nature protocols

https://doi.org/10.1038/s41596-025-01157-5

Protocol

Check for updates

Transcriptomic profiling of individual bacteria by MATQ-seq

Christina Homberger[1,5], Fabian Imdahl[2,3,5], Regan J. Hayward[3], Lars Barquist[2,3,4], Antoine-Emmanuel Saliba[1,2,3] & Jörg Vogel[1,2,3]

Abstract

Bacterial single-cell transcriptomics is revolutionizing our understanding of cell-to-cell variation within bacterial populations and enables gene expression profiling in complex microbial communities. Using the eukaryotic multiple annealing and dC-tailing-based quantitative single-cell RNA-sequencing (scRNA-seq) (MATQ-seq) approach, we have developed a robust bacterial scRNA-seq protocol, which integrates index sorting, random priming and rRNA depletion. This method stands out for its high rate of cell retention and its suitability for experiments with limited input material, offering a reliable method even for small sample sizes. Here we provide a step-by-step protocol covering the entire process of generating single-bacteria transcriptomes, including experimental and computational analysis. It involves (i) single-cell isolation via fluorescence-activated cell sorting (FACS) and cell lysis, (ii) reverse transcription and cDNA amplification using robotic liquid handling, (iii) rRNA depletion, (iv) indexing and sequencing, and (v) data processing steps to start comprehensive data analysis. Using model organisms such as *Salmonella enterica*, we show that the method achieves a retention rate of 95%, defined as the rate of initially sorted cells converted into effective sequencing libraries. This substantially surpasses other available protocols. The method robustly detects 300–600 genes per cell, highlighting its effectiveness in capturing a broad transcriptomic profile. The entire procedure from FACS-based single-cell isolation to raw data generation spans ~5 d. As MATQ-seq has already been proven robust in several bacterial species, it holds promise for the establishment of a streamlined microbial scRNA-seq platform.

Key points

- MATQ-sequencing-based bacterial single-cell RNA-sequencing enables high-resolution transcriptomic profiling of heterogeneous bacterial populations and is well-suited for experimental settings with low input material.
- The method offers high sensitivity and detection of a large number of genes per single bacterium, and successfully produces libraries for 95% of sorted cells.

Key references

Homberger, C., Hayward, R. J., Barquist, L. & Vogel, J. Improved bacterial single-cell RNA-Seq through automated MATQ-Seq and Cas9-based removal of rRNA reads. *MBio* **14**, e0355722 (2023): https://doi.org/10.1128/mbio.03557-22

Imdahl, F., Vafadarnejad, E., Homberger, C., Saliba, A.-E. & Vogel, J. Single-cell RNA-sequencing reports growth-condition-specific global transcriptomes of individual bacteria. *Nat. Microbiol.* **5**, 1202–1206 (2020): https://doi.org/10.1038/s41564-020-0774-1

[1]Institute of Molecular Infection Biology (IMIB), University of Würzburg, Würzburg, Germany. [2]Faculty of Medicine, University of Würzburg, Würzburg, Germany. [3]Helmholtz Institute for RNA-based Infection Research (HIRI), Helmholtz Centre for Infection Research (HZI), Würzburg, Germany. [4]Department of Biology, University of Toronto Mississauga, Mississauga, Ontario, Canada. [5]These authors contributed equally: Christina Homberger, Fabian Imdahl. e-mail: joerg.vogel@uni-wuerzburg.de

Introduction

Advances in single-cell RNA-sequencing (scRNA-seq) have enabled unbiased, high-throughput and high-resolution transcriptomic analysis of individual cells. While now routine for eukaryotic cells, application of scRNA-seq to bacteria is more challenging. Eukaryotic scRNA-seq methods cannot be easily adapted to bacteria because of the unique properties of bacterial RNA molecules, such as the absence of a poly(A) tail, the low quantities of RNA per cell (femtograms)[1] and the very short RNA half-life[2]. In addition, the presence of a robust cell envelope renders effective bacterial cell lysis difficult. Over the past years, these roadblocks have been overcome and several scRNA-seq protocols have become available[3–5]. Here, we detail our current scRNA-seq protocol, which comprises the capture of RNA molecules from FACS-sorted single bacteria, reverse transcription and cDNA amplification using the multiple annealing and dC-tailing-based quantitative scRNA-seq (MATQ-seq) workflow and rRNA depletion. The protocol is highly modular and, due to its high retention rate, it is applicable to profiling large bacterial communities as well as samples with very limited input material.

Development of the protocol

MATQ-seq was originally introduced by the laboratory of Chenghang Zong in 2017[6] as a eukaryotic scRNA-seq protocol designed to profile RNA molecules via random priming. We adapted the approach to profile the transcriptome of individual bacteria using model species such as *Salmonella enterica* and *Pseudomonas aeruginosa* and were able to define the transcriptomic adaptions of bacteria exposed to different stimuli. In the original paper[7] published in 2020, the method captured several hundred genes per bacterium, but the sequencing data were dominated by reads originating from rRNA. Therefore, in a follow-up study[8], we incorporated rRNA depletion using a Cas9-based cleavage protocol (depletion of abundant sequences by hybridization (DASH))[9,10], which allowed us to reduce the sequencing depth, driving down associated costs (Supplementary Table 1)[3]. To further enhance the robustness and sensitivity of our original bacterial MATQ-seq protocol, we also sought to improve it through the use of more effective enzymes, more efficient lysis and the implementation of automation[8]. The advanced protocol can reliably detect between 300 and 600 transcripts per single *S. enterica* cell, which represents a substantial improvement compared with the initial version of the protocol.

Overview of the procedure

Following culture, individual bacterial cells are first isolated by FACS and sorted directly into 96-well plates (Steps 1–8). After cell lysis, a sequential reverse transcription reaction is performed using MALBAC primers. These primers enable linear amplification and high efficiency through random priming[11]. They also have a unique sequence that induces looping of full amplicons, thereby reducing amplification bias by excluding full amplicons as templates for subsequent PCR reactions. Reverse transcription is followed by enzymatic digestion of primers and RNA. The single-stranded cDNA is then subjected to poly(C) tailing, which allows efficient and uniform cDNA capture using MALBAC primers for second-strand synthesis. cDNA amplification is carried out via PCR, followed by quality control (QC) (Steps 9–43). Libraries are generated through multiple steps that include DASH-based bacterial rRNA depletion before they are tagmented with unique Illumina indexes, pooled and sequenced (Steps 44–76). For data analysis, sequencing data are aligned to the respective reference genomes and single cells are associated to their individual indexes (Steps 77–91). A flowchart of the protocol and the library composition is provided in Fig. 1.

Advantages and limitations

The MATQ-seq protocol is unique in that it analyzes individual bacteria using a one-cell-per-well approach. It can easily be implemented with standard laboratory equipment and does not require prior fixation of bacterial cells. Its clear advantages are a high cell retention and a high

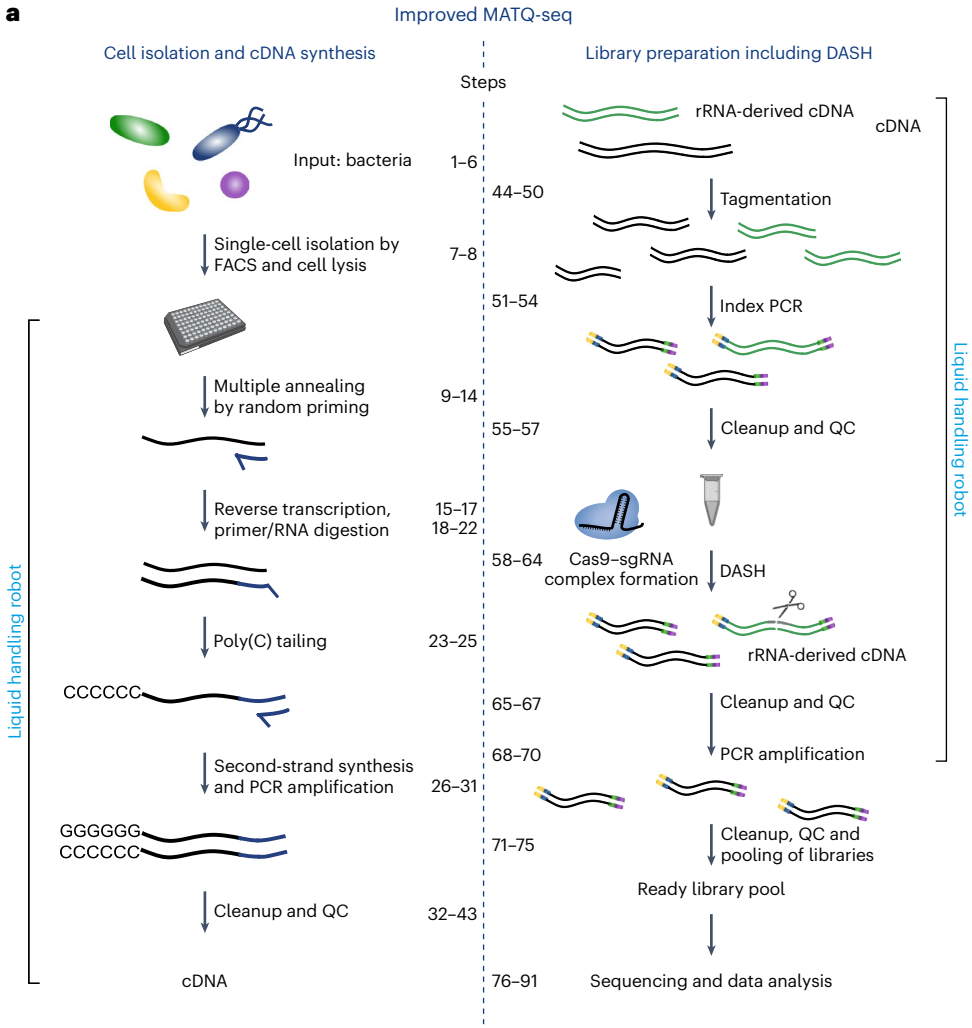

Fig. 1 | Overview of the MATQ-seq workflow. a, The workflow is structured into two main steps: cell isolation and cDNA synthesis (left) and library preparation including DASH for rRNA depletion (right). All pipetting steps can be automated using a liquid handling robot, except for all cleanup and QC steps. Adapted with permission from ref. 8, ASM. **b**, Illustration of the final MATQ-seq library using the Illumina Nextera XT library preparation kit. In addition to the region of interest containing the RNA-derived cDNA, the library includes Illumina and MATQ-seq specific sequences derived from dual indices, sequencing adapters and MALBAC primers. Sample indexing and demultiplexing is based on unique dual indexes implemented during the index PCR. The scheme illustrates sequencing in single-end mode for read 1, using 100 cycles.

number of detected genes per cell. It therefore offers an accessible approach that is an excellent choice for applications requiring a high resolution or for experimental settings with low input material. There are also drawbacks to the method. Due to the plate-based approach, MATQ-seq is

limited in throughput to the 100–1,000 cell range. Furthermore, isolation of individual bacteria is a prerequisite for the successful implementation of the protocol. Labour and costs per cell are higher compared with droplet-based approaches or split-sequencing-based methods, discussed further below (Supplementary Table 1)[3]. A disadvantage of DASH-based rRNA depletion is that it involves the design of tailored single-guide RNAs for each bacterial species.

