measurements to provide the denominator in Eq. (1). The labeling index P_L is then quasi-simultaneously calculated from Eq. (2) based on the C–D peak, and used to evaluate whether the cell is deuterium labeled according to a specified criterion ($P_L = 6.5$ and $P_L = 6.19$ in Fig. 7a,b and in Fig. 7d, respectively; see below for a guide to the selection of these threshold values). If labeled, the cell is translocated perpendicular to the flow direction to the release location (laser shutter 2 blocking the Raman laser to prevent photophoretic damage to the cell; Fig. 4a,f,g) and released by blocking laser shutter 1, so that the flow carries the cell into the collection outlet. The optical tweezers then move back to the capture location, the laser shutters open, and the process starts over to capture a new cell. If unlabeled, on the other hand, the cell is immediately released, and the flow stream in that location carries it to the waste outlet. This process is fully automated and is repeated to collect cells one by one with no human intervention.

The spectral window used for the labeling index P_L is tunable depending on the chemical signature of the cells of interest. For example, the ^{13}C labeling status of heterotrophic bacterial cells can be used as a sorting criterion, yielding as a natural choice:

$$P_{LC} = \frac{^{13}C}{^{13}C + {}^{12}C} = \frac{I_{965-977}}{I_{1,005-1,010} + I_{965-977}}, \quad (3)$$

where $I_{1,005–1,010}$ and $I_{965–977}$ are the integrated intensities that correspond to the phenylalanine peaks for ^{13}C and ^{12}C, respectively (Fig. 7f,g). The sum of ^{13}C and ^{12}C peak intensities is used in the denominator because the ^{12}C peak intensity decreases with increasing ^{13}C-labeling status of cells, and thus this sum can be an appropriate reference for the calculation. By contrast, the spectral region of 1,850–1,900 cm^{-1} is used to calculate the P_L value in Eq. (2) because the C–H peak (2,800–3,100 cm^{-1}) is strongly influenced by signals generated from PDMS (Fig. 5c); the sum of C–D and C–H peak intensities changes according to the small change in the vertical position of the laser spot within the microfluidic device. The use of dual sorting criteria (P_L and P_{LC}) is also possible, enabling one to probe a novel function of interest (via P_L), complemented by the tracking of carbon metabolism of individual cells (via P_{LC}) (Fig. 7f–h). Moreover, the raw spectral intensity can be used to sort microbes containing resonance-inducing compounds (e.g., carotenoids in photoautotrophic microorganisms) from other bacteria because their very strong signal is easily distinguished (Fig. 7i).

To provide even greater versatility, we have developed a new software platform for RACS based on a machine-learning algorithm, specifically '*K*-means clustering'[28,29], a simple and commonly used unsupervised algorithm (Supplementary Software 1 and 2) (Fig. 8c,d). Briefly, the software initially measures a certain number of cells (with the required number depending on the relative abundance of target cells in the sample, which can be assessed with the RACS platform as a 'counter'; see below) and classifies them into mutually exclusive clusters through iteration to minimize the Euclidean distance between the centroids of each cluster and the measured Raman data for the population of cells. To this end, the software initially selects the centroids of disjoint clusters (a random selection) and assigns the measured Raman spectra to the clusters of their closest centroids. Then it iteratively refines the locations of the clusters to minimize the mean Euclidean distance from each data point to its nearest centroid. This process ends when the centroids no longer move. During RACS, cells are assigned to clusters by calculating the Euclidean distances between the centroids identified beforehand and the Raman data measured in real time, and the cells belonging to the cluster or clusters of

interest are sorted. The choice of the number of clusters is key to the precision of the *K*-means clustering algorithm. Several strategies have been proposed, for example, elbow and silhouette methods, all of which are iterative approaches[69]. Our software uses a novel strategy, illustrated for the marine sediment sample containing cells that express high levels of cytochrome *c*: the software sets the number of clusters when the number of cells in the cluster containing cells that express the highest level of cytochrome *c* (the uppermost cluster, which is most distinct from all others, in our plots; e.g., cluster 1 in Fig. 7j,k) does not change with a further increase in the number of clusters (inset in Fig. 7j). This new software platform has great advantages for the sorting of samples in which the overall spectral background varies depending on the cells' ecophysiological condition (e.g., oxidation state of cytochrome, for the marine sediment sample[70]); the initial measurements to classify the cells into clusters provide a means for the RACS software to 'learn' the ecophysiological condition of the samples, thereby taking into account variation in the sample condition. Detection of cells expressing high levels of cytochrome *c* under different sample conditions can be achieved using Raman intensity values (e.g., as for P_L or P_{LC}) in conjunction with background subtraction. However, the background subtraction algorithms that are widely used to process Raman data are based on iteration[71,72] and thus would substantially increase calculation times of the sorting criteria. This would hinder applications such as RACS that require rapid determination of sorting criteria, despite being suitable for general processing of Raman data for analyses after multiple measurements. By contrast, our machine learning–based platform conducts the time-consuming clustering procedure before the RACS itself, so that during RACS it is necessary only to calculate the Euclidean distances between Raman data measured in real time and the centroids of each cluster to assign the data to the most appropriate cluster. As a result, this approach does not sacrifice sorting throughput, yet still properly takes into account variation in the sample conditions.

More generally speaking, a Raman spectrum is a multivariate signal, and although we use a small subset of this signal richness here, further development of machine learning would provide an ideal approach with which to harvest more of the complexity of the Raman spectrum.

Downstream analysis

Cells collected using RACS are viable and can be used for two broad categories of downstream analyses: cultivation and genomic analysis. In this protocol, we describe methods to cultivate the collected cells (Step 58). Cultivation of sorted cells enables further characterization of the taxa of interest, for example, to determine substrate usage patterns of cultivated microorganisms with a defined function in the ecosystem and investigate phenotypic traits such as resistance to antibiotics[40]. Techniques for downstream DNA analysis have been described in detail elsewhere[30], and many commercial research laboratories provide these services. Briefly, three approaches can be considered for investigating the sorted microbes (Fig. 1): (i) 16S rRNA gene amplicon sequencing (i.e., sequencing of PCR-amplified 16S rRNA genes) to compare the composition of microbial communities performing a particular function across different environments or time points; (ii) mini-metagenomics to identify and genomically characterize taxa performing a specific function within a community; and (iii) isolation of single cells on a multiwell plate (using FACS to perform the separation of the RACS-treated cells), followed by single-cell genomics to link the functional roles (phenotypes) and genotypes of individual microorganisms at very high resolution[30]. These techniques can be combined to complement each other, for example, as in our previous RACS-application[24], in which we used 16S rRNA gene amplicon sequencing to compare the composition of cells responsible for mucin degradation with that of the full community and mini-metagenomics to retrieve relevant genomic features of sorted bacteria in the mouse colon community.

Materials

Biological materials

- Complex natural community sample (e.g., colon contents of 6- to 8-week-old C57BL/6J mice bred at the Max F. Perutz Laboratories, University of Vienna, under specific pathogen-free conditions or marine sediments containing microbes expressing high levels of cytochrome *c*; see Step 1 for details about sample preparation) **! CAUTION** All animal experiments must be approved by the relevant institutional ethics committee and conducted in accordance with approved institutional regulations and guidelines. Our mouse colon specimens were collected under approval from the Institutional Ethics Committee of the University of Veterinary Medicine, Vienna, in accordance with Austrian laws (BMWF-66.006/0002-II/10b/2010).

- Laboratory model microorganisms (e.g., *Vibrio alginolyticus* 12G01 (wild type; NCBI:txid314288, kindly provided by M. Polz's group, University of Vienna, Austria)[73], *Vibrio cyclitrophicus* 1F111 (wild type; NCBI:txid1136159, kindly provided by M. Polz's group)[73], *Escherichia coli* NCM3722 Δ*motA* (non-motile mutant, kindly provided by S. Jun's group, University of California San Diego, USA)[74], and the free-living *Symbiodinium* species *Effrenium voratum* (National Center for Marine Algae and Microbiota (NCMA), cat. no. CCMP421)) **! CAUTION** All experiments using laboratory model microorganisms must be conducted in accordance with regulations and guidelines appropriate to their biosafety level.

Reagents

! CAUTION The Caution notes below provide summaries of potential dangers and precautions that should be taken. More detailed precautions for potentially harmful reagents are described in the material safety data sheets (MSDSs) that are provided on each manufacturer's website: Sigma-Aldrich (https://www.sigmaaldrich.com/); Instant Ocean (http://www.instantocean.com/); Carl Roth (https://www.carlroth.com/); and Alfa Aesar (https://www.alfa.com/). **▲ CRITICAL** None of the reagents used in this protocol are supplier sensitive; thus, any equivalent products available on the market can be used.

General

- EtOH (Sigma-Aldrich, cat. no. E7023) **! CAUTION** EtOH is a highly flammable liquid and vapor. Keep away from heat/sparks/open flames/hot surfaces. EtOH causes serious eye irritation. Wear protective gloves/eye protection/face protection. Wash skin thoroughly after handling.
- Milli-Q water (ultrapure water)
- Calcium chloride ($CaCl_2$; Sigma-Aldrich, cat. no. C5670) **! CAUTION** Calcium chloride causes serious eye irritation. Wear eye protection/face protection. Wash skin thoroughly after handling.
- Sodium chloride (NaCl; Carl Roth, cat. no. 3957)

Sample preparation (general)

- Deuterium oxide (D_2O; Sigma-Aldrich, cat. no. 151882)
- Glycerol (Sigma-Aldrich, cat. no. G7893)
- D-glucose (Sigma-Aldrich, cat. no. G8270)
- Sea salt (Instant Ocean, cat. no. SS15-10) **! CAUTION** Contact with Instant Ocean in dry form may cause skin or eye irritation. Wear protective gloves/eye protection/face protection.
- LB broth (BD Biosciences, cat. no. 244610)
- Agar for microbiology (Sigma-Aldrich, cat. no. 05039)

Sample preparation (mouse colon sample)

- Potassium chloride (KCl; Fluka, cat. no. 31248)
- Di-sodium hydrogen phosphate dihydrate ($Na_2HPO_4{\cdot}2H_2O$; Carl Roth, cat. no. 4984)
- Potassium phosphate, monobasic (KH_2PO_4; Sigma-Aldrich, cat. no. P5655)
- Mucin from porcine stomach (Sigma-Aldrich, cat. no. M2378)

Sample preparation (marine sediment sample)

- Biotin (Sigma-Aldrich, cat. no. B4639)
- Folic acid (Sigma-Aldrich, cat. no. F8758)
- Pyridoxine HCl (Sigma-Aldrich, cat. no. PHR1036) **! CAUTION** Toxic to aquatic life. Avoid release to the environment. Dispose of contents/container in an approved waste disposal plant.
- Thiamine HCl (Sigma-Aldrich, cat. no. V-014) **! CAUTION** Thiamine HCl is a highly flammable liquid and vapor. Keep away from heat/sparks/open flames/hot surfaces. It is toxic if swallowed, upon contact with skin or if inhaled. It causes damage to organs (eyes). Do not breathe dust/fumes/gas/mist/vapors/spray. Wash skin thoroughly after handling. Use only outdoors or in a well-ventilated area. Wear protective gloves/eye protection/face protection.
- Riboflavin (Sigma-Aldrich, cat. no. R9504)
- Nicotinic acid (Carl Roth, cat. no. 3815.1) **! CAUTION** Nicotinic acid causes serious eye irritation. Wear eye protection/face protection. Wash skin thoroughly after handling.
- DL-Pantothenic acid (Sigma-Aldrich, cat. no. 295787)
- 4-Aminobenzoic acid (Sigma-Aldrich, cat. no. A-9878) **! CAUTION** 4-Aminobenzoic acid causes skin irritation and may cause an allergic skin reaction. It causes serious eye irritation and may cause

respiratory irritation. Avoid breathing dust/fumes/gas/mist/vapors/spray. Wash skin thoroughly after handling. Use only outdoors or in a well-ventilated area. Wear protective gloves/eye protection/face protection.
- Choline chloride (Sigma-Aldrich, cat. no. C7527)
- Vitamin B_{12} (Fluka, cat. no. 95190)
- $NaNO_2$ (Sigma-Aldrich, cat. no. S2252)

Sample preparation (laboratory model microorganisms)
- LB broth (BD Biosciences, cat. no. 244610)
- Marine broth 2216 (BD Biosciences, cat. no. 279110)
- f/2 Media Kit (NCMA at Bigelow Laboratory, cat. no. MKF250L)
- M9 minimal salts (5×, Sigma-Aldrich, cat. no. M9956)
- Magnesium sulfate ($MgSO_4$; Sigma-Aldrich, cat. no. M2643)
- D-glucose–$^{13}C_6$ (Sigma-Aldrich, cat. no. 389374)

Fabrication of microfluidic device
- SYLGARD 184 silicone elastomer kit (base elastomer and curing agent; Dow, cat. no. 2065622)

Sorting of target cells
- DNA AWAY Surface Decontaminant (Thermo Fisher Scientific, cat. no. 7010)
- Tween 20 (Sigma-Aldrich, cat. no. P9416)
- Immersion oil (Immersol W 2010; Carl Zeiss, cat. no. 444969-0000-000) ▲ **CRITICAL** This has the same refractive index as water (RI = 1.33) but does not evaporate over time.

Collection of sorted cells and their downstream analysis (mouse colon sample)
- Bacto casitone (Gibco, cat. no. 225930)
- Yeast extract (Oxoid, cat. no. LP0021)
- Magnesium sulfate heptahydrate ($MgSO_4 \cdot 7H_2O$; Sigma-Aldrich, cat. no. 63138)
- Potassium phosphate, dibasic (K_2HPO_4; Sigma-Aldrich, cat. no. P8281)
- Potassium phosphate, monobasic (KH_2PO_4; Sigma-Aldrich, cat. no. P5379)
- Resazurin sodium salt (Sigma-Aldrich, cat. no. R7017)
- Sodium bicarbonate ($NaHCO_3$; Sigma-Aldrich, cat. no. S5761)
- L-Cysteine hydrochloride (Carl Roth, cat. no. 3468.2) ! **CAUTION** L-Cysteine hydrochloride causes skin irritation and serious eye irritation. It may cause respiratory irritation. Wear protective gloves/eye protection. Use only outdoors or in a well-ventilated area.
- Hemin (Fluka, cat. no. 51280)
- Acetic acid (Alfa Aesar, cat. no. 36289) ! **CAUTION** Acetic acid is a flammable liquid and vapor. It causes severe skin burns and eye damage. Use explosion-proof electrical/ventilating/lighting/equipment. Wash face, hands, and any exposed skin thoroughly after handling. Use only outdoors or in a well-ventilated area. Wear protective gloves/protective clothing/eye protection/face protection.
- Propionic acid (Sigma-Aldrich, cat. no. P1386) ! **CAUTION** Propionic acid is a flammable liquid and vapor. Keep away from heat/sparks/open flames/hot surfaces. It causes severe skin burns and eye damage and may cause respiratory irritation. Avoid breathing dust/fumes/gas/mist/vapors/spray. Wash skin thoroughly after handling. Use only outdoors or in a well-ventilated area. Wear protective gloves/protective clothing/eye protection/face protection.
- Isobutyric acid (Sigma-Aldrich, cat. no. I1754) ! **CAUTION** Isobutyric acid is a flammable liquid and vapor and is harmful if swallowed. It is toxic upon contact with skin and causes severe skin burns and eye damage. It is harmful to aquatic life. Keep away from heat/sparks/open flames/hot surfaces. Wash skin thoroughly after handling. Wear protective gloves/protective clothing/eye protection/face protection.
- *n*-Valeric acid (Sigma-Aldrich, cat. no. V0125) ! **CAUTION** *n*-Valeric acid causes severe skin burns and eye damage and is harmful to aquatic life, with long-lasting effects. Wash skin thoroughly after handling. Wear protective gloves/protective clothing/eye protection/face protection.
- Isovaleric acid (Sigma-Aldrich, cat. no. 129542) ! **CAUTION** Isovaleric acid is a combustible liquid. It causes severe skin burns and eye damage. Keep away from heat/sparks/open flames/hot surfaces. Wash skin thoroughly after handling. Wear protective gloves/protective clothing/eye protection/face protection.
- Biotin (Carl Roth, cat. no. 3822.2)

- Folic acid (Sigma-Aldrich, cat. no. F7876)
- Pyridoxine hydrochloride (Sigma-Aldrich, cat. no. PHR1036) **! CAUTION** Pyridoxine hydrochloride is harmful to aquatic life. Avoid release into the environment. Dispose of contents/container to an approved waste disposal plant.
- Thiamine hydrochloride (Sigma-Aldrich, cat. no. T1270)
- Riboflavin (Sigma-Aldrich, cat. no. R4500)
- Nicotinic acid (Carl Roth, cat. no. 3815.1) **! CAUTION** Nicotinic acid causes serious eye irritation. Wash skin thoroughly after handling. Wear protective gloves/protective clothing/eye protection/face protection.
- Calcium-D-pantothenate (Carl Roth, cat. no. 3812.2)
- Vitamin B_{12} (Fluka, cat. no. 95190)
- 4-Aminobenzoic acid (Sigma-Aldrich, cat. no. A-9878)
- Lipoic acid (Sigma-Aldrich, cat. no. T5625) **! CAUTION** Lipoic acid is harmful if swallowed. Wash skin thoroughly after handling.
- Hydrochloric acid solution (HCl; Sigma-Aldrich, cat. no. H9892) **! CAUTION** Hydrochloric acid solution causes severe skin burns and eye damage. Wash skin thoroughly after handling. Wear protective gloves/protective clothing/eye protection/face protection. Hydrochloric acid solution may be corrosive to metals. Keep only in original container.
- Sodium hydroxide solution (NaOH; Sigma-Aldrich, cat. no. S2770) **! CAUTION** Sodium hydroxide solution causes severe skin burns and eye damage. Wash skin thoroughly after handling. Wear protective gloves/protective clothing/eye protection/face protection. Sodium hydroxide solution may be corrosive to metals. Keep only in original container.

Cleaning after sorting

- Acetone (Sigma-Aldrich, cat. no. 179124) **! CAUTION** Acetone is a highly flammable liquid and vapor. It causes serious eye irritation. Keep away from heat/sparks/open flames/hot surfaces. Avoid breathing dust/fumes/gas/mist/vapors/spray. Wash skin thoroughly after handling. Use only outdoors or in a well-ventilated area. Wear protective gloves/eye protection/face protection.

Equipment

General

- Microcentrifuge (Eppendorf, model no. 5417R)
- 1.5-mL microcentrifuge tube (Eppendorf, cat. no. 0030120086)
- 0.2-mL PCR tube (Eppendorf, cat. no. 0030124332)
- 50-mL conical-bottom tube (Greiner Bio-One, cat. no. 227261)
- Polyethersulfone (PES) syringe filter (0.2-μm pore size, 13-mm diameter; Sigma-Aldrich, cat. no. Z741696)
- Petri dish (Greiner Bio-One, cat. no. 639160)
- Powder-free nitrile gloves (Sempermed, cat. no. 826781637)
- Blunt-end tweezers (Merck Millipore, cat. no. XX6200006P)
- Adhesive tape (Scotch Magic tape)

Sample preparation

- Shaking incubator (New Brunswick, model. no. Innova 4000)
- Anaerobic chamber (Coy Laboratory Products, Vinyl Type A)
- 2.0-mL glass vial (Wheaton, cat. no. 223683)
- Vial rubber stopper (Sigma-Aldrich, cat. no. 27224)
- Vial crimp seal (Sigma-Aldrich, cat. no. 27222-U)
- Steritop threaded bottle-top filter (0.2-μm pore size; EMD Millipore, cat. no. S2GPT05RE)
- Optical density meter (Implen, model no. OD600 DiluPhotometer)
- Disposable cuvette (Implen, model no. DiluCell)
- Nitrite test paper (CTL Scientific Supply, cat. no. 91311)

Fabrication of microfluidic device

- Master mold fabricated on a 4-inch silicon wafer **▲ CRITICAL** A user who has expertise in soft-lithography[66] and has access to a cleanroom facility can fabricate the master mold using a design file for the photomask (Supplementary Data 1) and standard microfabrication procedures (see ref. [75]).

Table 2 | Specifications of three commercial Raman systems (as of May 2020)

Manufacturer	Origin	Capability to integrate optical tweezers (1,064-nm laser)	Grating that can cover the 400–3,300 cm^{-1} spectral region with a single Raman measurement and is optimized WRT measurement with 532-nm laser	Control of system by third party software
Horiba Scientific	France	Yes	Yes (300 lines/mm)	Yes (via ActiveX)
Renishaw	UK	Yes	Yes (300 lines/mm)	Yes (via TCPIP interface)
Bruker	USA	Yes	Yes (400 lines/mm)	Yes (via HTTP protocol)

WRT, with regard to.

Alternatively, a master mold can be obtained by providing the design file to a commercial manufacturer, e.g., Darwin Microfluidics (https://darwin-microfluidics.com), microfluidic ChipShop (https://www.microfluidic-chipshop.com/services/design-development-manufacturing) or uFluidix (https://www.ufluidix.com/services/#design_and_development).

- Convection oven (capable of 75 °C; Quincy Lab, model no. 20GCE)
- Handheld corona treater (Electro-Technic Products, model no. 12051A)
- Balance (capable of 0.1-g resolution; A&D Weighing, model No. EK-1200i)
- Vacuum desiccator (Sigma-Aldrich, cat. no. Z119024)
- Vacuum pump (Gast Manufacturing, model no. DOA-P701-AC)
- Glass coverslip (Paul Marienfeld, cat. no. 0107242)
- Biopsy punch (0.75-mm diameter; World Precision Instruments, cat. no. 504529)
- Scalpel blade (Sigma-Aldrich, cat. no. S2646)
- Scalpel handle (Sigma-Aldrich, cat. no. S3021)
- Plastic cup
- Plastic knife
- Rubber dust air blower
- Air duster

Sorting of target cells

- Syringe pump (3 units; Cetoni, model no. neMESYS 290N) ▲ **CRITICAL** The equipment operates with a DC motor, whereas many others use a stepping motor. Unlike the stepping motor, a DC motor does not show discretized motion, and thus it provides stable fluidic conditions, especially when operating at the low flow rates (down to 0.06 μL/min) required by the RACS platform. Other syringe pumps that operate with a DC motor can be used as an alternative.
- Acrylic plate (McMaster-Carr, cat. no. 8560K358)
- Laser engraving and cutting machine (Universal Laser Systems, model no. PLS 6MW)
- Polyethylene (PE) tubing (sterile; Warner Instruments, cat. no. 64-0752)
- Microfluidic steel tubing (23G; Darwin Microfluidics, cat. no. LVF-3490)
- Blunt needle (23 gauge; Luer Lock; Transcodent, cat. no. 6012)
- 500-μL gastight glass syringe (Hamilton, cat. no. 1750 TLL)
- 1-mL gastight glass syringe (Hamilton, cat. no. 1001 TLL)
- Lens tissue paper (Glaswarenfabrik Karl Hecht, cat. no. 41019010)
- Scissors
- Pliers (two pair)
- Specimen clips

Confocal Raman microspectroscope

▲ **CRITICAL** The first seven components (from the confocal Raman microspectroscope to the grating) are available from the three Raman manufacturers (Horiba Scientific, Renishaw, Bruker) and the next four (from the objective to the laboratory retort stand) are available through a range of manufacturers. ▲ **CRITICAL** Three important technical modules of commercial Raman systems that are required to conduct this protocol are listed in Table 2.

- Confocal Raman microspectroscope (Horiba, model no. LabRAM HR800), based on upright microscope (Olympus, model no. BX-41)
- 532-nm neodymium-doped yttrium aluminum garnet (Nd:YAG) laser at 500 mW (Laser Quantum, model no. Ventus 532)
- 1,064-nm Nd:YAG laser at 500 mW (Laser Quantum, model no. Ventus 1064)
- Electron-multiplying charge-coupled device (EMCCD) spectrometer (Andor, model no. Newton DU970P-BV-328)
- Silicon wafer (for instrument calibration; VWR, cat. no. MSPP-452)
- Motorized *xyz*-stage (see the next subsection)
- 300-lines/mm grating that is optimized with respect to wavelengths between 500 and 600 nm (Horiba Scientific, part no. 51019140HR)
- High-power objective (Zeiss, cat. no. C-Apochromat 63×/1.20 W Korr UV-VIS-IR M27 63×; set 0.17 mm for coverslip correction collar)
- Two laser shutters (Thorlabs, cat. no. SH1)
- Two shutter controllers (Thorlabs, cat. no. SC10)
- Two laboratory retort stands (Allfine Medlab, cat. no. WIN-DL01)

Motorized *xyz* stage

- Motorized *xy* stage (Märzhäuser Wetzlar, cat. no. 00-26-401-0000)
- Motorized *z* stage (Märzhäuser Wetzlar, cat. no. 27-54-406-0000)
- 3-axis joystick (Märzhäuser Wetzlar, cat. no. 00-76-300-0820)
- Controller (Märzhäuser Wetzlar, cat. no. 00-76-150-1803)

Lower microscope

- Charge-coupled device (CCD) camera (PCO-TECH, model no. pco.1600)
- 6.5× magnification optical array (see the next subsection)
- Thread adaptor (external M26×36TPI thread and internal SM1 thread; Thorlabs, cat. no. SM1A28)
- Thread coupler (SM1 coupler; Thorlabs, cat. no. SM1T2)
- Beam-turning cubes (Thorlabs, cat. no. CM1-E02)
- Thread adaptor (external SM1 thread and internal RMS thread; Thorlabs, cat. no. SM1A3)
- 10× objective (Olympus, cat. no. MPLN10X)
- Short-pass filter (cutoff wavelength 500 nm; Thorlabs, cat. no. FESH0500)
- *Xyz* stage (Thorlabs, cat. no. LT3)
- Low-angle ring illuminator (Moritex, cat. no. MLRL-CB25) ▲ **CRITICAL** The low-angle ring illuminator creates quasi dark-field images.
- Ring illuminator controller (Moritex, cat. no. MLEK-A080W1LR-240V)
- Ring adaptor (McMaster-Carr, cat. no. 57485K77)
- Anodized aluminum plate (McMaster-Carr, cat. no. 7083T451)

6.5× magnification optical array

- C-mount coupler (Navitar, cat. no. 1-6010)
- 1.0× adaptor (Navitar, cat. no. 1-61400)
- Objective coupling lens (Navitar, cat. no. 1-60349D)
- Adaptor coupler (1.093-32TPI to M26×36TPI; Navitar, cat. no. 1-51896)

Collection of sorted cells and their downstream analysis

- Inoculation spreader (Sarstedt Inc., cat. no. 86.1569.005)

Cleaning after sorting

- Wash bottle (Sigma-Aldrich, cat. no. Z423149)
- Laboratory tissues (VWR, cat. no. 115-2074)

Software

- LabSpec 6 (https://www.horiba.com/en_en/products/detail/action/show/Product/labspec-6-spectroscopy-suite-software-1843/) to control the Raman microspectroscope (provided by the manufacturer, Horiba Scientific) ▲ **CRITICAL** Systems from other manufacturers can be controlled by their own software.
- neMESYS v.20191025 (https://www.cetoni.com/service/software/) to control the syringe pump (provided by the manufacturer, Cetoni)

- Camware v.4.04 (https://www.pco.de/company/whats-new/news/article/new-camware-version-404/) to control the lower microscope (provided by the manufacturer, PCO-TECH)
- RACS operation software platform written in MATLAB (https://www.mathworks.com/products/matlab.html) optimized for operation with the Horiba Raman microspectroscope. We have included descriptions throughout the code script, as well as a 'Horiba-specific' footnote for commands that are specific to the Horiba system (Supplementary Software 1 and 2). For Raman systems from other manufacturers, these commands can be replaced in accordance with the programming language appropriate for their system (Table 2)

Reagent setup

▲ **CRITICAL** D_2O at high concentration in the medium is known to have adverse effects on cell growth[17]. The effect is dependent on the sample type, and thus the experimental concentration used should be chosen on the basis of tests in which sample growth is compared at different D_2O concentrations (see Box 1).