Applications

MATQ-seq can readily be applied to profile heterogeneous transcriptional responses of isogenic bacteria to their environment or to monitor variations in host–pathogen interactions within bacterial populations. Owing to the high sensitivity and retention rate, MATQ-seq is uniquely suitable for experimental settings with limited input material such as intracellular bacteria or less abundant bacteria in tissues. Since the method relies on FACS to isolate individual bacterial cells, it can be coupled to index sorting. Additionally, markers such as antibodies targeting surface antigens and/or nucleic acid-binding dyes can be used to enrich for specific, potentially rare bacterial populations[12]. Bacteria can also be engineered to express fluorescent reporters to isolate bacterial populations based on their distinct fluorescence profiles, thus allowing multimodal profiling of bacteria[13]. So far, MATQ-seq has been applied to model organisms such as *S. enterica* and *P. aeruginosa*[7,8]. It should be applicable to other species as well, although some optimization might be required, for example, to achieve optimal bacterial lysis. The protocol includes several QC steps that can be used to adapt it to other bacterial species. Furthermore, the MATQ-seq workflow can be extended to a small-scale bulk approach with up to 1,000 bacteria per well.

Alternative methods

Since the initial reports of scRNA-seq in bacteria in 2020 and 2021[7,14,15], the number of protocols has increased with a focus on maximizing throughput and sensitivity. Here, we give a brief overview of these methods, but refer the reader to more comprehensive comparisons published elsewhere[3–5]. Several available protocols use a split-pool barcoding approach[16] to increase the throughput. Examples include microSPLiT[15] and PETRI-seq[14], which only require standard laboratory equipment. With multiple rounds of barcode ligations these methods are able to exponentially increase barcode combinations to individually label thousands of cells in a single experiment at a low cost per cell. Droplet-based approaches such as 10x Genomics or custom microfluidic systems have also been implemented, combining split-pool and droplet-based barcoding via microfluidics. These include smRandom-Seq[17], M3-seq[18] and BacDrop[19]. Microfluidic handling increases throughput by a factor of 10 compared with microSPLiT and PETRI-seq, allowing analysis of up to 100,000 cells. The high throughput nature of these methods comes with a high drop-out rate: for many individual bacterial transcriptomes none or very few RNA molecules are captured and most of the cells need to be discarded at the processing steps. Cell loss of up to 70% from the starting material due to several wash and ligation steps must be expected[3]. This limits the use of these methods to studies with sufficient starting material. In contrast to unbiased RNA molecule capture using random priming, ProBac-seq[20] uses a probe-based approach. This avoids rRNA priming and enhances specificity, but the need to design and synthesize a specific probe set for each bacterial strain to be analyzed makes this method less suitable for profiling complex bacterial communities.

In summary, when designing single-bacteria experiments, it is important to consider aspects such as available starting material, the required cell throughput and the desired resolution in terms of genes detected per cell, but also the investment in equipment. In contrast to methods aiming for high-throughput such as microSPLiT[15], M3-seq[18], BacDrop[19] and smRandom-seq[17], the MATQ-seq protocol described here offers high cell retention, versatility and adaptability without the need for custom equipment.

Experimental design

Here we describe an accessible end-to-end protocol for bacterial MATQ-seq (Fig. 1a), which involves the following steps:

Bacteria preparation (see 'Reagent setup—culturing of bacteria')
Bacteria can be isolated from tissues, biofilms or from cell cultures. It is recommended to add RNAlater before the cells are processed and transferred to the FACS to stop any transcriptomic changes.

FACS sorting (Steps 1–8)
Sorting of individual bacteria is challenging because of the necessity of reliably distinguishing micron-sized organisms from debris and dust. To sort the bacteria of interest, fluorescently labeled bacteria can be used to set the sorting gate (Fig. 2a–c). To achieve accurate sorting, FACS settings must be finely tuned to high sensitivity by setting high voltages for forward scatter (FSC) and side scatter (SSC) and by reducing the FSC threshold to 500–1,000.

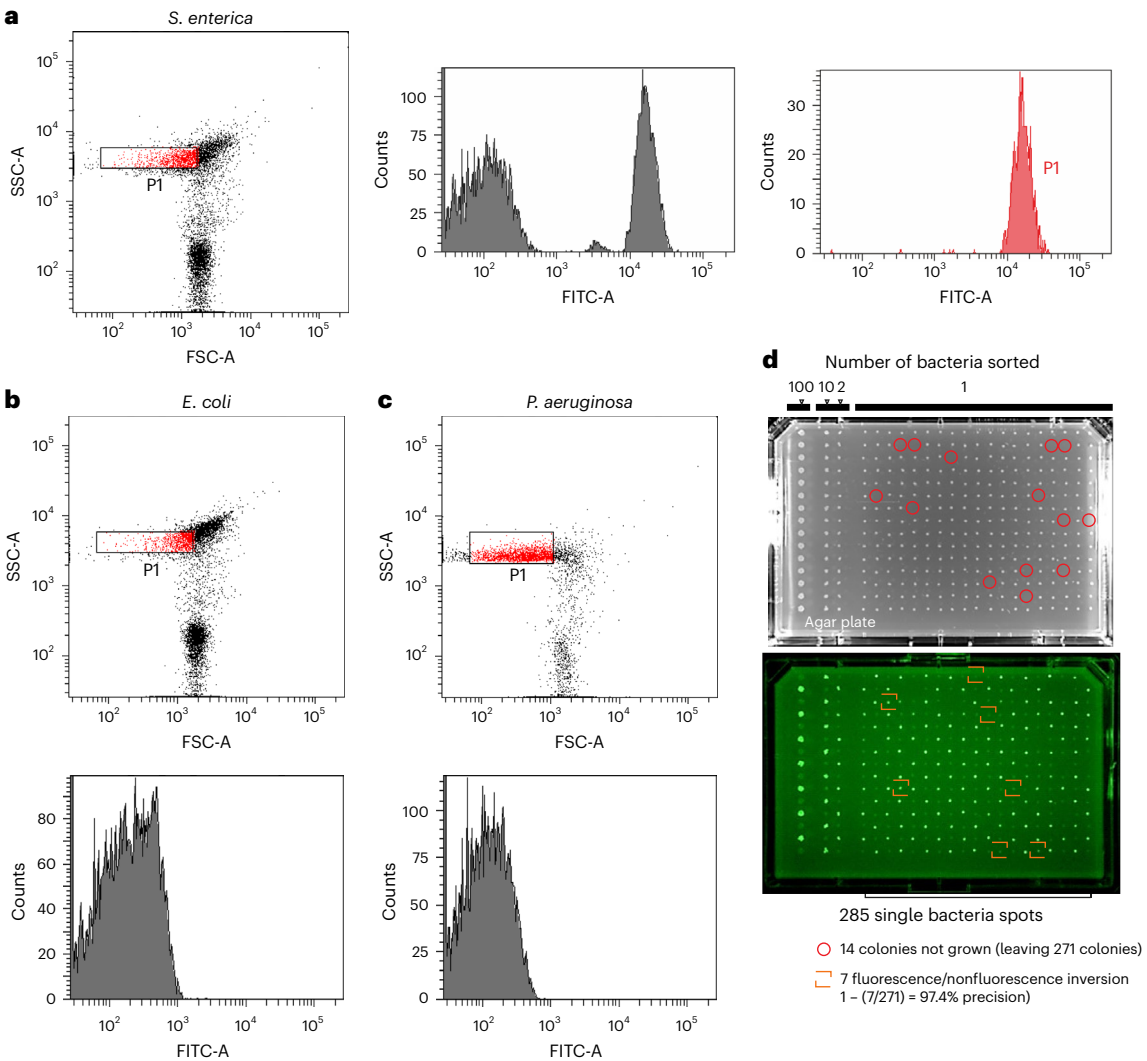

Fig. 2 | Single-cell isolation by FACS. a, Illustrative FACS plots demonstrating the gating strategy for *S. enterica* using a constitutively GFP-expressing SL1344 strain (P_{tet}::gfp integrated into the put locus in the chromosome)[29]. The sorting gate P1 shown on the FSC and SSC is set based on events with high emission in the FITC-A channel, highlighted in red on the count plot. The setup of the gating strategy can be conducted with any preferred fluorophore. Partially reprinted with permission from Homberger et al.[30] **b**,**c**, Examples of the gating strategy for *E. coli* (**b**) and *P. aeruginosa* (**c**). No changes in gating were required for *E. coli* compared to *S. enterica*, whereas adjustments were necessary for *P. aeruginosa* to account for variations in size and granularity. **d**, Assessment of the sorting efficiency for individual bacteria is conducted on LB agar moulded in a 384-well format. In addition to 100, 10 and 2 cells sorted on the left, 285 single GFP-fluorescent and nonfluorescent *Salmonella* were sorted systematically in an alternating pattern. Sorting precision was determined by quantifying the amount of colony mismatches detected under bright-field (red circles) and fluorescent (orange squares) imaging. Partially reprinted from Imdahl et al.[7], Springer Nature.

To validate our gating strategy and assess the reliability of the FACS, we conducted systematic sorting of 285 individual bacteria from a mixed culture of GFP-positive and wild-type (WT) bacteria onto an agarose plate in an alternating manner. The results show that 271 of the sorted cells grew single colonies and that there were 7 inversions between GFP and WT (Fig. 2d). This demonstrates a high sorting precision of 97.4%, confirming the effectiveness of our approach. We recommend performing similar tests for each new organism and gating strategy to ensure optimal accuracy and reliability in bacterial sorting.

In terms of sample input, there is no specific lower limit that is required for sorting. We have successfully used an input of only 100 bacteria and recovered ~60% after FACS. For higher sample input in the range of 100,000 cells, the recovery rate is even higher (73%). Thus, cell isolation by FACS is well suited for low input material and does not have a major impact on cell loss. However, low cell densities will result in an increase in total sorting time ranging from minutes to several hours, depending on cell density and the number of sorted cells. This needs to be taken into consideration when planning the experiment.

Lysis (see 'Reagent setup—lysis buffer')

Lysis of the bacterial cell envelope is essential for a successful experiment. This requires an efficient lysis buffer that is compatible with the downstream MATQ-seq workflow, that is, the buffer components even when diluted should not interfere with or inhibit the reverse transcription and amplification steps. The basis of the lysis buffer that we used is a formulated single-cell lysis buffer (Takara) supplemented with RNase inhibitor and PBS. For *S. enterica*, the addition of Ready lyse lysozyme was sufficient to achieve high lysis efficiency, determined by successful cDNA generation. In contrast to *S. enterica*, efficient lysis of *P. aeruginosa* required the addition of 0.5 mM EDTA and 10 mM DTT to the lysis buffer and an additional sonication step of 10 s (ref. 7).

In general, for Gram-negative species, it is advisable to vary the concentration of lysozyme and other chelating agents in addition to EDTA. However, a lysozyme concentration above 20 U has shown an adverse effect on the subsequent reverse transcription (C.H. and F.I., unpublished observation). For Gram-positive species, lysis is more challenging due to the need to disrupt the additional layers of peptidoglycan. Here, the use of freeze–thaw cycles after sorting has been shown to be effective for bulk lysis instead of sonication[21]. Another consideration is the replacement of lysozyme with other lysis enzymes, such as lysostaphin, or labiase or combining these enzymes with lysozyme. Since the lysis procedure varies between bacterial species, specific adaptations will be required for the species of choice. Two different lysis buffer formulations are listed in Table 1.

Table 1 | Lysis buffer formulations: can be complemented by mechanical rupture via sonication or freeze and thaw

Lysis buffer formulations	Reagent	Final concentration	Volume in µl for one reaction
Standard	10× lysis buffer		0.26
	Recombinant RNase Inhibitor (RRI)		0.03
	PBS (10×)	1×	0.26
	Lysozyme 50 U/µl	5 U	0.1
	Nuclease-free water		1.95
	Total		2.6
Modified with EDTA	10× lysis buffer	1×	0.26
	Recombinant RNase Inhibitor		0.03
	PBS (10×)	1×	0.26
	Lysozyme 50 U/µl	5 U	0.1
	DTT 1 M	100 mM	0.26
	EDTA 50 mM	0.5 mM	0.026
	Nuclease-free water		1.664
	Total		2.6

Reverse transcription, cDNA processing, amplification and QC (Steps 15–43)

Reverse transcription is one of the most critical steps in the protocol, as the efficiency of this reaction determines which transcripts are converted into cDNA for subsequent sequencing and analysis. To reduce amplification bias, we use MALBAC primers, which enable quasilinear amplification. Another crucial factor is the choice of reverse transcriptase (RT). During the optimization of this protocol, SuperScript IV has shown the best performance among the tested enzymes[8].

Further processing after reverse transcription includes the enzymatic digestion of primers and residual RNA by T4 DNA Polymerase and RNase H, respectively. For these reactions, very small volumes of viscous enzymes (100–200 nl) are added, requiring careful and slow pipetting. Inefficient digestion can negatively impact downstream processing.

After digestion, a poly(C) tailing step is performed, followed by second-strand synthesis and amplification of the cDNA product. It is important to note that QC can only be performed after PCR amplification, as the concentrations before this step are too low to be detected by Qubit or Bioanalyzer. To reduce costs per reaction, only a quarter of the product from the second-strand synthesis is used for the final amplification step in the MATQ-seq protocol. If PCR amplification does not work properly, the backup samples stored after single-strand synthesis can be used as input material to repeat the PCR.

Library preparation and sequencing (Steps 44–76)

Library preparation is based on the Illumina Nextera XT kit, which includes all necessary enzymes and buffers. Additionally, a tailored rRNA depletion step has been integrated, which is explained in more detail below and in Box 1.

For the library preparation, Unique Dual Index (UDI) primers distributed by Illumina are used. These are ready-to-use and provided in a 96-well plate format. Thus, it is recommended to perform the initial steps of library preparation, including optional pooling, in a 96-well plate. The sequences of the UDI primers required for demultiplexing are provided by Illumina. Ready pools are loaded and sequenced on an Illumina sequencing platform tailored to the number of samples.

Data processing and RNA quantification (Steps 77–91)

Following sequencing, reads must be quantified for further computational analysis. In principle, quantification of MATQ-seq is similar to quantifying any other RNA-seq experiment: adapter and primer sequences must be removed, reads mapped to the genomic sequence and then quantified by counting reads mapped to each genomic feature of interest (for example, mRNAs, small RNAs). However, special consideration should be given to QC given the low input RNA concentrations and complex structure of the resulting reads (Fig. 1b). We provide a protocol that includes multiple rounds of QC using FastQC and MultiQC[22] and adapter/primer sequence trimming using BBDuk, followed by read mapping using Bowtie2 (ref. 23) and quantification with featureCounts[24]. This produces a count matrix appropriate for further analysis using any of a variety of scRNA-seq analysis tools. We also provide code to produce an initial principal component analysis (PCA) plot based on the read counts, which can serve as both a QC measure to, for example, identify outliers, as well as a first step in the biological interpretation of the results.

Experimental setup and controls

Before FACS sorting, 96-well plates should be prepared with lysis buffer and stored on ice. If applying the protocol to a new species, it is recommended to initially assess the lysis buffer formulations and potential mechanical rupture by sorting several test plates using cultured bacteria (Table 1). Eight to sixteen wells can be used to test the lysis conditions. Pools of 100, 50 and 10 cells are systematically sorted down to the individual cell. For final plate sorting, it is highly recommended to include both positive and negative controls, that is, wells containing only lysis buffer and at least one well containing 50 pg of total RNA isolated from the bacterial species of interest, respectively. A random plate design can help to reduce bias in the final data analysis. Therefore, it is important to place negative and positive controls at different

BOX 1

Synthesis of guide RNA pool

● **TIMING 1.5 d**

The sgRNA pool is used for the DASH/library preparation protocol in Step 58

The synthesis of the sgRNA pool is excluded from the overall experimental timeline, as this step is a one-time process. The sgRNA pool can be used for multiple experiments

Dilute DASH oligo pool and fill-in-reaction oligonucleotide in nuclease-free water each to a final concentration of 10 µM.
To make the mastermix for PCR, mix 25 µl of KAPA HiFi Hot Start Ready Mix with 5 µl of oligo pool from the previous step, 10 µl of fill-in-reaction oligonucleotides from the previous step and 10 µl nuclease-free water in a PCR tube.
Include a negative control for PCR by omitting the fill-in-reaction oligonucleotide and adjusting the reaction volume to 50 µl with nuclease-free water.
▲ **CRITICAL STEP** To enhance the yield of the sgRNA pool for use in subsequent experiments, conduct multiple fill-in reactions simultaneously. This depends on the desired throughput and usage and is highly user-dependent. We recommend performing two to eight reactions in parallel. This is still easy to handle with minimal additional hands-on time compared to only one fill-in reaction.

Mix well by pipetting and briefly spin down.
Incubate PCR reactions in a thermal cycler using the following fill-in-reaction program:

Cycle number	Preincubation	Annealing	Filling	Hold
1	95 °C, 3 min			
2		60 °C, 30 s		
3			72 °C, 1 min	
4				10 °C, ∞

To perform a column purification using Zymo Oligo Clean & Concentrator kit, optionally pool up to two PCR reactions for the cleanup.
▲ **CRITICAL STEP** The Zymo Oligo Clean & Concentrator kit can be used for double-stranded DNA as well as RNA according to the manufacturer's recommendations and is not restricted to single-stranded DNA.

Insert the column into a collection tube and add PCR reaction to the column.
Centrifuge for 30 s at 10,000*g* and discard the flow through.
Add 750 µl of DNA wash buffer to the column and centrifuge for 30 s at 10,000*g*. Discard the flow through.
Centrifuge for another 1 min at max speed to get rid of the residual buffer.
Transfer the column to a microcentrifuge tube and add 15 µl nuclease-free water.
Centrifuge for 1 min at 13,000*g* to elute DNA.

▲ **CRITICAL STEP** The DNA pool should suffice for multiple in vitro transcription (IVT) reactions and can be stored long-term. For our experiments the DNA was sufficient for at least 2 and a maximum of 5 reactions.
■ **PAUSE POINT** Store DNA at −20 °C for up to 1 year or proceed with QC.

QC: perform the Qubit dsDNA HS assay of PCR products to determine concentration and check size by using the Bioanalyzer DNA 1000 kit by following manufacturer's instructions. The main peak should be in the range of ~145 bp.
◆ **TROUBLESHOOTING**
For the IVT using the MEGAshortscript kit, prepare a premix of 2 µl T7 10× reaction buffer, 8 µl NTPs (mix 2 µl each of the rNTPs provided in the kit) and 2 µl T7 enzyme mix per reaction.
Add a total of 300 ng DNA from the fill-in reaction to the premix and adjust the reaction volume to 20 µl with nuclease-free water.
Mix well by pipetting and briefly spin down.
Incubate IVT reaction at 37 °C for 4 h, then hold at 4 °C and procced to the next step.
Remove the DNA template by adding 1 µl of TURBO DNase included in the MEGAshortscript kit and incubate at 37 °C for 30 min, then hold at 4 °C and proceed to the next step.
To perform RNA clean up using the Monarch Spin RNA Cleanup kit, insert the columns into a collection tube and transfer the IVT product onto the column.
Centrifuge for 1 min at 10,000*g* and discard the flow through.
Add 500 µl RNA cleanup wash buffer to the column and centrifuge for 1 min at 10,000*g* before discarding the flow through.
Centrifuge for another 1 min at max. speed to get rid of the residual buffer.
Transfer the column to a microcentrifuge tube and add 50 µl nuclease-free water. Incubate columns for 5 min at room temperature.
Centrifuge for 1 min at 13,000*g* to elute RNA.

QC: Perform Qubit dsRNA HS Assay of sgRNA pool to determine concentration and check size by using Bioanalyzer RNA Pico kit following the manufacturer's instructions. The main peak is expected at 155–160 bp (Fig. 4).
◆ **TROUBLESHOOTING**
Put tubes on ice, prepare aliquots. Store the sgRNA pool immediately at −80 °C.
■ **PAUSE POINT** When stored in the correct conditions, sgRNAs can be kept at −80 °C for a maximum of 1 year.
▲ **CRITICAL STEP** It is advisable to combine sgRNA pools from multiple reactions post-QC. Aliquoting and storing them afterwards as described above ensures consistency in quality and concentration for all subsequent DASH reactions.

positions in the 96-well plate and to include both controls in every plate. If contaminations are observed, replicates of controls per plate are recommended. After FACS sorting, plates must be immediately sealed with proper microseals (Microseal 'F') that withstand low temperature (−80 °C). Sorted plates can be stored for several months at −80 °C.

In general, negative and positive controls should be included as early as possible, from FACS sorting into lysis buffer to processing and cDNA preparation. During library preparation, a negative control without cDNA should be included to ensure sterility and to rule out potential cross-contaminations. For sequencing, libraries of 10 or 100 cells can serve as controls and are used for QC. A sample pool is spiked with 1% PhiX control library. For sgRNA synthesis it is recommended to include a negative control without DASH fill-in reaction primer to eliminate potential DNA contaminations.