Isotonic liquid

The working fluid, in which the cells are suspended for RACS, should be of comparable osmolarity to that of the medium used for sample incubation. In general, glycerol at a certain molar concentration (balanced by Milli-Q water; see below) and ASW are used for samples that do not and do require salinity for viability, respectively (Fig. 3b).

- *Glycerol*. 0.3 M or 0.2 M solution, autoclaved (for the mouse colon sample and the pure intestinal bacterial culture (*E. coli*), respectively)
- *ASW*. ASW is 36 g of sea salt dissolved in Milli-Q water (3.6% (wt/vol)), autoclaved, and filtered with a Steritop threaded bottle-top filter (for the pure marine bacterial cultures *V. alginolyticus* and *V. cyclitrophicus*).

Freshly prepared isotonic liquid can be stored for several months at room temperature (20–25 °C). It should be regularly autoclaved to prevent the formation of air bubbles over time, which would result in flow fluctuation within the microfluidic device and thus perturb the RACS. Autoclaving the liquid every 2 weeks is recommended. ▲ **CRITICAL** The isotonic working fluid is necessary to minimize the osmotic stress on cells when they are resuspended in a fluorescence-free, transparent liquid for the RACS. Use of a non-isotonic liquid (e.g., Milli-Q water) substantially reduces the laser power that can be used for RACS without causing photophoretic damage to cells. ▲ **CRITICAL** For the collection syringe, Tween 20 (0.2–0.5% (vol/vol)) can be added to isotonic liquid for the easy recovery of the sorted cells (see Steps 31 and 58; Fig. 3c).

Medium for mouse colon sample

Prepare and autoclave 2× PBS (16 g/L NaCl, 0.4 g/L KCl, 0.48 g/L KH_2PO_4, and 3.6 g/L Na_2H-$PO_4{\cdot}2H_2O$, dissolved in Milli-Q water) for the resuspension and brief incubation of the test and control samples. Freshly prepared medium can be stored for several months at room temperature.

YCFA medium

Prepare and boil together for 10 min 990 mL Milli-Q water, 10 g casitone, 2.5 g yeast extract, 5 g D-glucose, 45 mg $MgSO_4{\cdot}7H_2O$, 90 mg $CaCl_2$, 0.45 g K_2HPO_4, 0.45 g KH_2PO_4, 0.9 g NaCl, and 1 mg resazurin sodium salt. Cool to room temperature under anaerobic conditions and add 4 g $NaHCO_3$, 1 g L-cysteine hydrochloride, 10 mg hemin, 1.9 mL acetic acid, 0.7 mL propionic acid, 90 µL isobutyric acid, 100 µL *n*-valeric acid, and 100 µL isovaleric acid. Adjust the pH to 6.8 using 0.2 M NaOH or 0.2 M HCl and autoclave. Cool to room temperature and add 10 mL vitamin solution (2 mg/L biotin, 2 mg/L folic acid, 10 mg/L pyridoxine hydrochloride, 5 mg/L thiamine hydrochloride, 5 mg/L riboflavin, 5 mg/L nicotinic acid, 5 mg/L calcium-D-pantothenate, 0.1 mg/L vitamin B_{12}, 5 mg/L 4-aminobenzoic acid, and 5 mg/L lipoic acid; filter-sterilize with a Steritop threaded bottle-top filter). Freshly prepared medium can be stored for several months at 4 °C.

YCFA agar plate

Prepare and boil for 10 min 990 mL Milli-Q water, 10 g casitone, 2.5 g yeast extract, 5 g D-glucose, 45 mg $MgSO_4{\cdot}7H_2O$, 90 mg $CaCl_2$, 0.45 g K_2HPO_4, 0.45 g KH_2PO_4, 0.9 g NaCl, and 1 mg resazurin sodium salt. Cool to room temperature under anaerobic conditions and add 4 g $NaHCO_3$, 1 g L-cysteine hydrochloride, 10 mg hemin, 1.9 mL acetic acid, 0.7 mL propionic acid, 90 µL

isobutyric acid, 100 μL *n*-valeric acid, and 100 μL isovaleric acid. Adjust the pH to 6.8 using 0.2 M NaOH or 0.2 M HCl, add 15 g/L agar, and autoclave. Cool to 50 °C and add 10 mL vitamin solution (2 mg/L biotin, 2 mg/L folic acid, 10 mg/L pyridoxine hydrochloride, 5 mg/L thiamine hydrochloride, 5 mg/L riboflavin, 5 mg/L nicotinic acid, 5 mg/L calcium-D-pantothenate, 0.1 mg/L vitamin B_{12}, 5 mg/L 4-aminobenzoic acid, and 5 mg/L lipoic acid; filter-sterilize with a Steritop threaded bottle-top filter). Pour on Petri dishes (filling 60% of the volume of each dish) and leave to solidify and dry; then transfer to an anaerobic tent for storage. Freshly prepared agar plates can be stored for several months at 4 °C.

LB medium

Prepare and autoclave 50% (vol/vol) D_2O-containing (balanced with Milli-Q water) and non-D_2O-containing LB medium (25 g/L LB broth dissolved in D_2O and/or Milli-Q water) in which to incubate *E. coli* overnight. Freshly prepared medium can be stored for several months at room temperature.

2216 media

Prepare and autoclave 50% (vol/vol) D_2O-containing (balanced with Milli-Q water) and non-D_2O-containing 2216 media (37.4 g/L marine broth 2216 dissolved in D_2O and/or Milli-Q water) in which to incubate *V. alginolyticus* or *V. cyclitrophicus* overnight. Freshly prepared medium can be stored for several months at room temperature.

M9 media

Prepare and autoclave 1 M $MgSO_4$ and 1 M $CaCl_2$. Prepare 1 M D-glucose and 1 M D-glucose–$^{13}C_6$ and filter with a PES syringe filter (0.2-μm pore size) to sterilize. Prepare three types of M9 media:

- 780 mL/L Milli-Q water, 200 mL/L M9 minimal salts (5×), 2 mL/L 1 M $MgSO_4$, 0.1 mL/L 1 M $CaCl_2$, and 20 mL/L 1 M D-glucose (for the control sample; ^{12}C in Fig. 7f–h).
- 780 mL/L Milli-Q water, 200 mL/L M9 minimal salts (5×), 2 mL/L 1 M $MgSO_4$, 0.1 mL/L 1 M $CaCl_2$, and 20 mL/L 1 M D-glucose–$^{13}C_6$ (for the ^{13}C-labeled sample; ^{13}C in Fig. 7f–h).
- 280 mL/L Milli-Q water, 500 mL/L D_2O, 200 mL/L M9 minimal salts (5×), 2 mL/L 1 M $MgSO_4$, 0.1 mL/L 1 M $CaCl_2$, and 20 mL/L 1 M D-glucose–$^{13}C_6$ (for the ^{13}C and D-labeled sample; ^{13}C+D in Fig. 7f–h).

Freshly prepared media can be stored for several months at room temperature.

f/2 medium

Prepare ASW (see above) supplemented with f/2 Media Kit components (1 mL/L NaH_2PO_4, 1 mL/L $NaNO_3$, 1 mL/L trace metals, 1 mL/L SiO_3, and 0.5 mL/L vitamins) and filter with a Steritop threaded bottle-top filter to sterilize. Freshly prepared medium can be stored for several months at 4 °C.

Vitamin solution medium

Prepare vitamin solution medium: 0.02 g/L biotin, 0.02 g/L folic acid, 0.10 g/L pyridoxine HCl, 0.05 g/L thiamine HCl, 0.05 g/L riboflavin, 0.05 g/L nicotinic acid, 0.05 g/L DL-pantothenic acid, 0.05 g/L P-aminobenzoic acid, 2.00 g/L choline chloride, and 0.01 g/L vitamin B_{12}; filter-sterilize with a Steritop threaded bottle-top filter. Freshly prepared medium can be stored for several months at 4 °C.

Medium for marine sediment sample

Prepare ASW (see above, but use 33.4 g sea salt instead of 36 g) and supplement it with 4.2 μL/L vitamin solution medium. Filter-sterilize with a Steritop threaded bottle-top filter Freshly prepared medium can be stored for several months at 4 °C.

Equipment setup

The acrylic supporter

The supporting holder (dimensions are in Fig. 6a) can be sculpted with a laser engraving and cutting machine, which enables the microfluidic device to be fitted onto the microscope stage. If an existing specimen holder for the microscope stage accommodates the microfluidic device, this item is not necessary.

Microfluidic steel tubing

Autoclave to sterilize before use.

Microcentrifuge tube, PCR tube, glass vial, vial rubber stopper, and vial crimp seal
Autoclave and UV-irradiate 1.5-mL microcentrifuge tubes, 0.2-mL PCR tubes, 2.0-mL glass vials, vial rubber stoppers, and vial crimp seals to sterilize before use.

Lower microscope
Assemble eight components: CCD camera, optical array (6.5× magnification), short-pass filter, thread adaptor, thread coupler, beam-turning cubes, thread adaptor, and 10× objective (Fig. 4a,c). Fabricate a supporting aluminum plate with a milling machine (dimensions are in Fig. 4e) and, using four screws, attach it to the *xyz* stage that is fixed to the optical table (the *xyz* stage has a screw thread; the orientation is as shown in Fig. 4c,e). Install the microscope assembly on the other side of the aluminum plate, using screws (CCD camera and beam-turning cubes have a screw thread). Detach the *xyz* stage from the optical table, install the assembly under the microscope stage (Fig. 4d), and adjust the field of view (in the *xy* plane) to match that of the objective in the upright microscope (Fig. 4a).

Laser shutter
Mount the laser shutters on each laboratory retort stand and place them within the laser path: shutter 1 blocks both lasers and shutter 2 blocks the 532-nm laser alone (Fig. 4a,f,g).

Procedure

▲ **CRITICAL** All steps should be performed under sterile conditions using equipment and materials that have been autoclaved, cleaned using 70% (vol/vol) EtOH, or flame treated wherever possible. Wear gloves and wash with 70% (vol/vol) EtOH before handling materials.

Sample preparation

1 Prepare the samples of interest. For deuterium (D)-incorporation into the mouse colon sample, follow option A. For D-incorporation into pure bacterial cultures, follow option B. For ^{13}C or ^{13}C+D incorporation into the pure bacterial culture, follow option C. For the enrichment procedure for a marine sediment sample containing microbes expressing high levels of cytochrome *c*, follow option D.

(A) **Deuterium (D)-incorporation into the mouse colon sample** ● **Timing 8 h**

(i) Remove the colon contents of a 6- to 8-week-old mouse in an anaerobic tent.

(ii) Suspend the contents in 4 mL 2× PBS and homogenize by vortexing.

(iii) Wait 10 min to let the large particles settle, then transfer half of the particle-free supernatant into an equal volume of D_2O and the other half into Milli-Q water so that the resulting samples are the colon contents in 50% (vol/vol) D_2O-containing PBS (test sample) and non-D_2O-containing PBS (control), respectively.

(iv) Supplement the test sample with 2.5 mg/mL mucin from porcine stomach and the control sample with 2.5 mg/mL glucose.

▲ **CRITICAL STEP** Mucin can alternatively be used for the control. Glucose is suggested as a reference sugar source for the control in studies in which there are several test samples, each of which is supplemented with a different chemical stimulus of interest (e.g., galactose, fucose), providing a means to set a universal threshold value for RACS of the test samples (see Step 52B).

(v) Incubate cells for 6 h at 37 °C under anaerobic conditions, add glycerol (20% (vol/vol)) and then seal the vials with vial crimp seals for storage at −80 °C until further processing.

■ **PAUSE POINT** The mouse colon samples can be stored at −80 °C for up to 2 years before proceeding to the next step.

(B) **D-incorporation into pure bacterial cultures** ● **Timing 24 h**

(i) Inoculate a single bacterial colony picked from a pre-grown culture on an agar plate into either 50% (vol/vol)-D_2O-containing or non-D_2O-containing medium (5-mL volume in a 50-mL conical-bottom tube; the residual 45-mL allows gas transfer).

▲ **CRITICAL STEP** Use 2216 medium for *V. alginolyticus* or *V. cyclitrophicus*; use LB medium for *E. coli*.

(ii) Incubate overnight to prepare D-labeled or unlabeled bacterial species.

■ **PAUSE POINT** The pure bacterial cultures can be stored at 4 °C for up to 1 month before proceeding to the next step.

(C) **^{13}C or ^{13}C+D incorporation into pure bacterial cultures** ● **Timing 24 h**

(i) Inoculate a single *E. coli* colony picked from a pre-grown culture on an agar plate into each of the three M9 media (Reagent setup; use a 5-mL volume in a 50-mL conical bottom tube).

(ii) Incubate overnight to prepare the unlabeled control (i.e., ^{12}C-labeled), ^{13}C-glucose-labeled, and ^{13}C-glucose+D-labeled *E. coli*.

PAUSE POINT The pure bacterial cultures can be stored at 4 °C for up to 1 month before proceeding to the next step.

(D) **Enrichment procedure for marine sediment samples containing microbes expressing high levels of cytochrome *c*** ● **Timing 8 weeks**

(i) Maintain marine sediment samples in the dark and at room temperature until further processing in the laboratory.

(ii) Transfer a small spatula of sediment (0.5 g) to the vitamin-supplemented ASW medium (50 mL) supplemented with 1 mM $NaNO_2$ (to stimulate nitrite-oxidizing, cytochrome *c*-containing microbes) and incubate it in the dark at 28 °C without agitation.