Protocol automation and avoiding contamination

The MATQ-seq protocol can be performed either manually or in an automated fashion using a liquid handling robot (for example, I.DOT, Dispendix). The use of a liquid handler not only saves time and increases throughput, but also reduces pipetting biases. Here, we have used the I.DOT because it is a non-contact liquid handler, which eliminates sample loss and allows dispensing volumes in the nanoliter range. If the protocol is performed manually, only standard laboratory equipment is required with the exception of a FACS sorter. It is strongly recommended to use a PCR workstation for all steps before PCR amplification. Sample processing does not require specialized skills, but care must be taken to avoid contamination. All surfaces, including pipettes, should be cleaned with RNaseZap and 70% ethanol before each experiment. In addition, dedicated PCR cyclers for the pre- and postamplification steps should be used and all pipetting steps need to be carried out with filter tips. All reagents are aliquoted in PCR-grade tubes to minimize the risk of contamination.

rRNA depletion (Steps 58–71 and Box 1)

rRNA depletion is not mandatory but greatly reduces the cost of sequencing and enables more efficient sample processing. We have integrated rRNA depletion via DASH[9,10] into the MATQ-seq workflow[8], removing noninformative rRNA molecules at the level of cDNA. DASH relies on Cas9-mediated cleavage guided by a pool of sgRNAs. The sgRNAs are designed to bind to complementary rRNA regions adjacent to a protospacer adjacent motif (PAM) 'NGG', which is required for Cas9 activity. Cas9 then forms a complex with the sgRNAs and produces double-strand breaks at their target regions. In this manner, DASH allows the selective removal of rRNA with high sequence specificity and is directly integrated into library preparation using the Nextera XT kit. The synthesis of the sgRNA pool intended for the DASH reaction is outlined in Box 1. It is performed independently of the actual scRNA-seq experiment.

By using DASH, we achieved a reduction of rRNA reads to an average of 75% of total reads[8]. RNase H has also been successfully used for rRNA depletion in bacterial scRNA-seq pipelines. M3-seq and BacDrop use an RNase H-mediated depletion protocol applied either to amplified libraries or directly after cell permeabilization[18,19]. Both RNase H-based methods allow a reduction of rRNA reads to 35–40% of total reads. The incorporation of rRNA depletion into the bacterial scRNA-seq pipeline is a considerable improvement over previously developed protocols[7,14,15].

Materials

Biological materials

- *S. enterica* Typhimurium strain SL1344 (DSMZ Collection of Microorganisms and Cell Culture, DSM 24522)
- *P. aeruginosa* strain PA01 (DSMZ Collection of Microorganisms and Cell Culture, DSM 19880)

Reagents

- 10× Lysis buffer (Takara, cat. no. 635013)
  ▲ **CRITICAL** (proprietary composition suitable for single-cell lysis)
- AMPure XP beads (Beckman Coulter, cat. no. A63881)
- Cas9 Nuclease (NEB, cat. no. M0386M)
- dCTP (Invitrogen, cat. no. 10217016)
- Deep Vent (exo-) DNA Polymerase (NEB, cat. no. M0259L)
  ▲ **CRITICAL** (no suitable alternative)
- DNA 1000 Kit (Agilent Technologies, cat. no. 5067-1504)
- DNA Polymerase I (NEB, cat. no. M0209S)
- dNTPs (Thermo Fisher Scientific, cat. no. R0182)
- DTT (Invitrogen, cat. no. D1532)
- EDTA (Thermo Fisher Scientific, cat. no. 15575020)
  ▲ **CAUTION** Harmful if inhaled and can cause severe skin burns and eye damage. Wear protective gloves and eye protection.
- High Sensitivity DNA kit (Agilent Technologies, cat. no. 5067-4626)
- Illumina DNA/RNA UDI sets A–D, Tagmentation (v3) (Illumina, cat. nos. 20091654, 20091656, 20091658, 20091660)—formerly called IDT for Illumina DNA/RNA UDI sets A–D, Tagmentation (v2) (Illumina, cat. nos. 20027213, 20027214, 20042666, 20042667)
  ▲ **CRITICAL** (guarantees high quality and compatibility with library preparation kit)
- KAPA HiFi HotStart ReadyMix (Roche, cat. no. 07958927001)
- MEGA Shortscript T7 Transcription kit (Invitrogen, cat. no. AM1354)
- Monarch Spin RNA Cleanup kit (500ug) (NEB, cat. no. T2050S)
- Nextera XT DNA Library Preparation kit (Illumina, cat. no. FC-131-1096)
  ▲ **CRITICAL** (no alternative available)
- NSQ500/2000; NovaSeq Sequencing kits (Illumina, various)
- Nuclease-free water (Ambion, cat. no. AM9937)
- Oligo Clean and Concentrator kit (Zymo Research, cat. no. D4060)
- PhiX Control Library (Illumina, FC-110-3001 or -3002)
- PMSF (Carl Roth, cat. no. 6367.1)
  ▲ **CAUTION** Toxic if swallowed and causes severe skin burns and eye damage. Wear protective gloves, protective clothing and eye protection.
- Proteinase K (NEB, cat. no. P8107S)
  ▲ **CAUTION** May cause allergic skin reaction and breathing difficulties if inhaled. Wear protective gloves and eye protection.
- Qubit 1× dsDNA High Sensitivity (Invitrogen, cat. no. Q33231)
- Qubit RNA High Sensitivity (Invitrogen, cat. no. Q32855)
- Ready lyse lysozyme solution (Epicentre, cat. no. R1804M)
  ▲ **CRITICAL** (no suitable alternative)
- Recombinant RNase Inhibitor (Takara, cat. no. 2313A)
- RNA 6000 Nano Kit (Agilent Technologies, cat. no. 5067-1511)
- RNA 6000 Pico Kit (Agilent Technologies, cat. no. 5067-1513)
- RNAlater/RNAprotect Tissue Reagent (Qiagen, cat. no. 76104)
- RNase H (NEB, cat. no. M0297L)
- RNase If (NEB, cat. no. M0243S)
- RNaseZap RNase decontamination solution (Invitrogen, cat. no. AM9780)
- SuperScript IV (Invitrogen, cat. no. 18090050)
  ▲ **CRITICAL** (exhibited best performance in comparative experiments[8])
- T4 DNA Polymerase (NEB, cat. no. M0203S)
- Terminal Transferase (NEB, cat. no. M0315S)

Primers and oligonucleotides

▲ **CRITICAL** All oligonucleotides are resuspended in nuclease-free water at a concentration of 100 µM. It is recommended to use oligonucleotides from IDT (desalted quality or higher) because oligonucleotides from other suppliers have been observed to impact yield adversely.

- Forward primer for DASH amplification: 5′-AATGATACGGCGACCACCGAGAT-3′ (Illumina)
- Reverse primer for DASH amplification: 5′-CAAGCAGAAGACGGCATACGA-3′ (Illumina)
- GAT27dt: 5′-GTGAGTGATGGTTGAGGATGTGTGG AGNNNNNTTTTTTTTTTTTTTTTTT-3′ (ref. 6)
- GAT27 5N3G: 5′-GTGAGTGATGGTTGAGGATGTGTGGAGNNNNNGGG-3′ (ref. 6)
- GAT27 5N3T: 5′-GTGAGTGATGGTTGAGGATGTGTGGAGNNNNNTTT-3′ (ref. 6)
- GAT21 6N3G: 5′-GATGGTTGAGGATGTGTGGAGNNNNNNGGG-3′ (ref. 6)
- GAT27 PCR: 5′-GTGAGTGATGGTTGAGGATGTGTGGAG-3′ (ref. 6)
- DASH oligo pool, pool of 797 oligonucleotides (ref. 10)
- DASH fill-in-reaction 5′-GGCATACTCTGCGACATCGT-3′ (ref. 10)
- Illumina DNA/RNA UDI sets A–D, Tagmentation

Reagent setup

Culturing of bacteria

Bacterial culture preparation for scRNA-seq is variable and depends on the experimental setup. This includes the organism of choice, the specific growth conditions and the desired timepoint for collecting the sample. Therefore, culture preparation time can vary from 2 to 10 h or longer (for example, for anaerobic species).

▲ **CRITICAL** The time needed for bacterial culture preparation is an important aspect to consider when designing the experimental plan to align smoothly with single cell isolation by FACS. For optimal sample processing it is recommended to start with the preparation of the lysis buffer plates (Steps 1 and 2) ~1–2 h before the bacterial cultures are ready to be processed.

Lysis buffer

The lysis buffer contains 0.26 µl of Takara 10× Lysis buffer, 0.03 µl Recombinant RNase Inhibitor, 0.1 µl lysozyme (50 U/µl), 0.26 µl 10× PBS and 1.95 µl nuclease-free water. Lysis buffer should always be freshly prepared on the day of FACS sorting. Another lysis buffer formulation can be found in Table 1. If the standard lysis buffer formulation is not effective for the organism of choice, the addition of EDTA and DTT can further enhance the lysis by chelating metal ions and reducing disulfides. The modified lysis buffer combined with sonication has shown high efficiency for *P. aeruginosa*[7].

RT mix

Prepare the RT–primer mix containing 0.006 µl GAT27 dt, 0.02 µl GAT27 5N3G, 0.02 µl 5N3T and 0.354 µl nuclease-free H_2O. It is advisable to prepare a mix for at least 20 reactions for ease of pipetting. The primer mix should always be freshly prepared.

dNTP mix

Prepare a dNTP mix by adding the individual dNTPs supplied in an equimolar ratio to a final concentration of 10 mM for each dNTP. The dNTP mix can be stored at −20 °C for up to 1 year.

RNA digestion mix

For manual pipetting, it is advisable to prepare an RNA digestion mix containing an equal volume of RNase H and RNase If (0.1 µl each per reaction) in a larger volume (for example, 10 µl each, sufficient for 96 reactions). It can be stored at −20 °C for up to 1 month. This simplifies pipetting of small amounts of the viscous enzyme mix.

▲ **CRITICAL** This mix can be omitted when using the I.DOT.

Equipment

- 48-well PCR plates, PCR plate no-skirt (Brand, cat. no. BR781415)
  ▲ **CRITICAL** When using 48-well plates, it is advisable to skip the outer columns due to imperfect sealing in those areas that may lead to increased evaporation. Use Microseal ‘F’ for steps including freezing and Microseal ‘B’ for all other steps.
- 96-well PCR plates PCR plate non-skirted (Brand, cat. no. 781368)
- Bioanalyzer 2100 Bioanalyzer (Agilent)
- FACS Aria III (BD Biosciences)
- I.DOT liquid handler (Dispendix, only for automated workflow)

- I.DOT well plates, Pure Plate 100 (Dispendix, only if using I.DOT liquid handler, cat. no. D16110021801)
- I.DOT well plates, Pure Plate 200 (Dispendix, only if using I.DOT liquid handler, cat. no. D16110021807)
- Microseal 'B' (Bio-Rad, cat. no. MSB-1001)
- Microseal 'F' (Bio-Rad, cat. no. MSF-1001)
- Qubit flex (Invitrogen, cat. no. Q33327)
- Qubit assay cap stripes MicroAmp optical 8-tube cap (Life Technologies, cat. no. 4323032)
- Qubit assay stripes MicroAmp optical 8-tube strip (Life Technologies, cat. no. 4316567)
- Sequencer NextSeq 500, NextSeq 2000, NovaSeq 6000 (Illumina)
- Vortex Bioanalyzer MS 3 (IKA)
- Local computer or a computing cluster for bioinformatic analysis
  ▲ **CRITICAL** Bioinformatic analysis can be performed on either a local computer or a computing cluster. The main differences lie in the hardware and job scheduling software used by clusters to effectively distribute multiple computational tasks. To reduce task execution time, one can increase the number of CPU cores available for certain processes. CPU cores are responsible for executing instructions and each core may be capable of handling multiple threads concurrently, generally two per core on current consumer hardware. The amount of RAM and the number of threads or cores to use depends on the available system resources. For instance, a typical laptop may have 4–8 cores and 8–16 GB RAM, while a cluster server might have more than 24 cores and over 16 GB RAM.