(iii) Monitor nitrite concentration at weekly intervals and, when depleted, transfer 10% volume of culture to fresh medium with 1 mM $NaNO_2$.

(iv) Perform four sequential transfers before RACS to increase the initially low abundance of the taxa of interest (cytochrome *c*–expressing cells) within the community.

PAUSE POINT The marine sediment sample can be stored at room temperature for up to 1 week before proceeding to the next step.

Fabrication of microfluidic device ● Timing 3 h

2 Fix the master mold in the center of a Petri dish with adhesive tape (Fig. 9a,b).

3 Clean the surface of the master mold with an air duster and close the Petri dish to avoid contamination.

4 Pour the PDMS base elastomer and curing agent at a ratio of 10:1 (wt. %) into a plastic cup so that the resulting weight is 90 g, and thoroughly mix with a plastic knife.

CRITICAL STEP The preparation will appear whitish after mixing because of the formation of many tiny air bubbles.

5 Pour ~20 g of the PDMS mixture (a circle of 5-cm diameter) onto the master mold from Step 3, cover the Petri dish with a lid, and spread the PDMS mixture over the entire Petri dish by manually rocking it with slow movements to form an ~1-mm-thick layer (PDMS layer 1).

CRITICAL STEP It is essential that PDMS layer 1 be thin in order to obtain high-quality images through the PDMS layer in Step 43.

CRITICAL STEP Air bubbles in PDMS layer 1 can be removed by gently blowing air onto the surface with a rubber dust air blower (without vacuum processing).

6 Pour the rest of the PDMS mixture (~70 g) into an empty Petri dish to form an ~ 5-mm-thick layer (PDMS layer 2).

7 Place PDMS layer 2 in a desiccator (with the Petri dish lid open) under vacuum for 5 min so that the air bubbles disappear or accumulate near the surface. De-vacuum the desiccator and then open it. Burst remaining bubbles by gently blowing air onto the surface with a rubber dust air blower. Repeat this step three times until there are no visible air bubbles.

8 Bake the two PDMS layers in a convection oven at 75 °C for >30 min to allow them to polymerize (solidify).

PAUSE POINT The PDMS layers can be kept in the convection oven at 75 °C for a few days or kept at room temperature for several months.

9 For PDMS layer 1, cut around each microfluidic device (along the red dotted lines in Fig. 9a) using a scalpel and gently peel them from the master mold with blunt-end tweezers.

CRITICAL STEP Maintain an angle of <30° between the blade and the PDMS layer to avoid damage to the master mold during cutting; a small crack easily leads to breakage of the entire master mold.

CRITICAL STEP While removing the PDMS layer from the master mold, gently peel it from the left as far as the 'analysis region', and then from the right as far as the analysis region (Fig. 9a). If not removed in this way, the thin PDMS layer can tear within the analysis region, making the device unusable.

10 For PDMS layer 2, use a scalpel to cut two identically sized PDMS bricks (5 × 10 × 5 mm^3; width × length × thickness) for each microfluidic device (Figs. 6a, 9c,d, 10c).

11 Clean the surface of the microfluidic device layer from Step 9 (PDMS layer 1; the side without channels) and one side of the two bricks from Step 10 (PDMS layer 2) where they will be attached together (in Step 13) by attaching and then detaching adhesive tape. Repeat this step three times with new, clean adhesive tape each time.

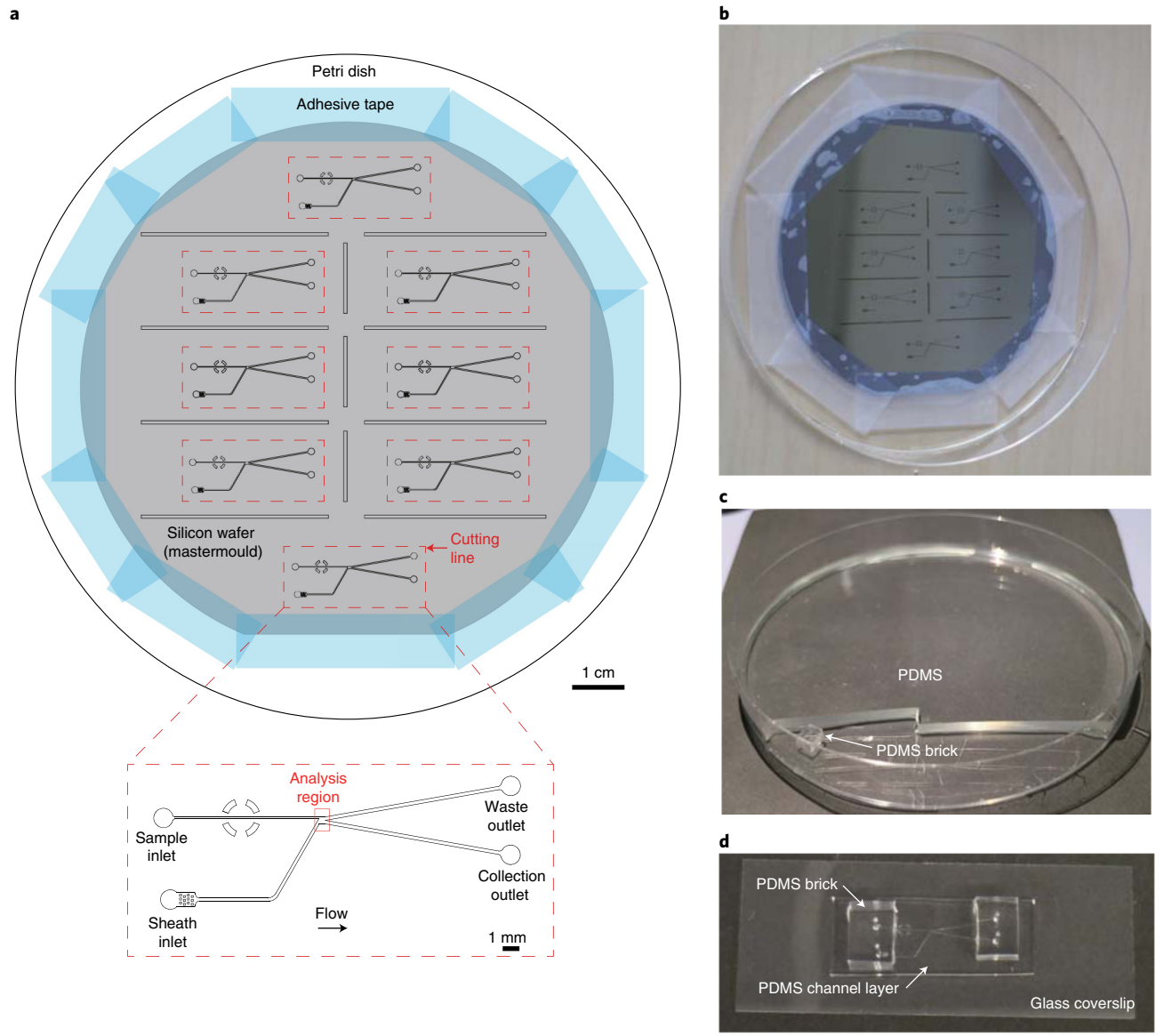


Fig. 9 | Design of the master mold (for the fabrication of eight identical devices) and configuration for the microfluidic device fabrication. **a**,**b**, Schematic (**a**; inset: magnified image of a single device) and photograph (**b**) of the master mold. **c**, A flat PDMS layer (to be attached to the PDMS channel layer) on which to assemble the microfluidic tubing with a thin PDMS channel layer (see main text). **d**, An assembled microfluidic device showing the three layers (from top: PDMS brick, PDMS channel layer, glass coverslip).

12 Treat the cleaned surfaces of the microfluidic device layer and the two bricks for 1 min with a handheld corona treater.

▲ **CRITICAL STEP** The entire surface must be homogeneously treated to achieve good attachment; otherwise, leaks may occur from poorly bonded areas when flow is introduced.

13 Place the two bricks onto the microfluidic device layer (onto the plasma-treated side lacking channels), so that one covers the two inlet ports and one covers the two outlet ports (Fig. 9d).

14 Place the assembly on a clean Petri dish with the bricks facing downward (to facilitate detachment from the Petri dish after baking because of the smaller surface area in contact) and bake it in a convection oven at 75 °C for >30 min to allow chemical linkage.

■ **PAUSE POINT** The microfluidic assembly can be kept in the convection oven at 75 °C for a few days or kept at room temperature for several months.

15 Punch holes at the four ports (two inlets, two outlets) with a biopsy punch, pushing the punch all the way through the PDMS device.

? TROUBLESHOOTING

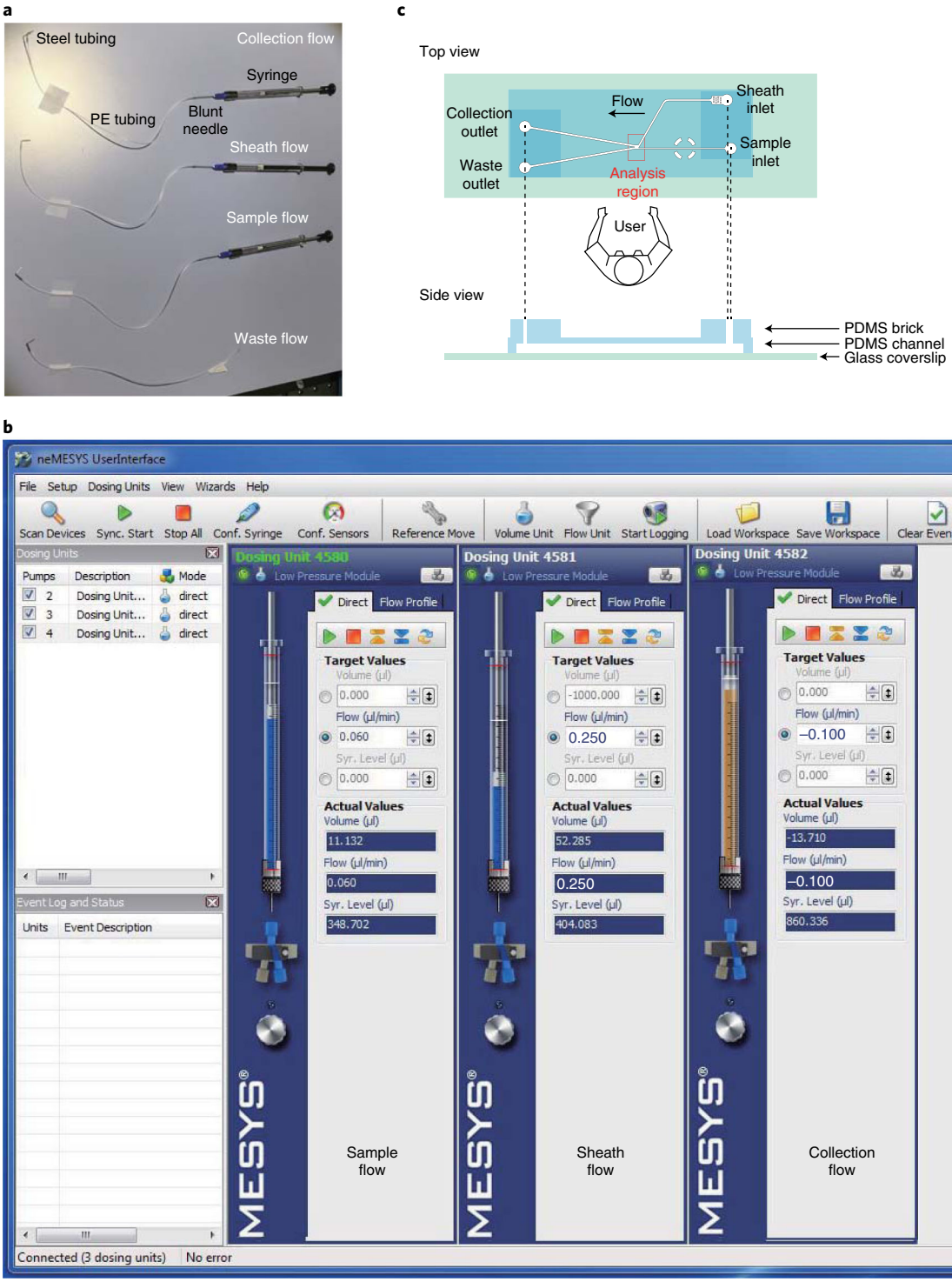


Fig. 10 | Three components required to operate the RACS. a, Photograph of the tubing assemblies: the top three are to carry the sample, the sheath, and the collection fluids, and the bottom assembly is to carry the waste. **b**, Software window to operate the syringe pump. **c**, Orientation of the microfluidic device for the attachment of the tubing assemblies. In comparison with Fig. 6a, the device is mirrored on the *x* axis, because it will be flipped over when it is installed on the Raman microscope system.

16 Clean the surface of the PDMS device (microchannel side) and one surface of a glass coverslip (where they will be attached together in Step 18) by carefully attaching and then detaching adhesive tape. Repeat this step three times with new, clean adhesive tape each time.

17 Treat the cleaned sides of the PDMS device and the coverslip for 1 min using a handheld corona treater as described in Step 12.

18 Place the PDMS device on the glass coverslip (with the surfaces treated by the corona in contact) and bake the assembly in a convection oven at 75 °C for >30 min to allow chemical linkage.

▲ **CRITICAL STEP** Given the thinness of the microfluidic device cut from PDMS layer 1, place the device onto the coverslip carefully, starting from one side and gradually bringing the rest into contact to avoid the formation of wrinkles or air bubbles.

■ **PAUSE POINT** The assembly can be kept in the convection oven at 75 °C for a few days or kept at room temperature for several months. Store the device in a sterile Petri dish until used. Although not autoclavable, the device will be exposed to fluid flow, providing vigorous cleaning conditions, before the actual sorting (see Step 42).

Sorting of target cells ● Timing ~ 4 h (1.5–2 h hands-on)

19 *Starting the Raman system and preparing the working area (Steps 19–23).* Turn on all equipment that is required for the RACS. The main power supply of the Raman system (powering CCD camera 1, the confocal pinhole, the grating motor, and the spectrometer; Fig. 4a) is always kept on. Turn on (i) the two lasers (532 nm and 1,064 nm for the Raman measurement and the optical tweezers, respectively), (ii) the two laser shutters (shutter 1 blocks the two lasers simultaneously and shutter 2 blocks only the Raman laser; both open as default), and (iii) camera 2 and the low-angle ring illuminator for the lower microscope (Fig. 4).

▲ **CRITICAL STEP** The two lasers require an ~10-min warm-up.

20 Calibrate the system with the LabSpec software. First, place a silicon wafer on the microscope stage and turn on camera 1. Change the laser density filter to <1% and focus the Raman laser on the silicon wafer with the 10× objective (the laser spot size is at its minimum when focused), then change to the 100× objective. Block the laser, change the laser density filter back to 100%, turn off camera 1, withdraw the removable beam splitter (Fig. 4a; the system is then in 'Raman measurement' mode), and click the 'AC' (meaning 'auto-calibration') button. Then select 'Current laser/grating'.

21 Place the removable beam splitter back into the laser path and change the settings in 'LabSpec': 100 for 'slit' and 300 for 'pinhole'.

22 Fix the LabSpec software bug that causes the RACS software to freeze: go to the folder 'C:\Horiba\LabSpec_6_3_40\Parameters\COMMON' (version '6_3_40'; this may vary depending on the system) and check whether the file 'OLActiveX.bin' is ~2 MB. If not, overwrite the existing file by that with the same file name in the folder one level higher 'C:\Horiba\LabSpec_6_3_40\Parameters' (see also Step 54).

? TROUBLESHOOTING

23 Clean the working area (microscope stage, optical table, and computer desk) with 70% (vol/vol) EtOH and then DNA AWAY. Clean all metal tools (tweezers, pliers, scissors) with 70% (vol/vol) EtOH followed by flaming, and let them cool before use.

▲ **CRITICAL STEP** This step is intended to minimize contamination, which can affect downstream cultivation for further ecological evaluation and DNA analysis (Step 58).

24 *Resuspending the sample for analysis (Steps 24–26).* Wash the sample with 1 mL isotonic liquid once by centrifugation, using a microcentrifuge at room temperature and 3,287*g* (3 min; mouse colon sample from Step 1A), 3,000*g* (3 min; pure bacterial cultures from Step 1B/C), or 4,500*g* (20 min; marine sediment sample from Step 1D).

▲ **CRITICAL STEP** If a Raman peak of residual D_2O is present (spectral region 2,200–2,600 cm^{-1}; Fig. 3a), wash the sample twice with 1 mL isotonic liquid.

25 Resuspend the sample in 1–2 mL isotonic liquid in a 50-mL conical-bottom tube.

▲ **CRITICAL STEP** Based on a theoretical calculation (Supplementary Fig. 17 in ref. [24]), the maximum throughput can be achieved with a cell density of 10^8 cells/mL (calculation based on *E. coli*). Given that this cell density is not easy to achieve in general, prepare the sample to be as dense as possible while minimizing cell aggregates.

26 Prepare a 15-mL aliquot of the same isotonic liquid in a 50-mL conical bottom tube (for the sheath and collection fluids).

▲ **CRITICAL STEP** For the collection syringe, isotonic liquid can be supplemented by Tween 20 (0.2–0.5% (vol/vol)) for the easy recovery of the sorted cells (see Steps 31 and 58; Fig. 3c).

27 *Assembling the microfluidic device onto the Raman system (Steps 27–40).* Cut four pieces of PE tubing (identical length) and place them into a clean Petri dish (to prevent contamination).
▲ **CRITICAL STEP** The tubing length depends on the distance between the stand where the syringes/syringe pumps are placed during the RACS and the microscope stage where the microfluidic device is assembled (Fig. 6). They must be long enough to allow loose and flexible connection between the two parts; however, twists and entanglements among them may occur if they are too long. It is advisable to use tubing lengths ~10% longer than the shortest distance between the two parts.

28 Gently bend four 23G microfluidic steel tubing connectors to 90° (at their mid-point) with two pairs of pliers.

29 Connect one of the pieces of PE tubing to one of the bent-steel tubing connectors at one end and to a 23-gauge blunt needle at the other. Prepare three of these tubing assemblies, one each for the sample, sheath, and collection flows (Fig. 10a).

30 Connect the remaining piece of PE tubing with a bent steel tubing connector at one end and leave the other end free. This assembly is used for the waste flow (Fig. 10a).

31 Fill three glass syringes with the sample from Step 25 (sample syringe with 500-µL volume) and the isotonic liquid from Step 26 (sheath fluid and collection syringes with 1-mL and 500-µL volumes, respectively). Begin by filling the collection syringe, followed by the sheath and then the sample syringes to minimize the possibility of contamination of the collection syringe by unsorted microbes from the sample.
▲ **CRITICAL STEP** We highly recommend that the syringes be used for only a single application (i.e., users keep their syringes only for their specific applications) and that syringes be changed periodically (every 2 months, but this strongly depends on the experiment frequency).
▲ **CRITICAL STEP** To avoid trapping air bubbles in the syringe while filling, fill half of the syringe volume with fluid and then vigorously shake it with the head facing upward. Large bubbles will gather and merge at the top, and small bubbles can be made to rise by tapping the syringe body. Push the plunger until all bubbles are removed and then fill with fluid again.

32 Attach each of the three syringes to the blunt needle (Luer Lock) of one of the tubing assemblies (prepared in Step 29).
▲ **CRITICAL STEP** A PES syringe filter can be placed between the sheath fluid syringe and the blunt needle to prevent the contamination that may be introduced because the sheath fluid passes through the collection outlet (Fig. 2 and Supplementary Video 1).

33 Push the plungers of the three syringes until the tubing is filled with the fluids and carefully check whether the assembly (steel tubing, PE tubing, blunt needle, and syringe) is totally bubble-free.

34 Start the syringe pump software (neMESYS; Fig. 10b). Click 'Scan device', so that the software connects to the hardware. Click the orange arrow button pointing upward to move each of the three syringe-holder units to the top. Click each syringe module and click 'Conf. Syringe' to set its configuration as follows:
- Left unit: Hamilton, 500 µL (sample)
- Middle unit: Hamilton, 1 mL (sheath fluid)
- Right unit: Hamilton, 500 µL (collection)

35 Install the syringes onto the syringe pump holders (Fig. 6). Ensure that they are in a perfectly vertical position and are well secured. The units where the syringes are hooked up are as follows:
- Rear unit: Hamilton, 500 µL (sample)
- Middle unit: Hamilton, 1 mL (sheath fluid)
- Front unit: Hamilton, 500 µL (collection)

▲ **CRITICAL STEP** The collection syringe should be close to the user to facilitate recovery of the collected cells after the RACS (see Step 56).

36 Push the plungers of the three syringes (using the neMESYS software), using flow rates of 2, 7, and 2 µL/min for the sample, the sheath fluid, and the collection syringes, respectively. Stop the flow once a small droplet appears at the tip of the steel tubing to maintain the bubble-free condition.

37 Place the microfluidic device onto the microscope stage (ensuring that the orientation matches Fig. 10c) and connect the four tubing assemblies (sample, sheath fluid, collection, waste) to the corresponding ports using the steel tubing connectors.
▲ **CRITICAL STEP** Ensure that the four lengths of tubing are not entangled and thus will not interfere with each other when the microfluidic device is flipped over to be immobilized on the microscope stage (see Step 39).

38 Turn on camera 1 in the LabSpec software and change to the 10× objective.

39 Flip over the microfluidic device and fix it on the acrylic supporting plate with adhesive tape and then on the microscope stage using specimen clips.
▲ **CRITICAL STEP** Align the sample channel in the microfluidic device parallel to the *x* axis in the video (obtained using camera 1) so that movement in the *y* direction during sorting is perpendicular to the flow direction (see Step 50).
▲ **CRITICAL STEP** Ensure that the 'collection outlet' is on the side closest to the user after flipping over the microfluidic device (Fig. 6). This enables easy disconnection of the collection tubing after RACS to recover the collected cells (see Step 56).

40 Add 0.5 ml of the isotonic fluid to a microcentrifuge tube (placed in a small rack) and put the end of the waste fluid tubing into this reservoir (Fig. 6).
▲ **CRITICAL STEP** The RACS platform uses very low flow rates (~500 μm/s flow velocity in the analysis region). If the end of the waste fluid tubing were exposed to the air (rather than a liquid reservoir), flow in the device would fluctuate over time because of the growth and then detachment of the droplet.

41 *Running the RACS (Steps 41–52).* Find the analysis region in the microfluidic device (Figs. 2,6), using the 10× objective and camera 1.

42 Start the flow in the neMESYS software controlling the syringe pumps in the following order: collection, then sheath fluid, then sample, using flow rates of 2, 7, and 2 μL/min, respectively.
▲ **CRITICAL STEP** This order is necessary to prevent the sample from entering the collection or sheath fluid channels. Start each subsequent flow when the previous one has reached the analysis region.
▲ **CRITICAL STEP** The flow rates are ~10 times higher than those used during the actual sorting to rapidly fill the system (channels in the microfluidic device and the tubing for the waste) with the fluids while avoiding leakage due to the pressure increase.
▲ **CRITICAL STEP** An air bubble may form and become trapped in the analysis region (typically at the side wall between the channels for the sheath and collection fluids; Fig. 2). In this case, use higher flow rates for the sheath and collection flows (10 and 5 μL/min, respectively). The bubble should shrink and disappear within a few minutes; PDMS is gas permeable.
? TROUBLESHOOTING

43 Start the camera software for the lower microscope and check the flow in the microfluidic device. Click 'View B/W' → 'Live Preview' → 'Convert Control' (hovering over icons will display their names). Move the lower microscope upward and click 'Auto Range Peak' periodically (every few hundred microns of movement). When the objective is focused within the microfluidic device, adjust the contrast, using the lookup table (LUT) in the 'Convert Control' window. Once the three flows (sample, sheath fluid, and collection) are introduced and stabilized in the analysis region, withdraw the lower microscope.

44 Move the microscope stage 6 mm downward, place 2–3 drops of immersion oil Immersol W 2010 onto the microfluidic device (on the coverslip side, where it will contact the 63× objective), and change the objective (of the Raman system) to the 63× water-immersion objective.

45 In the LabSpec software, set the laser density filter to 1% and turn on the Raman laser. Move the stage upward and focus the Raman laser on the analysis region in the microfluidic device. Two sequential laser foci (at the upper and lower surfaces of the coverslip) are shown in the video (camera 1 in LabSpec). Move the stage 15 μm above the second focus to where the RACS will take place.
▲ **CRITICAL STEP** The proper laser focus can be identified as when the beam spot size becomes minimal.
▲ **CRITICAL STEP** There is a discrepancy between the physical thickness (170 μm) and the optical thickness (~150 μm; measured by the distance between the two laser foci) of the coverslip.

46 Turn off camera 1 in the LabSpec software, set the laser density filter back to 100%, and withdraw the removable beam splitter (which places the system in 'Raman' mode; Fig. 4a).
▲ **CRITICAL STEP** Camera 1 in the Raman system will not work from this point onward, so any visual inspection must be performed using the lower microscope (camera 2).

47 Once the waste fluid tubing is fully filled with fluid, ensure that the system is bubble-free. Carefully check the microfluidic device, syringes, and tubing assemblies, particularly the end of the waste

fluid tubing in the microcentrifuge waste reservoir. Secure the waste fluid tubing in the reservoir using adhesive tape to prevent it being pulled out during RACS.

48 Move the lower microscope upward and check the stability of the sample flow. As in Step 43, click 'Auto Range Peak' periodically (every few hundred microns of movement) in the Camware software and adjust the contrast using the LUT in the 'Convert Control' panel.

49 Decrease the flow rates (in the neMESYS software) in the order sample, then in the sheath fluid, and then in the collection channel to 0.06, 0.25, and −0.1 (withdrawal) μL/min, respectively. Ensure that the flow is stable and that all cells pass through the waste outlet by default.

? TROUBLESHOOTING

50 Place the laser spot at the side wall between the channels for the sheath and collection fluids. Move the laser spot 300 μm away from the side wall (in the *y* direction). In the *x* direction, place the laser spot at the edge of the cell stream (see Supplementary Video 1), so as not to lose the captured cell during the sorting (in Step 52) as a result of collision with other cells during translocation to the release location. Turn on the optical tweezers laser. Check whether a cell is captured in the laser spot (the optical tweezers). Release the captured cell by blocking the tweezers laser (using the laser shutter 1). Repeat cell capture and release several times. Once it has been confirmed that all works well, turn off the 'Live Preview' (in the Camware software) and close the LabSpec software.

▲ **CRITICAL STEP** If the next cell is captured immediately after releasing the previous one, the cell concentration in the sample fluid can be considered sufficiently high.

▲ **CRITICAL STEP** Raman laser power for the RACS should be adjusted on the basis of the sample conditions (see Box 1).

51 Open 'MATLAB' and select the 'RACS_ver1.fig' file ('RACS_ver2.fig' for the machine learning–based software platform). Right-click and select 'Open in GUIDE'. Click the green 'play triangle' button and the GUI working platform appears. The word GUIDE stands for Graphical User Interface Development Environment. The '.fig' format is for the working platform and the '.m' format contains the corresponding code.

▲ **CRITICAL STEP** If MATLAB and LabSpec are simultaneously open while trying to control the Raman system, the system will crash.

▲ **CRITICAL STEP** The system is in the RACS mode now. From this point onward, the microscope stage should be controlled by the MATLAB code only; do not move it manually or by using a joystick.