Equipment Setup

I.DOT liquid handler

The dispensing robot requires the setup of individual liquid classes for different reagents. The glycerol 50% liquid class is used for the following reagents: proteinase K, Cas9 nuclease, Deep Vent (exo-) DNA Polymerase, RNase H, RNase If, Recombinant RNase Inhibitor, SuperScript IV, T4 DNA Polymerase, Terminal Transferase.

Individual liquid classes with different pressure settings due to varying viscosity properties were necessary for the following reagents: Cas9 buffer, Thermopol buffer, Terminal Transferase buffer. These additional liquid classes were set up according to the manufacturer's instruction manual.

All remaining liquids were dispensed using the pre-installed water liquid class.

▲ **CRITICAL** If using the I.DOT dispensing robot, its source well plates can be effectively reused up to five times. Between MATQ-seq runs. plates must be stored at −20 °C. However, after five cycles, it is recommended to exchange all used wells of the I.DOT plates to ensure optimal performance and reliability. For streamlined handling of I.DOT plates and reagents, it is advisable to dedicate a single I.DOT plate for each pipetting step.

FACS setup

The recommended nozzle size for bacterial cell isolation is 70 µm. The sorting device should also be equipped with an integrated sample cooling system that allows temperature control of ~4 °C throughout processing. Note that FSC and SSC voltages must be set high to detect bacteria. With our setup, we used a voltage of 566 for the FSC and 374 for the SSC. For the discrimination of bacteria versus debris or dust, it is recommended to use fluorescently labeled bacteria for gating (Fig. 2a–c). It is also advisable to use a coupled threshold of 500 for FSC and SSC. It is very important to dilute the FACS sample initially stored in RNAlater in 1× PBS and not in RNAlater right before sorting.

▲ **CRITICAL** To achieve an event rate of 1,000–1,500 events per second, a standard dilution of 1:20 (for example, 50 µl culture in 950 µl 1× PBS) was used for cells collected at an OD of 0.3–0.5 and further adjusted as required.

Thermal cyclers

All thermal cycler programs were carried out with a lid temperature of 105 °C.

Centrifugation

All centrifugation steps were carried out at room temperature (20–25 °C).

Procedure

▲ **CRITICAL** This procedure is optimized for automation using I.DOT and lists the respective volumes of reagents required for each reaction. Any other type of liquid handling equipment or manual pipetting can be used as well. If pipetting is performed manually, a master mix for the individual steps should be prepared in advance for the target number of wells +10% access volume.

▲ **CRITICAL** Single-cell sequencing requires cleanliness and extra care throughout the entire experimental process. The use of filter tips and gloves to minimize the risk of contamination is necessary. Surfaces should be cleaned before and after work using ethanol and RNase zap. Ideally, a PCR hood should be used to prevent contamination. Adhering to these guidelines is essential for obtaining accurate and reliable results for bacterial scRNA-seq.

Cell preparation for FACS analysis and single-cell isolation (day 1)

● **TIMING 4–6 h (culture preparation is not included)**

1. Prepare a 48- or 96-well plate with 2.6 µl of lysis buffer per well using a pipette or a liquid handling robot.
2. Seal plates with a Microseal 'B' (Bio-Rad) and quickly spin them down for 5 s in a plate centrifuge to make sure that there are no bubbles. Do not forget to add wells for a positive and a negative control (wells prefilled with lysis buffer that will not receive a cell (in Step 7)). Lysis buffer plates can be stored on ice until they are used.
   ▲ **CRITICAL STEP** Efficient cell lysis is critical for the success of the MATQ-seq protocol. The best lysis strategy might be species dependent and requires optimization (see 'Experimental design').
3. Take 1 ml of your bacterial culture (see 'Reagent setup') and centrifuge it for 4 min at room temperature at 17,000*g*. Discard the supernatant and resuspend the bacterial pellet in 1 ml 1× PBS.
4. Repeat the centrifugation and resuspend again in 1 ml 1× PBS.
5. Repeat the centrifugation and resuspend in 1 ml RNAlater–PBS mix (1:1 vol/vol).
6. Store samples on ice until FACS sorting.
   ▲ **CRITICAL STEP** Washing in 1× PBS can lead to alterations in the transcriptome, and recently published data show that bacterial mRNA has a much shorter half-life than previously thought[2]. To overcome this, direct resuspension and washing in RNAlater is an alternative. For low input samples, washing steps (Steps 3 and 4) with 1× PBS or RNAlater can be omitted, allowing direct resuspension in RNAlater or even direct sorting from a planktonic culture to avoid sample loss. However, working with very low input scales ($<10^3$ cells) may increase the overall sorting time. It is important to note that storing samples on ice may affect the transcriptome and alter the distribution of subpopulations. To account for this, the FACS instrument can be pre-set to allow immediate sorting after sample resuspension.
7. FACS sort your cells using a BD Aria III and the 70 µm nozzle. It is recommended to include negative and positive controls on each plate (in Step 12). In the BD FACSDiva software, set the sorting precision to 'single cell'. Dilute your sample in 1× PBS to achieve a reasonable event rate (1,000–1,500 events/s). For sorting, the well plate prepared with lysis buffer in Step 2 is unsealed and placed in the plate holder of the FACS. To assess the accuracy of your sorting, you can sort single cells on a freshly prepared agarose dish, incubate it overnight and judge the precision of your bacterial sorting procedure (Fig. 2d).
   ▲ **CRITICAL STEP** This sample preparation protocol using RNAlater as preservation reagent for FACS samples has previously been successfully evaluated for RNA-seq[25]. It is very important to dilute the FACS sample initially stored in RNAlater in 1× PBS and not in RNAlater right before sorting. A high concentration of RNAlater will interfere with the fluorescence signals used in FACS and may cause clogging or malfunction due to its viscosity. RNAlater is only used as preservation reagent after cell collecting and before the actual sorting procedure.

8. Immediately after sorting, seal the plate (Microseal 'F') and spin it down to make sure all droplets find their way to the lysis buffer. Keep plates on ice and freeze them at −80 °C.
   ▲ **CRITICAL STEP** Some bacteria, especially those that are already difficult to lyse in bulk approaches (for example, Gram-positive bacteria) might require additional mechanical lysis. Plates can be sonicated in a sonication bath for 10 s (ref. 7) or be suspended in liquid nitrogen for freeze and thaw cycles. A combination of both is also possible and has proven to be very effective. After treatment, keep plates on ice and freeze them at −80 °C.
   ■ **PAUSE POINT** Sorted bacteria in plates can be kept at −80 °C for several months.

MATQ-seq protocol automated with I.DOT (day 2)

● **TIMING total 8–10 h (without sample purification. TIMING depends on the number of samples and if the protocol is carried out automated or by manual pipetting; the use of a dispensing robot drastically reduces the required hands-on time from ~6 h to 3 h and increases sample throughput. For manual pipetting it is not recommended to process more than 24 samples at once; the automated protocol can handle 96 samples or even 384 at once.)**

▲ **CRITICAL** It is of high importance to promptly transfer and store samples on ice once the PCR program is completed, and to perform all pipetting steps on a cooling block unless otherwise specified in the protocol. Extremely clean work is a requirement for the success of the protocol. Neglecting a high standard of cleanliness can compromise sample integrity, emphasizing the critical need for immediate and proper sample handling. Bear in mind that the individual wells contain a femtogram amount of RNA that is lost at the slightest contamination.

Reaction preparation

● **TIMING 45 min to 1 h including incubation times**

9. Take sorted plates out of the −80 °C and let them thaw.
   ▲ **CRITICAL STEP** All following steps until the PCR cleanup should be performed on a clean PCR workbench.
10. Preheat the thermal cycler to 72 °C.
11. Mix 0.4 µl of the RT–primer mix with 0.05 µl DTT and 0.12 µl dNTP mix to create the pre-RT mix.
12. Add 0.5 µl of 100 pg/µl total RNA of the desired organism as a positive control into the intended well. Add 0.5 µl of nuclease-free water into the intended well for the negative control.
13. Use the liquid handling robot to dispense 570 nl of the pre-RT mix into each well. Seal your plate and spin it down.
14. Incubate the plate in the thermal cycler for 3 min at 72 °C. Place your plate on ice for at least 1 min before dispensing the RT reagents.

Reverse transcription

● **TIMING 2–2 h 15 min, including PCR program**

15. Distribute 0.8 µl SSIV buffer, 0.15 µl SSIV enzyme, 0.1 µl RNase Inhibitor, 0.2 µl DTT and 1.15 µl nuclease-free water to each well, using the liquid handling robot.
    ▲ **CRITICAL STEP** If using the I.DOT, keep the source wells containing enzymes at −20 °C until usage.
16. Seal plate and spin it down.
17. Place the plate in the thermal cycler and start the following RT program:

Cycle number	Primer annealing	Elongation	Looping	Hold
1–13	8 °C, 12 s			
	15 °C, 45 s			
	20 °C, 45 s			
	30 °C, 30 s			
		42 °C, 2 min		
			50 °C, 3 min	
14			50 °C, 3 min	
15				4 °C, ∞

Primer and RNA digestion

● **TIMING 3–3.30 h including PCR programs**

18. Remove plate from cycler at 50 °C and add 0.2 µl of T4 DNA polymerase at room temperature using the liquid handling robot. Alternatively, if the cycler has already cooled down to 4 °C, preheat the plate to 50 °C for 1 min before adding 0.2 µl of T4 DNA polymerase at room temperature.
19. Seal the plate and spin it down.
20. Place the plate in the thermal cycler and start the following primer digestion program:

Cycle number	Incubation	Inactivation	Hold
1	37 °C, 40 min		
2		75 °C, 20 min	
3	37 °C, 40 min		
4		75 °C, 20 min	
5			4 °C, ∞

21. Add 0.1 µl RNase H and 0.1 µl RNase If to each well using the liquid handling robot.
22. Place the plate in the thermal cycler and incubate for 15 min at 37 °C and 15 min at 75 °C. Then keep the temperature at 4 °C.

Poly(C) tailing

● **TIMING 45 min to 1 h including PCR programs**

23. Using the liquid handling robot, add 0.1 µl of TdT, 0.4 µl of TdT buffer and 0.4 µl of dCTP and 3.13 µl nuclease-free water to each well.
24. Seal the plate and spin it down.
25. Place the plate in the thermal cycler and incubate for 15 min at 37 °C and 15 min at 75 °C. Then keep the temperature at 4 °C.

Second-strand synthesis

● **TIMING 30–45 min including PCR programs**

26. To synthesize the second strand, dispense 0.125 µl dNTP (mix of 10 mM each), 0.125 µl GAT21 6N3G primer and 12.925 µl nuclease-free water. Using the I.DOT, switch to S200 I.DOT well plates to dispense 1.5 µl Thermopol buffer. These wells have larger pores, which reduce foam formation during the dispensing process. Since the Thermopol buffer contains a proprietary detergent component, using larger pores with the I.DOT is highly recommended.
27. Seal, spin down and incubate the plate for 1 min at 95 °C, then cool down to 48 °C.
28. Take the plate out and immediately add 0.4 µl Deep Vent (exo-) DNA Polymerase at room temperature.
29. Incubate using the following second-strand synthesis PCR program:

Cycle number	Denaturation	Annealing	Elongation	Hold
1	98 °C			
2		48 °C		Add 0.4 µl Deep Vent (exo-)
3		48 °C, 20 s	72 °C, 1 min	
4			72 °C, 2 min	
5				4 °C, ∞

■ **PAUSE POINT** Store samples at −20 °C or directly continue with PCR amplification. Plates can be stored for several months at −20 °C.

PCR amplification

● **TIMING 2–2.5 h including PCR programs**

30. Transfer 6.55 µl of the cDNA obtained in the previous step to a new well plate and place on ice. The remaining cDNA is kept at −20 °C as a backup. It can be stored for at least 1 month.
31. For PCR, add 28.5 µl nuclease-free water, 0.75 µl dNTP (mix of 10 mM each), 0.75 µl Deep Vent Polymerase, 0.2 µl GAT 27 PCR primer to the newly prepared plate containing 6.55 µl cDNA from the previous step. Using the I.DOT, switch to S200 I.DOT well plates to dispense 3.25 µl Thermopol buffer.
32. Seal, spin down and incubate using the following MATQ-seq PCR amplification program:

Cycle number	Denaturation	Annealing	Elongation	Hold
1	95 °C, 30 s			
2–26	95 °C, 15 s			
		62 °C, 20 s		
			72 °C, 2 min	
27			72 °C, 5 min	
28				4 °C, ∞

■ **PAUSE POINT** Store samples at −20 °C or directly continue with cleanup. Samples should not be stored for more than 1 week at −20 °C.

Purification (day 3)

● **TIMING 1–2 h (highly depends on number of samples)**

33. Store AMPure XP magnetic beads at room temperature for at least 30 min and vortex well before usage.
34. Add 40 µl of beads to each sample (1:1 vol/vol ratio). Mix well by pipetting or gentle vortexing.
35. Incubate samples for 10 min before transferring plate on magnetic stand. Wait for ~2–3 min until the beads are well separated.
    ▲ **CRITICAL STEP** It is of high importance to wait until all magnetic beads have settled and the solution is clear before proceeding to the next step!
36. Remove and discard supernatant carefully while the plate is still placed on the magnetic stand.
37. Keep the plate on the magnetic stand and wash beads twice with 200 µl of freshly prepared 80% (vol/vol) ethanol in water. Incubate for 30 s at room temperature in between and remove and discard supernatant carefully.
    ▲ **CRITICAL STEP** Ensure complete removal of any residual ethanol; this is crucial for the successful elution in the subsequent steps.
38. Dry beads for 5–10 min on a magnetic stand. Do not overdry the beads and continue as soon as the beads are matte in appearance.
    ▲ **CRITICAL STEP** The drying time may vary depending on the room temperature and the efficiency of the ventilation system.
39. Detach the plate from the magnetic stand and add 17 µl nuclease-free water. Pipette up and down 10 times to ensure proper resuspension of beads. Incubate for 2 min at room temperature.
40. Move the plate back on magnetic rack and wait for 2–3 min until the beads are well separated.
41. Transfer 15 µl of eluate into a new well plate and discard the plate containing magnetic beads.
    ■ **PAUSE POINT** cDNA samples are either directly processed for QC or can be stored at −20 °C for several months.

Quality control

▲ **CRITICAL** The quality and quantity of the generated cDNA is assessed using a Qubit flex (quantity) and a Bioanalyzer (quality). Representative Bioanalyzer traces for *S.* Typhimurium and *P. aeruginosa* can be found in Fig. 3.

42. Measure the cDNA concentration of each sample by Qubit flex using the 1× dsDNA High Sensitivity kit following the manufacturer's instructions.
◆ **TROUBLESHOOTING**
43. Dilute samples to a concentration of 1–2 ng/µl in nuclease-free water and test samples covering the available concentration spectrum on a High Sensitivity DNA Bioanalyzer chip following the manufacturer's instructions.
◆ **TROUBLESHOOTING**
■ **PAUSE POINT** cDNA samples can be stored at −20 °C for up to several months.

Library preparation (day 4)

● **TIMING 1–2 d (depends on number of samples and use of a pipetting/dispensing robot)**
▲ **CRITICAL** The library preparation protocol relies on the Nextera XT kit (Illumina), but includes several modifications and the integration of the DASH protocol.
44. Dilute all cDNA samples to 0.4 ng/µl in nuclease-free water.
45. Thaw all reagents from the Nextera XT kit stored at −20 °C and keep them on ice. Store Neutralize Tagment buffer at room temperature.
46. Thaw UDI plate(s) on ice.
47. Mix 1.25 µl of diluted cDNA (0.4 ng/µl) with 2.5 µl Tagment DNA buffer.
48. Add 1.25 µl Amplicon Tagment Mix and mix samples well by pipetting.
49. Spin down samples and incubate for 10 min at 55 °C. Cool down samples to 10 °C and immediately add 1.25 µl Neutralize Tagment buffer.
50. Mix samples well by pipetting, spin down and incubate for 5 min at room temperature.
51. Vortex and spin down UDI plates.
52. Add 3.75 µl of Nextera PCR Master Mix and 2.5 µl of UDIs to each sample.
▲ **CRITICAL STEP** It is crucial to keep track of all used UDIs and the related samples for subsequent data analysis.
53. Mix samples using a microplate shaker or by gentle vortexing, then spin them down.
54. Run the library PCR program as follows:

Cycle number	Denaturation	Annealing	Elongation	Hold
1			72 °C, 3 min	
2	95 °C, 30 s			
3–15	95 °C, 10 s	55 °C, 30 s		
			72 °C, 30 s	
16			72 °C, 5 min	
17				10 °C, ∞

55. Perform cleanup of libraries using 12.5 µl AMPure XP beads as described above (Steps 33–38).
56. Elute libraries with 13.1 µl nuclease-free water and transfer 11 µl of eluate into a new well plate.
57. Measure library concentrations of each sample by Qubit flex using the 1× dsDNA High Sensitivity kit following the manufacturer's instructions.
◆ **TROUBLESHOOTING**
■ **PAUSE POINT** Libraries can be stored at −20 °C for at least 1 year or directly used for DASH.

Ribosomal depletion via DASH

58. Incubate the sgRNA pool (Box 1) for 3 min at 94 °C followed by 5 min at 4 °C.
59. Mix 1.85 µl of 3.1 buffer (NEB, included in the Cas9 enzyme kit) with sgRNA and Cas9. The provided ratio calculator will calculate the necessary volumes of sgRNA and Cas9 used for the following DASH reaction (Supplementary Table 2). Add nuclease-free water to a total volume of 10.5 µl.
▲ **CRITICAL** The ratio calculation relies on the input cDNA for Nextera index PCR. To adapt the ratio calculation for other library preparation protocols, it is necessary to adjust the ratio in accordance with the amount of cDNA enriched during PCR.
60. Incubate for 15 min at 37 °C to allow sgRNA–Cas9 complex formation.

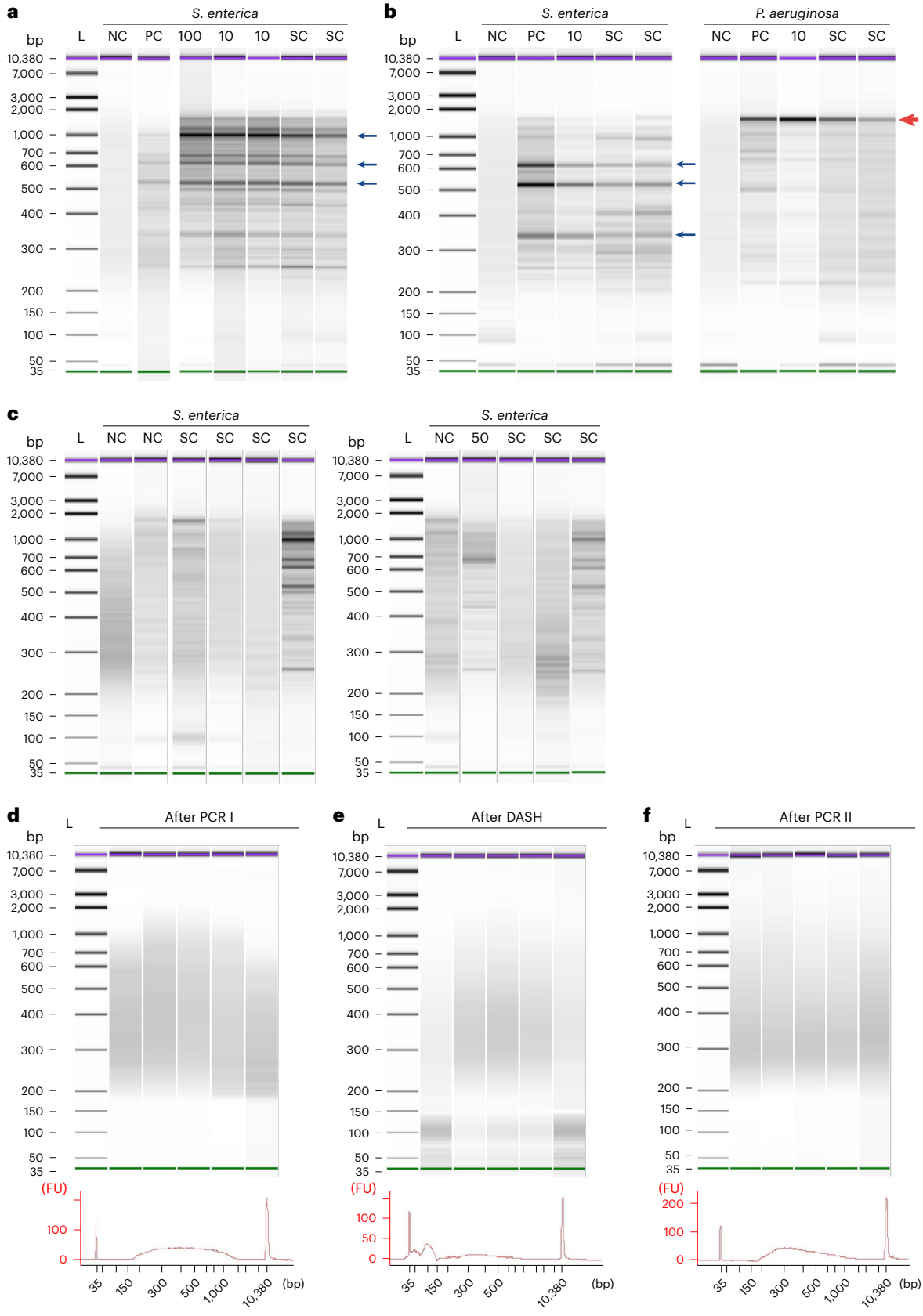
a
S. enterica
L
NC
PC
100
10
10
SC
SC
bp
10,380
7,000
3,000
2,000
1,000
700
600
500
400
300
200
150
100
50
35
b
S. enterica
L
NC
PC
10
SC
SC
P. aeruginosa
NC
PC
10
SC
SC
c
S. enterica
L
NC
NC
SC
SC
SC
SC
S. enterica
L
NC
50
SC
SC
SC
d
L
After PCR I
(FU)
100
0
35
150
300
500
1,000
10,380
(bp)
e
L
After DASH
(FU)
100
50
0
35
150
300
500
10,380
(bp)
f
L
After PCR II
(FU)
200
100
0
35
150
300
500
1,000
10,380
(bp)