52 When the GUI opens, the two laser shutters are initially closed. Open them by pushing the physical 'open' button at the shutter controllers. To run 'RACS_ver1.fig' to count cells, follow option A; to run 'RACS_ver1.fig' to sort cells, follow option B; to run 'RACS_ver2.fig', follow option C.

▲ **CRITICAL STEP** With 'RACS_ver1.fig' (workflow displayed in Fig. 8b), the platform can be used as a cell counter (option A) or as a cell sorter (option B). As a cell counter, it can be used to determine the threshold value for P_L and to evaluate the relative abundance of taxa of interest within the microbial community sample when running the RACS with a new sample. As a sorter, it provides a means to sort taxa of interest from within the microbial community sample. 'RACS_ver2.fig' performs a *K*-means clustering based on the analysis of an initial sample of cells and then sorts the remaining cells on the basis of that clustering (workflow displayed in Fig. 8d).

▲ **CRITICAL STEP** Operating the platform with respect to the specific sorting criteria (deuterium labeling, ^{13}C-incorporation, the possession of cytochrome *c*, or carotenoids) is possible by adjusting the spectral window of interrogation. The relevant parts are marked within the accompanying scripts (RACS_ver1.m; RACS_ver2.m).

(A) **'RACS_ver1.fig' to count cells**

(i) Click 'Spectra calibration'. Laser shutter 1 is closed; the stage moves 70 μm in the *y* direction; laser shutter 1 opens; the spectroscope records the calibration Raman spectrum (measuring the working fluid in the absence of a cell, averaged over 10 measurements); the stage moves back to the initial position; and the calibration spectrum is shown in two panels (panels 16 and 17 in Fig. 8a).

(ii) Click 'Start the process' and the RACS begins. Set 'PL_input' (panel 6 in Fig. 8a) to 100, so that all captured cells are considered to be 'not labeled'; the software records the P_C and P_L values of the cells in the sample without sorting.

? TROUBLESHOOTING

(B) **'RACS_ver1.fig' to sort cells**

(i) Click 'Spectra calibration'. Laser shutter 1 is closed; the stage moves 70 μm in the *y* direction; laser shutter 1 opens; the calibration Raman spectrum is measured; the stage moves back to the initial position; and the calibration spectrum is shown in two panels.

(ii) Click 'Start the process' and the RACS begins; no human intervention is needed until the end of the sorting.

▲ **CRITICAL STEP** The threshold value for P_L is set to 100 by default, so that no cells are sorted until an appropriate threshold value is set when the RACS begins. On the basis of these measurements in the 'Counter mode' (option A), an appropriate threshold value can be set in panel 6 after clicking the 'start the process' button (see below for a guide to the selection of the threshold value).

? TROUBLESHOOTING

(C) **'RACS_ver2.fig' for *K*-means clustering and sorting**

(i) Click 'Calibration'. Laser shutter 1 is closed; the stage moves 70 μm in the *y* direction; laser shutter 1 is opened; the calibration Raman spectrum is measured; the stage moves back to the initial position; and the calibration spectrum is shown in two panels (panels 15 and 16 in Fig. 8c).

(ii) Click 'Clustering', and the software measures cells in the same way as in 'Counter mode' (i.e., the software accumulates the P_C and P_L values of the cells in the sample without sorting).

(iii) Click 'Stop' and the software conducts the calculation for the *K*-means clustering.

(iv) Click 'Start RACS' and the software asks for a file that contains the data measured during the optimization of the 'Clustering'. Upon selection of a file 'Clustering_Spectra_1981-10-07 19_02_07.txt' (named using the date and time '1981-10-07 19_02_07'; this will vary depending on the timing of the RACS operation), the RACS begins based on the clusters; no human intervention is needed until the end of the sorting.

? TROUBLESHOOTING

Collection of sorted cells and their downstream analysis ● Timing 24.5–48.5 h for downstream cell cultivation; variable for DNA analysis

53 Click 'Stop' and then 'Exit' to terminate the sorting (panels 14 and 15 in Fig. 8a; panels 13 and 14 in Fig. 8c). Wait for 5 min to allow all the sorted cells to reach the collection tubing. While waiting, calculate the fluid volume that was collected in the collection tubing during the sorting; for example, if the process was run for 1 h, 0.1 μL/min × 60 min = 6 μL. Also, turn on the lower microscope and check that there are no fluctuation in the flow and no bubbles in the system.

▲ **CRITICAL STEP** Ensure that cells do not enter the collection channel by default as a result of flow instability. If the isotonic liquid is not regularly autoclaved or is not kept at room temperature (see Reagent setup), or if the platform encounters large temperature shifts during the sorting, air dissolved in the isotonic liquid could form bubbles that lead to flow instability.

54 Fix the LabSpec software bug: overwrite the file 'OLActiveX.bin' in the folder 'C:\Horiba\LabSpec_6_3_40\Parameters\COMMON' (v. '6_3_40'; this may vary depending on the system) by that with the same file name in the folder one level higher, 'C:\Horiba\LabSpec_6_3_40\Parameters'.

? TROUBLESHOOTING

55 Withdraw the lower microscope from the microfluidic device.

▲ **CRITICAL STEP** Ensure that the microfluidic device and the connected tubing assemblies are not touched during the movement of the lower microscope. The system uses an ultra-low flow rate and is highly susceptible to interference that may cause flow instability and non-target cells to enter the collection channel.

56 Grip the steel tubing as closely as possible to the microfluidic device with blunt-end tweezers and rapidly detach the collection tubing from the microfluidic device (Fig. 6a).

57 Use the neMESYS software to stop all flows.

58 To use the sorted cells for subsequent cultivation, follow option A or B. For DNA analysis, follow option C.

(A) **Cultivation of cells in liquid medium**

(i) Put the steel tubing (at the end of the collection tubing) into a microcentrifuge tube filled with 500 µL culture medium (e.g., YCFA medium for the mouse colon sample).

(ii) Set the flow rate to 2 µL/min ('infusion') and eject the collected cells until the volume ejected reaches double the volume collected during the sorting (to ensure that all sorted cells are recovered); for example, if the RACS has been run for 1 h, 0.1 µL/min × 60 min × 2 = 12 µL should be collected.

(iii) Culture the sorted cells under appropriate conditions, for example, 37 °C under anaerobic conditions for 24 h for the mouse colon sample.

■ **PAUSE POINT** The sorted cells can be stored for several years at −80 °C until further processing. Cells from the same sorting process can be used simultaneously for downstream genetic characterization and cultivation.

(B) **Cultivation of cells on an agar plate**

(i) Put the steel tubing (at the end of the collection tubing) onto the agar plate, set the flow rate to 2 µL/min ('infusion'), and eject the collected cells until the volume ejected reaches double the volume collected during the sorting (but >50 µL to allow spreading the cells on the agar plate in the next step).

(ii) Spread the cells gently over the surface area using an inoculation spreader.

(iii) Culture the sorted cells under appropriate conditions, for example, 37 °C under anaerobic conditions for 48 h for the mouse colon sample.

■ **PAUSE POINT** The sorted cells can be stored for several years at −80 °C until further processing. Cells from the same sorting process can be used simultaneously for downstream genetic characterization and cultivation.

(C) **DNA analysis**

(i) Put the steel tubing (at the end of the collection tubing) into a PCR tube.

(ii) Set the flow rate to 2 µL/min ('infusion') and eject the collected cells until the volume ejected reaches double the volume collected during the sorting. As controls, also collect the sample and sheath fluids separately from the syringes.

(iii) For mini-metagenomics ('Downstream analysis' in Fig. 1), as in our demonstration[24], the collected cells can be lysed and subjected to whole genome amplification (WGA) using a REPLI-g Single Cell Kit (Qiagen) according to the manufacturer's instructions. Shotgun libraries are generated using the amplified DNA from WGA reactions (sorted fractions) and libraries are sequenced for example with a HiSeq 3000 (Illumina) in 2 × 150 bp mode. For single-cell genomics, each individual cell can be isolated on a multi-well plate using FACS and then sequenced[30].

■ **PAUSE POINT** The sorted cells can be stored for several years at −80 °C until further processing. Cells from the same sorting process can be used simultaneously for downstream genetic characterization and cultivation.

Cleaning after sorting ● Timing 30 min

▲ **CRITICAL** All components, except for the Hamilton glass syringes (i.e., microfluidic device, tubing assemblies, Petri dish), are for single use and should be disposed of after each experiment.

59 Separate the plunger and the barrel of the Hamilton glass syringes and thoroughly clean them with Milli-Q water using a wash bottle.

▲ **CRITICAL STEP** This procedure removes salt crystals if ASW was used as the working fluid. It is important to remove them carefully in this step because if present, they may precipitate when in contact with acetone (Step 62).

60 Assemble the plunger and the barrel, fill and empty each syringe with Milli-Q water three times, and repeat this three times using fresh Milli-Q water.

61 Separate the plunger and the barrel and thoroughly clean them with acetone, using a wash bottle.

62 Assemble the plunger and the barrel, fill and empty each syringe with acetone three times, and repeat this three times using fresh acetone.

63 Separate the plunger and the barrel and dry them under UV light to provide additional sterilization.

■ **PAUSE POINT** The syringes can be stored on clean laboratory tissues during frequent experiments (more than once per week) or in the original syringe packaging box.

Troubleshooting

Troubleshooting advice can be found in Table 3.

Table 3 | Troubleshooting table

Step	Problem	Possible reason	Solution
15, 42	Microfluidic device leaks	Biopsy punch is old and thus blunted	Carefully inspect the biopsy punch before use to ensure that the hollow tip (to punch the PDMS) is sharp. Periodically change the punch (after fabricating 40 microfluidic devices)
22, 54	RACS software freezes	Software bug: when the LabSpec is controlled with third-party software (e.g., code built in-house using MATLAB, as in this protocol), it accumulates binary log data into the file 'OLActiveX.bin' over the RACS operation, which decreases the speed of the code operation over time and eventually freezes the operation	Before the first RACS operation in a new Raman system, copy the file 'OLActiveX.bin' in the folder 'C:\Horiba\LabSpec_6_3_40 \Parameters\COMMON' (v. '6_3_40'; this may vary depending on the system) and paste it into the folder one level higher ('C:\Horiba \LabSpec_6_3_40\Parameters'). Then, for RACS operation, follow the instructions in Steps 22 and 54
49	Cell stream does not exit into the waste channel by default	There is an undetected bubble in the system or the system has leakage (particularly at the intersections between the microfluidic device and the tubing assemblies)	Use a new microfluidic device
		The syringes are old; thus the plunger movement is not smooth/stable	Replace the syringes with new ones
52A–C	Calibration Raman spectrum is not measured	Software bug: when third-party software (the code built in-house using MATLAB in this protocol) initially controls the LabSpec software, a first Raman measurement (i.e., that for the calibration) measures nothing, and a spectrum of a noisy signal is acquired	Click 'Spectra calibration' (panel 3 or 1 in Fig. 8a or c, respectively) one more time
	Fluorescence is dominant; thus the Raman signal is veiled (Fig. 3a,b)	There is residual D_2O or medium for incubation within the sample after washing	Wash the sample one time more (Step 24)
		The water-immersion objective has damage on the surface	Check the objective surface by eye and measure the Raman spectra within the immersion oil and in the microfluidic device. If the Raman signal in the immersion grows over time and it contributes to the signal measured in the microfluidic device (Supplementary Fig. 2), the objective surface is damaged, and thus should be replaced
	The cellular Raman signal is weak	Laser focus difference between the Raman and optical tweezers lasers is large (Fig. 5 and Table 1)	Use the recommended objective (Zeiss, cat. no. C-Apochromat 63×/1.20 W Korr UV-VIS-IR M27 63×) and set the coverslip correction collar to 0.17 mm
		Confocality of the Raman system is poorly corrected	Check whether the pinhole size is correctly set as 300 (Step 21)

Timing

Step 1A, deuterium (D)-incorporation into the mouse colon sample: 8 h
Step 1B, D-incorporation into pure bacterial cultures: 24 h
Step 1C, ^{13}C or ^{13}C+D incorporation into pure bacterial cultures: 24 h
Step 1D, enrichment procedure for marine sediment samples containing microbes expressing high levels of cytochrome *c*: 8 weeks
Steps 2–18, fabrication of microfluidic device: 3 h
Steps 19–23, starting the Raman system and preparing the working area: 30 min
Steps 24–26, resuspending the sample for analysis: 0.5–1 h
Steps 27–40, assembling the microfluidic device onto the Raman system: 30 min
Steps 41–52, running the RACS: on the order of 2–2.4 h, depending on the number of cells to be sorted

Steps 53–58, collection of sorted cells and their downstream analysis: 24.5–48.5 h for downstream cell cultivation; variable for DNA analysis
Steps 59–63, cleaning after sorting: 30 min

Anticipated results

Choice of sorting parameters

The platform should be used to evaluate the characteristics of the sample before sorting in order to set sorting parameters appropriate to each application (i.e., used as a cell counter, Step 52A). Measurements of deuterium-labeled and unlabeled samples of three pure bacterial cultures show that, regardless of labeling status, $P_C > 1.0$ when a single cell is captured in the optical tweezers (Fig. 7a), with P_C values of 1.42 ± 0.16 (mean ± s.d.; $n = 52$), 1.37 ± 0.12 ($n = 55$), and 1.32 ± 0.17 ($n = 60$) for *V. alginolyticus*, *V. cyclitrophicus*, and *E. coli*, respectively. Using $P_L = 6.5$, the platform differentiated the D-labeled cells from the unlabeled cells (Fig. 7a): P_L values were 7.47 ± 0.48 ($n = 23$), 7.49 ± 0.33 ($n = 26$), and 7.40 ± 0.48 ($n = 30$) for the labeled cells, and 6.07 ± 0.13 ($n = 29$), 5.86 ± 0.11 ($n = 29$), and 5.85 ± 0.18 ($n = 30$) for unlabeled cells (for *V. alginolyticus*, *V. cyclitrophicus*, and *E. coli*, respectively, in each case). These types of measurements using pure bacterial cultures are useful when users establish the RACS platform in their system to check its performance before application to the environmental samples of interest.

To apply the RACS system to a natural community sample, a measurement using a control must precede the actual sorting in order to estimate the threshold for P_L; a value greater than the values measured in the control sample should be used during RACS. For the mouse colon microbiota, for example, control cells incubated in non-D_2O-containing PBS supplemented with glucose displayed $P_C > 1.1$ ($P_C = 1.37 \pm 0.38$; $n = 151$) upon single-cell capture and $P_L < 6.19$ (chosen on the basis of 'mean + 2 s.d.'; $P_L = 5.92 \pm 0.13$).

Our software platform can be tuned to measure and sort cells in terms of other molecular parameters (in lieu of deuterium-labeling status). We examined the potential of RACS to sort with respect to ^{13}C, alone and coupled with deuterium labeling, and with respect to the possession of carotenoids. The ^{13}C-labeling status can be identified by measuring a phenylalanine peak that is shifted from 1,007 to 971 cm^{-1} when a cell is ^{13}C-labeled (see the plots for ^{13}C and ^{12}C in Fig. 7f)[15]. To quantify this, we determined the 'carbon labeling index' P_{LC} (Eq. (3)). Measurements of *E. coli* cells cultured in ^{13}C and ^{12}C show that, using $P_{LC} = 0.66$, the platform could distinguish the ^{13}C cells (0.69 ± 0.01; $n = 92$) from the ^{12}C cells (0.63 ± 0.01; $n = 110$) (Fig. 7g). Meanwhile, for cells labeled with both ^{13}C and deuterium, the use of $P_{LC} = 0.66$ identified 73.3% (111 out of 150) of cells, because the deuterium labeling decreases the intensity of the ^{13}C-phenylalanine peak (see the plots for '^{13}C +D' and '^{13}C' in the inset in Fig. 7f); this may occur because several hydrogen atoms (H) of the phenylalanine ($C_9H_{11}NO_2$) molecule are replaced by deuterium (D)[76]. By contrast, the use of $P_L = 6.5$ still distinguished all *E. coli* cells labeled with both ^{13}C and D from those labeled with only ^{13}C or non-labeled *E. coli* cells (Fig. 7h): values were 6.00 ± 0.13, 6.11 ± 0.14, and 8.58 ± 0.60 for ^{12}C, ^{13}C, and ^{13}C+D-labeled cells, respectively.

Use of the software platform as a means to identify and sort carotenoid-containing cells is straightforward because carotenoids generate a resonance Raman signal that is ~1,000 times stronger than regular spontaneous Raman signals (Fig. 7i)[45]. Given this, it is possible to sort on the basis of the raw signal intensity, rather than an intensity ratio between two spectral regions to normalize as in Eqs. (1)–(3). For example, using $I_{1,500\text{-}1,560} = 5{,}000$ (where I represents an integrated intensity) easily distinguished carotenoid-containing *E. voratum* (*Symbiodinium* CCMP421) cells (22,645.5 ± 2,234.7; $n = 20$) from non-carotenoid-containing *V. alginolyticus* cells (498.0 ± 73.6; $n = 80$).

RACS results

For the pure bacterial cultures, on the basis of the sample evaluation (Fig. 7a), we chose 1.1 and 6.5 as threshold values for P_C and P_L, respectively, and conducted a 1-h RACS experiment using a 1:1 mixture of labeled and unlabeled *V. alginolyticus* cells. The P_C values were 1.61 ± 0.17 and 1.47 ± 0.28 for the 'captured and selected' and 'captured and rejected' fractions, respectively. $P_L > 6.5$ distinguished the selected cells (6.96 ± 0.22) from the rejected cells (6.10 ± 0.18). All the captured and selected fraction ($n = 193$) showed a clear C–D peak at 2,040–2,300 cm^{-1}, and the captured and rejected fraction ($n = 232$) showed only a weak signal intensity (coming from water) in this spectral region (Fig. 7b).

For the mouse colon microbiota, on the basis of the measurement of a control sample (Fig. 7c,d), we analyzed 215 cells in the test sample using $P_C > 1.1$ to identify the capture of a cell (with overall value 1.39 ± 0.39 and values of 1.34 ± 0.40 and 1.40 ± 0.39 for the 'captured and selected' and 'captured and rejected' fractions, respectively). In this sample, in which cells were labeled for the ability to degrade mucin, 14.9% of the cells (32 out of 215) were identified as D-labeled and sorted (using $P_L > 6.19$, with $P_L = 6.38 \pm 0.21$ and $P_L = 5.94 \pm 0.17$ for the 'captured and sorted' and 'captured and rejected' fractions, respectively).

Further flexibility of the pipeline was achieved by implementing the machine-learning algorithm (*K*-means clustering) in the RACS platform (Fig. 8c,d). We examined this sorting algorithm using a marine sediment sample in which cells expressing a high level of cytochrome *c* were our taxa of interest to be sorted. The Raman spectra of this sample changes depending on its condition (e.g., oxidation state[70]) and this new platform takes this variation into account by initially measuring the sample under its current condition before the RACS itself begins. The algorithm selected a clustering into five clusters on the basis of the criterion that the number of cells in the uppermost cluster (cluster 1) does not change when increasing the number of clusters (see inset in Fig. 7j). This strategy was chosen because cytochrome *c* generates a strong Raman signal (due to resonance Raman scattering) and thus cells containing cytochrome *c* occupy the uppermost cluster (cluster 1 in Fig. 7j,k). The level of cytochrome *c* expression will differ depending on the species present within the community. Users can choose the number of clusters to be sorted in the graphical user interface (panel 6 in Fig. 8c), and iteration between the RACS and the downstream DNA analysis must be carried out to select the number of clusters to be sorted to meet the user's demand. The machine learning–based RACS software is also applicable to sorting carotenoid-containing cells, such as *E. voratum*, from non-carotenoid-containing cells (*V. alginolyticus*; Supplementary Fig. 3). Unlike what we did for the marine sediment sample, we set two clusters without conducting iteration because the measurements of those two populations are clearly distinct; the difference in the signal intensity between resonance and regular Raman scattering is on the order of 10^3. Clustering using either a wide spectral region (400–3,300 cm^{-1}) or a narrow region corresponding to C=C stretching (1,500–1,560 cm^{-1}; representative of carotenoids) was able to successfully distinguish between the two populations. The trade-off between sensitivity and computation time must be considered for future use; analysis using a wide spectral region yields a greater separation but requires longer computation time (Supplementary Fig. 3b,c).

Collection of taxa of interest and their downstream analysis

The collected cells after RACS, all of which share a specific ecological function, can be used for downstream molecular characterization (DNA analysis using 16S rRNA gene amplicon sequencing, metagenomics, or single-cell genomics after isolating single cells using FACS; Fig. 1) or cultivation for further ecological characterization, such as measuring the response to changes in various physico-chemical gradients[40]. In this protocol, we present the use of cell cultivation using the mouse colon sample. The collected cells after RACS were both spread on agar plates and cultivated in liquid medium, where growth was quantified after 48 h growth by colony counts and after 24 h growth by cell density, respectively (Fig. 7e). Although the agar plate and culture medium were not optimized for the growth of the cells from the mouse colon community, which has great heterogeneity in physiology, this trial demonstrates the possibility of cultivation after sorting from a complex community. The recovery efficiency of cultivable cells was previously quantified using a pure bacterial culture (*M. adhaerens*, HP15), with an efficiency of 81.8 ± 5.9%[24].

Reporting Summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The datasets generated or analyzed in this protocol are available from the corresponding authors upon request. The source data used to create Figs. 3, 5c, and 7 and Supplementary Figs. 2 and 3c,d are provided as Source Data files. Source data are provided with this paper.

Code availability

The MATLAB code to operate the RACS is included in Supplementary Software 1 and 2 of this paper and is also available from GitHub (https://github.com/harubang2/MATLAB-platform-for-Raman-activated-cell-sorting-RACS).

References

1. Lloyd-Price, J. et al. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. *Nature* **569**, 655–662 (2019).
2. Blainey, P. C., Mosier, A. C., Potanina, A., Francis, C. A. & Quake, S. R. Genome of a low-salinity ammonia-oxidizing archaeon determined by single-cell and metagenomic analysis. *PLoS ONE* **6**, e16626 (2011).
3. Thomas, T., Gilbert, J. & Meyer, F. Metagenomics -- a guide from sampling to data analysis. *Microb. Inform. Exp.* **2**, 3 (2012).
4. Horgan, R. P. & Kenny, L. C. 'Omic' technologies: genomics, transcriptomics, proteomics and metabolomics. *Obstet. Gynaecol* **13**, 189–195 (2011).
5. Prosser, J. I. Dispersing misconceptions and identifying opportunities for the use of 'omics' in soil microbial ecology. *Nat. Rev. Microbiol.* **13**, 439–446 (2015).
6. Yu, F. B. et al. Microfluidic-based mini-metagenomics enables discovery of novel microbial lineages from complex environmental samples. *eLife* **6**, e26580 (2017).
7. Mukherjee, S. et al. Genomes OnLine database (GOLD) v.7: updates and new features. *Nucleic Acids Res* **47**, D649–D659 (2019).
8. Woyke, T., Doud, D. F. R. & Schulz, F. The trajectory of microbial single-cell sequencing. *Nat. Methods* **14**, 1045–1054 (2017).
9. Berry, D. & Loy, A. Stable-isotope probing of human and animal microbiome function. *Trends Microbiol* **26**, 999–1007 (2018).
10. Manefield, M., Whiteley, A. S., Griffiths, R. I. & Bailey, M. J. RNA stable isotope probing, a novel means of linking microbial community function to phylogeny. *Appl. Environ. Microbiol.* **68**, 5367–5373 (2002).
11. Dumont, M. G. & Murrell, J. C. Stable isotope probing—linking microbial identity to function. *Nat. Rev. Microbiol.* **3**, 499–504 (2005).
12. Wilhelm, R. C., Singh, R., Eltis, L. D. & Mohn, W. W. Bacterial contributions to delignification and lignocellulose degradation in forest soils with metagenomic and quantitative stable isotope probing. *ISME J* **13**, 413–429 (2019).
13. Wang, Y., Huang, W. E., Cui, L. & Wagner, M. Single cell stable isotope probing in microbiology using Raman microspectroscopy. *Curr. Opin. Biotechnol.* **41**, 34–42 (2016).
14. Haider, S. et al. Raman microspectroscopy reveals long-term extracellular activity of chlamydiae. *Mol. Microbiol* **77**, 687–700 (2010).
15. Huang, W. E. et al. Raman-FISH: combining stable-isotope Raman spectroscopy and fluorescence in situ hybridization for the single cell analysis of identity and function. *Environ. Microbiol.* **9**, 1878–1889 (2007).
16. Wagner, M. Single-cell ecophysiology of microbes as revealed by Raman microspectroscopy or secondary ion mass spectrometry imaging. *Annu. Rev. Microbiol.* **63**, 411–429 (2009).
17. Berry, D. et al. Tracking heavy water (D_2O) incorporation for identifying and sorting active microbial cells. *Proc. Natl Acad. Sci. USA* **112**, E194–E203 (2015).
18. Malmstrom, R. R. & Eloe-Fadrosh, E. A. Advancing genome-resolved metagenomics beyond the shotgun. *mSystems* **4**, e00118–e00119 (2019).
19. Neufeld, J. D. et al. DNA stable-isotope probing. *Nat. Protoc.* **2**, 860–866 (2007).
20. Jing, X. et al. Raman-activated cell sorting and metagenomic sequencing revealing carbon-fixing bacteria in the ocean. *Environ. Microbiol.* **20**, 2241–2255 (2018).
21. Wang, Y. et al. Raman activated cell ejection for isolation of single cells. *Anal. Chem.* **85**, 10697–10701 (2013).
22. Singer, E., Wagner, M. & Woyke, T. Capturing the genetic makeup of the active microbiome *in situ*. *ISME J* **11**, 1949–1963 (2017).
23. Huang, W. E., Ward, A. D. & Whiteley, A. S. Raman tweezers sorting of single microbial cells. *Environ. Microbiol. Rep* **1**, 44–49 (2009).
24. Lee, K. S. et al. An automated Raman-based platform for the sorting of live cells by functional properties. *Nat. Microbiol.* **4**, 1035–1048 (2019).
25. Lee, K. S., Wagner, M. & Stocker, R. Raman-based sorting of microbial cells to link functions to their genes. *Microb. Cell* **7**, 62–65 (2020).
26. Premvardhan, L., Bordes, L., Beer, A., Büchel, C. & Robert, B. Carotenoid structures and environments in trimeric and oligomeric fucoxanthin chlorophyll a/c_2 proteins from resonance Raman spectroscopy. *J. Phys. Chem. B* **113**, 12565–12574 (2009).
27. Takano, H. The regulatory mechanism underlying light-inducible production of carotenoids in non-phototrophic bacteria. *Biosci. Biotechnol. Biochem.* **80**, 1264–1273 (2016).
28. Wagstaff, K., Cardie, C., Rogers, S. & Schrödl, S. Constrained k-means clustering with background knowledge. in *Proc. 18th International Conference on Machine Learning* (eds Brodley, C. E. & Danyluk, A. P.) 577–584 (Morgan Kaufmann, 2001).