Fig. 3 | QC of MATQ-seq and library preparation including DASH. a, DNA High Sensitivity Bioanalyzer profile of cDNA obtained after MATQ-seq. Each profile represents the cDNA prepared from either a single cell (sc), 10 sorted or 100 sorted cells. The positive control (PC) was performed with a spike-in of 50 pg of total RNA. Characteristic bands are indicated with blue arrows. Adapted with permission from ref. 8, ASM. **b**, Comparison of cDNA profiles from *S. enterica* and *P. aeruginosa*, highlighting species-specific bands that are highly reproducible among samples by blue (*S. enterica*) and red (*P. aeruginosa*) arrows. **c**, Illustrative Bioanalyzer profiles with undesired outcomes in both the negative control (NC) and sorted cell conditions. These profiles represent examples of nonspecific amplification and suboptimal MATQ-seq efficiency. The last lane of both blots shows an optimal outcome for comparison. **d**–**f**, Bioanalyzer profiles at each step of the library preparation process are presented for five representative single-cell libraries in each of the lanes 2–6. An exemplary electropherogram of lane 2 is displayed below: the Bioanalyzer profiles after index PCR are identical to those obtained from normal Nextera XT library preparation protocol with an average library size in the range of 400–500 bp (**d**); libraries after Cas-9 mediated cleavage by DASH characteristically show a reduction of the main peak due to cleavage of rRNA-derived cDNA and the appearance of an additional smaller peak in the range of 50–150 bp (**e**); and the reamplification of non-cleaved libraries and the subsequent library purification leads to the elimination of the smaller peak and the recovery of the Bioanalyzer pattern as in **d** (**f**). L, ladder. Partially adapted with permission from ref. 30, Humana.

61. Add 8 µl of library or pooled libraries (libraries can be normalized based on their concentration and a maximum of 12 samples can be pooled within a single reaction). For pooling, all samples are adjusted to the smallest measured library concentration using nuclease-free water and pooled in same volumes (1 µl each for 8–12 samples, 2 µl each for 4–7 samples). Skip samples with a library concentration below 5 ng/µl for pooling.
62. Incubate for 2 h at 37 °C to allow the DASH reaction to proceed.
63. Inactivate Cas9 by adding 1 µl proteinase K immediately after the DASH reaction.
64. Incubate for 15 min at 37 °C. Immediately add 0.5 µl of 40 mM PMSF to inactivate proteinase K.
65. Perform cleanup of libraries using 12.5 µl AMPure XP beads, as described in Steps 33–38.
66. Elute libraries with 13.1 µl nuclease-free water and transfer 11 µl of eluate into a new well plate.
67. Measure library concentrations of each sample by Qubit flex using the 1× dsDNA High Sensitivity kit following the manufacturer's instructions.
    ◆ **TROUBLESHOOTING**
    ■ **PAUSE POINT** Store cDNA at −20 °C until further use for up to several months.
68. Normalize libraries to 0.5 ng/µl in nuclease-free water and use 6.25 µl each. Mix libraries with 3.75 µl Nextera PCR Master Mix and add 1.25 µl of each DASH amplification primer (forward and reverse) to reach a reaction volume of 12.5 µl.
69. Mix samples using a microplate shaker or by gentle vortexing, then spin them down.
70. Run the following library PCR program:

Cycle number	Denaturation	Annealing	Elongation	Hold
1			72 °C, 3 min	
2	95 °C, 30 s			
3–15	95 °C, 10 s	55 °C, 30 s		
			72 °C, 30 s	
16			72 °C, 5 min	
17				10 °C, ∞

71. Repeat Steps 55 and 56.
    ■ **PAUSE POINT** Store libraries at −20 °C until further use for at least one year.

Pooling, final QC and sequencing (day 5)

● **TIMING 4–8 h (depends on the number of samples and usage of a pipetting/dispensing robot)**

72. Determine library concentrations using a Qubit flex as described in Step 42.
    ◆ **TROUBLESHOOTING**
73. Check average library sizes (in bp) on a High Sensitivity DNA Bioanalyzer chip following the manufacturer's instructions. Set the length range manually from 170 to 2,000 bp in the Bioanalyzer Expert software. Representative Bioanalyzer traces for *S.* Typhimurium and *P. aeruginosa* can be found in Fig. 3.
    ◆ **TROUBLESHOOTING**

74. Adjust libraries to a final concentration of 5 nM using sequencing buffer (10 mM Tris–HCl, pH 8, 0.1% (vol/vol) Tween-20) and pool in equimolar ratios. Determine molarities using the following formula:

$$\text{concentration (nM)} = \frac{\text{concentration (ng/µl)} \times 10^6}{\text{avg.}\frac{\text{library}}{\text{pool}}\text{size (bp)} \times 660\ \text{g/mol}}$$

75. Perform QC of the final pool on a Qubit flex (measurement in triplicate is recommended) and High Sensitivity DNA Bioanalyzer as described in Steps 42 and 43.
■ **PAUSE POINT** Sequencing pool can be stored at −20 °C until further use for at least 1 year.
▲ **CRITICAL STEP** If libraries or sequencing pools are stored at −20 °C for longer than 1 month, we recommend requantifying previous results from the Qubit and Bioanalyzer before proceeding with sequencing.

Sequencing

● **TIMING 24 h (depends on the sequencing platform)**

76. Spike the sequencing pool with 1% PhiX control library and perform sequencing in single-end mode with 100 cycles. Our recommendation is to use a minimum of 75 cycles, as 50 cycles has been shown to have a negative impact on MATQ-seq data quality due to lower coverage of the cDNA insert in the final library.
▲ **CRITICAL STEP** The optimal sequencing depth for a particular experiment varies based on factors such as the research question, number of samples and specific genes under investigation. A previously described downsampling experiment[7] can provide valuable guidance for selecting a sequencing depth tailored to the individual experiment.

Data analysis

Demultiplexing

● **TIMING 1 working day (30 min hands-on time). This highly depends on the sequencer output and the hardware specifications.**

77. Perform demultiplexing using the bcl2fastq2 Conversion Software from Illumina (v20.2) generating a FASTQ sequencing file per sample.

Installing Ubuntu-based software

● **TIMING 1–2 min**

78. Use the following command, in which `curl` is used to download files. `parallel` allows multiple tasks to run in parallel.

```
$ sudo apt-get install curl parallel
```

▲ **CRITICAL STEP** The 'apt-get' command is a front-end for the package manager used by Debian-derived Linux distributions, such as Ubuntu. Similar package managers exist for most popular UNIX operating systems, such as RPM for Fedora and Red Hat-derived Linux distribution (for example, CentOS), or Homebrew for Mac OS X, but may require additional installation.

Install Bioconda and bioinformatics tools

● **TIMING 5–10 min**

79. Bioconda is a bioinformatics channel for the Conda package manager and provides an interface for installing software[26]. Install Bioconda using the following command:

```
$ curl https://repo.anaconda.com/miniconda/Miniconda3-latest-Linux-x86_64.sh
-o miniconda.sh
$ bash miniconda.sh
#This refreshes the terminal startup items, so it doesn't have to be
reopened to start using bioconda
$ source ~/.bashrc
```

▲ **CRITICAL STEP** Note: during the install, you will be required to accept the licence terms, accept the default install location and select Yes for Conda init.

80. Add the bioconda channels and install software as follows:

```
#Channels are where different software is stored
$ conda config --add channels defaults
$ conda config --add channels bioconda
$ conda config --add channels conda-forge
#Install required software using conda
$ conda install -c bioconda fastqc multiqc bowtie2 subread bbmap
samtools
```

Download bacterial genome and annotation

● **TIMING 5 min**

81. Download the genome and annotation files for the bacteria being analyzed. We are using files downloaded from NCBI (https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_000210855.2/), using the GenBank annotation. The genome of *Salmonella* Typhimurium SL1344 includes a chromosome and three plasmids, which need to be concatenated for downstream use. We have already prepared these concatenated genome and annotation files, which can be downloaded using the following code:

```
#Make and change into the MATQSeq directory where all files will #be
saved
$ mkdir MATQSeq
$ cd MATQSeq
$ mkdir -p Genome_annotation
#Download GFF
$ curl https://raw.githubusercontent.com/BarquistLab/MATQ-seq_2023/
main/Protocol/salmonella_SL1344.fa -o Genome_annotation/salmonella_
sl1344.fa
#Download GFF
$ curl https://raw.githubusercontent.com/BarquistLab/MATQ-seq_2023/
main/Protocol/salmonella_SL1344.gff3 -o Genome_annotation/salmonella_
sl1344.gff3
```

Download sequence files

● **TIMING 1–1.5 h**

82. The sequencing reads from our updated protocol[7] are stored in the European Nucleotide Archive (ENA)[27] under the project accession PRJNA904727. From the ENA, access a tab-separated (.tsv) report containing a download link for each cell and the corresponding file names from this link: https://www.ebi.ac.uk/ena/browser/view/PRJNA904727.
83. This file has already been prepared and can be downloaded with the following line of code:

```
$ curl https://github.com/BarquistLab/MATQ-seq_2023/blob/main/
Protocol/files.tsv -o filelist.tsv
```

84. Use the following command to loop through each line of filelist.tsv, starting from the second line (that is, ignoring the column headings). Column 15 is used to name the downloaded file and column 12 contains the ftp download link.

```
$ mkdir sequenced_reads
$ awk -F'\t' 'NR>1 {system("curl -o sequenced_reads/" $15 ".fq.gz "
$12)}' filelist.tsv
```

QC of sequenced reads

● **TIMING 30–45 min**

85. FastQC creates a report for each FASTQ file, generating a range of useful metrics for data quality. These include sequence and base quality scores, GC content, overrepresented sequences and adaptor content (more information can be found at https://www.bioinformatics.babraham.ac.uk/projects/fastqc/Help/). To use FastQC in combination with `parallel`, allowing us to run multiple instances of FastQC at the same time to efficiently process all sequence files. Here we are using 8 threads (*-j*)—this can be increased depending on the computational resources available. Use the following command to perform the QC:

```
$ mkdir -p fastqc
$ ls sequenced_reads/*.fq.gz | parallel -j 8 'fastqc {} -o fastqc'
```

Read trimming, removal of adaptors and primers

● **TIMING 1 h**

▲ **CRITICAL** To remove the sequencing adapters and primers, we use BBDuk (https://archive.jgi.doe.gov/data-and-tools/software-tools/bbtools/bb-tools-user-guide/bbduk-guide/). An advantage of using BBDuk is the bundled file containing common sequencing adapters, such as the Nextera adapters used in the MATQseq protocol[6]. The most efficient way of removing sequencing adapters and MATQseq specific primers is to run BBDuk in two phases: (1) remove MATQseq primers from the 5′ read end while also trimming any poly(A) tails from the 3′ end and (2) remove sequencing adapters and MATQseq primers from the 3′ end.