29. Kanungo, T. et al. An efficient k-means clustering algorithm: analysis and implementation. *IEEE Trans. Patt. Anal. Mach. Intell.* **24**, 881–892 (2002).
30. Rinke, C. et al. Obtaining genomes from uncultivated environmental microorganisms using FACS-based single-cell genomics. *Nat. Protoc.* **9**, 1038–1048 (2014).
31. Bonner, W. A., Hulett, H. R., Sweet, R. G. & Herzenberg, L. A. Fluorescence activated cell sorting. *Rev. Sci. Instrum.* **43**, 404–409 (1972).
32. Ha, B. H., Lee, K. S., Jung, J. H. & Sung, H. J. Three-dimensional hydrodynamic flow and particle focusing using four vortices Dean flow. *Microfluid. Nanofluid.* **17**, 647–655 (2014).
33. Chu, H., Doh, I. & Cho, Y.-H. A three-dimensional (3D) particle focusing channel using the positive dielectrophoresis (pDEP) guided by a dielectric structure between two planar electrodes. *Lab Chip* **9**, 686–691 (2009).
34. Gao, C. et al. Single-cell bacterial transcription measurements reveal the importance of dimethylsulfoniopropionate (DMSP) hotspots in ocean sulfur cycling. *Nat. Commun.* **11**, 1942 (2020).
35. Kitzinger, K. et al. Single cell analyses reveal contrasting life strategies of the two main nitrifiers in the ocean. *Nat. Commun.* **11**, 767 (2020).
36. Majed, N., Chernenko, T., Diem, M. & Gu, A. Z. Identification of functionally relevant populations in enhanced biological phosphorus removal processes based on intracellular polymers profiles and insights into the metabolic diversity and heterogeneity. *Environ. Sci. Technol.* **46**, 5010–5017 (2012).
37. Fernando, E. Y. et al. Resolving the individual contribution of key microbial populations to enhanced biological phosphorus removal with Raman–FISH. *ISME J* **13**, 1933–1946 (2019).
38. Milucka, J. et al. Zero-valent sulphur is a key intermediate in marine methane oxidation. *Nature* **491**, 541–546 (2012).
39. Hatzenpichler, R. et al. Visualizing in situ translational activity for identifying and sorting slow-growing archaeal–bacterial consortia. *Proc. Natl Acad. Sci. USA* **113**, E4069–E4078 (2016).
40. Schiessl, K. T. et al. Phenazine production promotes antibiotic tolerance and metabolic heterogeneity in *Pseudomonas aeruginosa* biofilms. *Nat. Commun.* **10**, 762 (2019).
41. Gleizer, S. et al. Conversion of *Escherichia coli* to generate all biomass carbon from CO_2. *Cell* **179**, 1255–1263 (2019).
42. Dong, T. G., Ho, B. T., Yoder-Himes, D. R. & Mekalanos, J. J. Identification of T6SS-dependent effector and immunity proteins by Tn-seq in *Vibrio cholerae*. *Proc. Natl Acad. Sci. USA* **110**, 2623–2628 (2013).
43. Dolinšek, J., Lagkouvardos, I., Wanek, W., Wagner, M. & Daims, H. Interactions of nitrifying bacteria and heterotrophs: identification of a *Micavibrio*-like putative predator of *Nitrospira* spp. *Appl. Environ. Microbiol.* **79**, 2027–2037 (2013).
44. Pätzold, R. et al. In situ mapping of nitrifiers and anammox bacteria in microbial aggregates by means of confocal resonance Raman microscopy. *J. Microbiol. Methods* **72**, 241–248 (2008).
45. Wei, L. & Min, W. Electronic preresonance stimulated Raman scattering microscopy. *J. Phys. Chem. Lett.* **9**, 4294–4301 (2018).
46. Gruber-Vodicka, H. R. et al. Paracatenula, an ancient symbiosis between thiotrophic Alphaproteobacteria and catenulid flatworms. *Proc. Natl Acad. Sci. USA*. **108**, 12078–12083 (2011).
47. Lenz, R., Enders, K., Stedmon, C. A., MacKenzie, D. M. A. & Nielsen, T. G. A critical assessment of visual identification of marine microplastic using Raman spectroscopy for analysis improvement. *Mar. Pollut. Bull.* **100**, 82–91 (2015).
48. Gillibert, R. et al. Raman tweezers for small microplastics and nanoplastics identification in seawater. *Environ. Sci. Technol.* **53**, 9003–9013 (2019).
49. Choy, C. A. et al. The vertical distribution and biological transport of marine microplastics across the epipelagic and mesopelagic water column. *Sci. Rep.* **9**, 7843 (2019).
50. Zhang, P. et al. Raman-activated cell sorting based on dielectrophoretic single-cell trap and release. *Anal. Chem.* **87**, 2282–2289 (2015).
51. McIlvenna, D. et al. Continuous cell sorting in a flow based on single cell resonance Raman spectra. *Lab Chip* **16**, 1420–1429 (2016).
52. Folick, A., Min, W. & Wang, M. C. Label-free imaging of lipid dynamics using coherent anti-stokes Raman scattering (CARS) and stimulated Raman scattering (SRS) microscopy. *Curr. Opin. Genet. Dev.* **21**, 585–590 (2011).
53. Hiramatsu, K. et al. High-throughput label-free molecular fingerprinting flow cytometry. *Sci. Adv.* **5**, eaau0241 (2019).
54. Suzuki, Y. et al. Label-free chemical imaging flow cytometry by high-speed multicolor stimulated Raman scattering. *Proc. Natl Acad. Sci. USA* **116**, 15842–15848 (2019).
55. Nitta, N. et al. Raman image-activated cell sorting. *Nat. Commun.* **11**, 3452 (2020).
56. Eek, K. M., Sessions, A. L. & Lies, D. P. Carbon-isotopic analysis of microbial cells sorted by flow cytometry. *Geobiology* **5**, 85–95 (2007).
57. Dyksma, S. et al. Ubiquitous Gammaproteobacteria dominate dark carbon fixation in coastal sediments. *ISME J* **10**, 1939–1953 (2016).
58. Ling, L., Zhou, F., Huang, L. & Li, Z.-Y. Optical forces on arbitrary shaped particles in optical tweezers. *J. Appl. Phys.* **108**, 073110 (2010).
59. Bonessi, D., Bonin, K. & Walker, T. Optical forces on particles of arbitrary shape and size. *J. Opt. A Pure Appl. Opt.* **9**, S228–S234 (2007).
60. Ashkin, A. Forces of a single-beam gradient laser trap on a dielectric sphere in the ray optics regime. *Biophys. J.* **61**, 569–582 (1992).

61. Novotny, L., Bian, R. X. & Xie, X. S. Theory of nanometric optical tweezers. *Phys. Rev. Lett.* **79**, 645–648 (1997).
62. Dholakia, K. & Reece, P. Optical micromanipulation takes hold. *Nano Today* **1**, 18–27 (2006).
63. Kim, S., Kang, I., Seo, J.-H. & Cho, J.-C. Culturing the ubiquitous freshwater actinobacterial acI lineage by supplying a biochemical 'helper' catalase. *ISME J* **13**, 2252–2263 (2019).
64. Li, T. et al. Simultaneous analysis of microbial identity and function using NanoSIMS. *Environ. Microbiol.* **10**, 580–588 (2008).
65. Huang, W. E., Griffiths, R. I., Thompson, I. P., Bailey, M. J. & Whiteley, A. S. Raman microscopic analysis of single microbial cells. *Anal. Chem.* **76**, 4452–4458 (2004).
66. McDonald, J. C. et al. Fabrication of microfluidic systems in poly(dimethylsiloxane). *Electrophoresis* **21**, 27–40 (2000).
67. Schuster, K. C., Reese, I., Urlaub, E., Gapes, J. R. & Lendl, B. Multidimensional information on the chemical composition of single bacterial cells by confocal Raman microspectroscopy. *Anal. Chem.* **72**, 5529–5534 (2000).
68. Dochow, S. et al. Quartz microfluidic chip for tumour cell identification by Raman spectroscopy in combination with optical traps. *Anal. Bioanal. Chem.* **405**, 2743–2746 (2013).
69. Kodinariya, T. M. & Makwana, P. R. Review on determining number of Cluster in K-Means Clustering. *Int. J. Adv. Res. Comput. Sci. Manag. Stud.* **1**, 90–95 (2013).
70. Bjerg, J. T. et al. Long-distance electron transport in individual, living cable bacteria. *Proc. Natl Acad. Sci. USA.* **115**, 5786–5791 (2018).
71. Zhao, J., Lui, H., McLean, D. I. & Zeng, H. Automated autofluorescence background subtraction algorithm for biomedical Raman spectroscopy. *Appl. Spectrosc.* **61**, 1225–1232 (2007).
72. Beier, B. D. & Berger, A. J. Method for automated background subtraction from Raman spectra containing known contaminants. *Analyst* **134**, 1198–1202 (2009).
73. Hehemann, J.-H. et al. Adaptive radiation by waves of gene transfer leads to fine-scale resource partitioning in marine microbes. *Nat. Commun.* **7**, 12860 (2016).
74. Taheri-Araghi, S. et al. Cell-size control and homeostasis in bacteria. *Curr. Biol.* **25**, 385–391 (2015).
75. Mazutis, L. et al. Single-cell analysis and sorting using droplet-based microfluidics. *Nat. Protoc.* **8**, 870–891 (2013).
76. Wang, Y. et al. Reverse and multiple stable isotope probing to study bacterial metabolism and interactions at the single cell level. *Anal. Chem.* **88**, 9443–9450 (2016).
77. Yuan, X. et al. Effect of laser irradiation on cell function and its implications in Raman spectroscopy. *Appl. Environ. Microbiol.* **84**, e02508–e02517 (2018).

Acknowledgements

We acknowledge support from a US Department of Energy Joint Genome Institute Emerging Technologies Opportunity grant (DE-AC02-05CH11231 to R.S. and M.W.). R.S. acknowledges support from a Gordon and Betty Moore Foundation Marine Microbial 1775 Initiative Investigator Award (GBMF3783), a Gordon and Betty Moore Symbiosis in Aquatic Systems Initiative Investigator Award 1776 (GBMF9197), a grant from the Simons Foundation (542395) as part of the Principles of Microbial Ecosystems (PriME) Collaborative and a grant (#315230_176189) from the Swiss National Science Foundation. F.C.P. and D.B. were supported by the Austrian Science Fund (FWF; P26127-B20 and P27831-B28) and the European Research Council (Starting Grant: FunKeyGut 741623). F.C.P. was also supported by the European Union's Horizon 2020 Framework Programme for Research and Innovation (grant no.658718). M.P. and M.W. were supported by the European Research Council via the Advanced Grant project NITRICARE 294343 and the FWF Wittgenstein award (to M.W.). L.B. was supported by grants from the Swedish Research Council (2019-04401) and the Science for Life Laboratory. We gratefully acknowledge funding from the European Molecular Biology Organization (EMBO; ALTF 1109-2016) and from the Human Frontier Science Program (HFSP; LT001209/2017) to U.A. We thank Horiba, Renishaw, and Bruker for providing their system specifications. We thank Cetoni for permission to display their software window. We thank R. Naisbit for scientific editing and M. Schmid for help with maintaining the Raman systems at the University of Vienna. We thank the M. Polz group, University of Vienna, Austria, for providing *Vibrio alginolyticus* 12G01 (wild type; NCBI:txid314288) and *Vibrio cyclitrophicus* 1F111 (wild type; NCBI:txid1136159). We thank S. Jun's group, University of California, San Diego, for providing *Escherichia coli* NCM3722 ΔmotA (non-motile mutant).

Author contributions

K.S.L., F.C.P., M.P., L.B., U.A. and D.B. designed the protocol and performed the experiments. R.S. and M.W. supervised the research. K.S.L. and R.S. wrote the manuscript. All authors have approved the final version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41596-020-00427-8.

Correspondence and requests for materials should be addressed to K.S.L., M.W. or R.S.

Peer review information *Nature Protocols* thanks Wei Huang and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 2 June 2020; Accepted: 25 September 2020;
Published online: 11 December 2020

Related links
Key references using this protocol
Lee, K. S. et al. *Nat. Microbiol.* **4**, 1035–1048 (2019): https://doi.org/10.1038/s41564-019-0394-9
Pereira, F. C. et al. *Nat. Commun.* **11**, 5104 (2020): https://doi.org/10.1038/s41467-020-18928-1

nature research

Corresponding author(s): Roman Stocker, Kang Soo Lee, and Michael Wagner

Last updated by author(s): Aug 25, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☒	☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Software developed: MATLAB GUI (graphical user interface) was used to build the sorting platform.
Data analysis	No software was used.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The datasets generated or analyzed in this protocol are available from the corresponding authors upon request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The sample size was constrained by cells analysed during the sorting. We used all samples measured during the sorting in our developed platform.
Data exclusions	No data were excluded from the analyses.
Replication	All experiments were conducted with at least three replicates. All attempts at replication were successful.
Randomization	The sample from the each culture tube was randomly collected and used for the sorting.
Blinding	Our sorting platform works in a fully automated manner. We were totally blinded to group allocation during data collection and/or analysis.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	For the mouse colon microbial community sample, the colon contents of a 6–8 week old C57BL/6J mouse was harvested in an anaerobic tent. Randomization and blinding are not applicable as animals were not distributed into groups.
Wild animals	The study did not involve wild animals.
Field-collected samples	The marine sediment samples were collected from the Pacific Ocean off the coast of Vancouver, Canada. The samples were maintained in the dark and at room temperature until further processing in the laboratory. A small spatula of sediment (0.5 g) was transferred to the vitamin-supplemented ASW medium (50 mL) supplemented by 1 mM $NaNO_2$ and incubated in the dark at 28 ^{o}C without agitation. Nitrite concentration was monitored at weekly intervals and when depleted, 10% volume of culture was transferred into fresh medium with 1 mM $NaNO_2$. Four sequential transfers were performed before Raman-activated cell sorting (RACS) to increase the initially low abundance of taxa of interest (cytochrome c-expressing cells) within the community.
Ethics oversight	The mouse colon specimens were collected under approval from the Institutional Ethics Committee of the University of Veterinary Medicine, Vienna, in accordance with Austrian laws (BMWF-66.006/0002-II/10b/2010).

Note that full information on the approval of the study protocol must also be provided in the manuscript.