86. Prepare adapters and primers files as follows:

```
#Download MATQ-seq primers
$ curl https://raw.githubusercontent.com/BarquistLab/MATQ-seq_2023/
main/Protocol/matqseq_primers.fa -o matqseq_primers.fa
#Copy BBDuk adapters file to current folder
$ cp $HOME/miniconda3/opt/bbmap-39.01-0/resources/adapters.fa nextera_
and_primers.fa
#Add MATQ-seq primers to the end of the BBDuk adapters file
$ cat matqseq_primers.fa >> nextera_and_primers.fa
```

▲ **CRITICAL STEP** $HOME is a system variable that contains the default user home directory. This is generally /home/user on Linux systems.

87. Run BBDuk as follows, with the following parameters:

```
#Create BBDuk directories
$ mkdir -p BBDuk_L BBDuk_L_R
#BBDuk parameters
$ bbduk_opts="-Xmx16g t=8 minlen=18 qtrim=rl \
trimq=20 k=17 mink=11 hdist=1"
#Process all sequenced cells
$ for file in sequenced_reads/*.fq.gz; do
filename = "${file##*/}" # remove path from file name
filename = "${filename%.fq.gz}" # remove file extension
read = "${filename}.fq.gz"
# Trim left
bbduk.sh $bbduk_opts \
in=$file out=BBDuk_L/$read \
ref=matseq_primers.fa ktrim=l trimpolya=30
# Trim right
bbduk.sh $bbduk_opts \
```

```
in=BBDuk_L/$read out=BBDuk_L_R/$read \
ref=nextera_and_primers.fa ktrim=r
done
```

Parameter	Description
XmxNNg	The maximum amount of RAM to use (NN in GB, e.g., 16 in the above example)
t	The number of threads to use
minlen	Discard trimmed reads shorter than this length
qtrim	Quality-trim reads at the 5′ (l) 3′ (r) or both ends (rl)
trimq	Quality trim using the Phred algorithm
ktrim	Adapter trimming direction 5′ (l) 3″ (r)
k	*k*-mer size to use for adapter trimming
mink	Allows shorter *k*-mers to be used at the end of reads
hdist	Hamming distance (no. of mismatches allowed)
trimpolya	Trim poly(A) tails

▲ **CRITICAL STEP** FastQC can be run again at this point to examine the results after running BBDuk.

Alignment and feature counting

● **TIMING 2.5 h**

88. We use Bowtie2 (ref. 23) for read alignment and `featureCounts`[24] to count the number of reads recovered for each feature. The first step is to generate a genome index using Bowtie2 as follows:.

```
$ bowtie2-build Genome_annotation/salmonella_sl1344.fa Genome_
annotation/sl1344_index
```

89. Run Bowtie2 and convert the output to a BAM file using `samtools`[28]. We then count the number of assigned reads to each feature using `featureCounts` and the Salmonella annotation (.gff) as follows, with the following parameters:

```
#Create output directories
$ mkdir -p bowtie2_aligned featureCounts
#Loop over the trimmed files
$ for file in BBDuk_L_R/*.fq.gz; do
filename="${file##*/}" # remove path from file name
filename="${filename%.fq.gz}" # remove file extension
# Set the output filenames
bam_file="${filename}.bam"
count_file="${filename}.count"
# Align with Bowtie2 and convert SAM to BAM with Samtools
bowtie2 -p 8 --local \
-x Genome_annotation/sl1344_index \
-U $file | samtools view -@ 8 -bS -h - > \
bowtie2_aligned/$bam_file
#Count the number of reads assigned to each genomic feature
featureCounts -T 8 \
-a Genome_annotation/salmonella_sl1344.gff3 \
-t "gene,rRNA,tRNA,pseudogene" -g "ID" \
-o featureCounts/$count_file \
bowtie2_aligned/$bam_file
```

Parameter	Description
Bowtie2	
local	Local alignment mode
x	Path to the genome index
N	Number of mismatches allowed in alignment
U	Path to single-end read file
P	Number of threads to use
Samtools	
@	Number of threads to use
bS	Input is SAM format and output in BAM
h	Include the header in the BAM file
featureCounts	
T	Number of threads to use
t	Genomic feature (3rd column in GFF)
g	Unique identifier for each feature (9th column in GFF)

▲ **CRITICAL STEP** Depending on the bacterium and annotation used, different feature names (for example, gene, rRNA, pseudogene in the above example) may be used. You should examine the annotation file and if necessary, adjust the arguments given to `featureCounts` for features (-t) and identifier (-g) so it matches the specific annotation.

Generate sequence report

● **TIMING 1–2 min**

90. MultiQC is an efficient tool to produce interactive QC reports based on the logs produced by the tools run in the previous steps across all cells, generating a single dynamic HTML report[22]. Use MultiQC as follows to provide an overview of the results and indicate if any samples may have failed:

```
$ multiqc -o multiqc/ .
```

Post-processing

● **TIMING 20 min**

91. Collate the count files obtained using featureCounts into a count matrix where rows are genes and columns are cells. Use this count matrix in downstream analysis. After obtaining quantifications, various analytical approaches can be employed depending on the specific biological question. For instance, PCA is used for visualizing high-dimensional data by reducing the dimensionality into principal components that capture uncorrelated sources of variation in the data. Plotting cells on these components can identify cell populations that differ from one another for further exploration. We provide R code to generate the PCA plot shown in Fig. 5 (Supplementary Code).

Troubleshooting

Troubleshooting advice can be found in Table 2.

Table 2 | Troubleshooting

Step	Problem	Possible reason	Solution
Steps 42 and 43	Positive control is not working	RNA quality	Check RNA on Nanodrop and on a gel, Bioanalyzer, etc
		RNA quality	Dilutions of total RNA (1 ng/µl and lower) stored at −20 °C tend to degrade over time. Prepare new dilutions frequently

Table 2 (continued) | Troubleshooting

Step	Problem	Possible reason	Solution
		Issue with one of the reagents	Use new aliquots of water, buffers and primer. If this does not help, enzymes can be exchanged one by one as a last possibility
		I.DOT dispensing robot displays inconsistent dispensing volumes and/or drop detection errors with one or more reagents	If I.DOT is used, perform experiments manually to make sure it is not a dispensing problem
		Sterility issue—contamination and degradation of RNA	Decontaminate the working area and wipe all materials with RNAseZap
		Bead cleanup	Make sure beads have not been overdried. To test, measure the Qubit before cleanup
Steps 42 and 43	Negative control is higher that 2 ng/µl and shows band pattern on Bioanalyzer	Sterility issue—contamination and degradation of RNA	Assess the sterility of the workflow. This includes ensuring a clean RNA-seq workbench, separating work involving living organisms from this area and considering FACS as a potential contamination source. Make sure the FACS tubing has been properly flushed and cleaned before use. Remove and clean the nozzle in a sonicater bath before re-use. Use ultraclean water for cleaning and run a water-only sample to check for contamination
		Cross-contamination	Use strip tubes with individual caps instead of well plates. Physically separate negative and positive controls (leave empty wells in between)
Steps 42 and 43	cDNA is obtained for 100/ 10 cells but not for single cells	If cDNA yield is only an issue for single cells, cell lysis might not be efficient	Try alternative lysis options and make sure that the MATQ-seq workflow is running well with a spike-in of low amounts of total RNA (e.g., 10 pg)
Steps 42 and 43	cDNA concentration of single-cells is below 3 ng/µl	Timing issue	Make sure all pipetting steps are performed fast and efficiently. Keep samples on ice at all times
Steps 42 and 43	Bioanalyzer shows major peaks below 150 bp	Excess of primer dimers	Make sure ethanol evaporated before elution and beads have not been overdried
Step 57	Library concentration is below 5 ng/µl	Efficiency issue	Check library on Bioanalyzer. If no broad peak is detectable, library preparation did not work. Repeat the PCR step and prepare new dilutions of cDNA input material
Step 67	Library concentration is above 3 ng/µl	DASH did not work properly	Check libraries on Bioanalyzer. There should be only a broad pattern and no peak. In case of a peak, check whether sgRNA was degraded. Minimize freeze–thawing of sgRNA and repeat the DASH protocol
Steps 72 and 73	Library concentration is below 10 ng/µl	Library amplification after DASH did not work	Verify the concentration/dilution of the input material and correct primer concentration. Repeat PCR amplification
Steps 72 and 73	Bioanalyzer shows major peaks below 150 bp	Excess of primer dimers	Make sure ethanol evaporated before elution and beads have not been overdried
Box 1	Concentration of DNA products is below 30 ng/µl	Inefficient fill-in reaction	Make sure primer pool and fill-in oligonucleotide have correct concentration. Repeat PCR and measure Qubit before proceeding with cleanup
Box 1	Additional shorter peaks appear	Issue with cleanup	Check consumables used in the purification kit, make sure all steps were correctly executed. Repeat PCR and cleanup and make sure all residual buffer components are removed before elution
Box 1	Low concentration of sgRNA pool (<100 ng/µl)	Inefficient IVT	Check quality of DNA template, make sure DNA is not degraded and verify dNTP concentration. Repeat IVT and purification, place sgRNA on ice immediately after elution and stored it at −80 °C
Box 1	Additional shorter peaks appear	Issue with cleanup	Check consumables used in the RNA cleanup kit, make sure all steps were correctly executed. Repeat IVT and cleanup and make sure all residual buffer components are removed before elution

Timing

Timing information can be found in Table 3. Timing for data analysis has been estimated using a Linux laptop running Ubuntu 22.04 LTS with 8 CPUs, 16 GB RAM and a solid-state hard drive. Timing may vary depending on your computational environment. All code can be run on a computing cluster; increasing computational resources can reduce the overall run times.

Table 3 | Timing

	Steps			Timing	Stop and store
Varies	Bacterial culture preparation			4–12 h	
Day 1, 4–6 h	**Sample preparation and single-cell isolation**				
	1–2	Prepare lysis buffer and fill plates		50 min	
	3–6	Prepare bacterial culture for FACS		30 min	
	7–8	FACS sort cells		100 min	Store at −80 °C
Day 2, 8–10 h	**MATQ-seq protocol automated with I.DOT**				
	9–14	Thaw plates and prepare initial mixes		60 min	
	15–17	Reverse transcription		2.15 h	
	18–20	Primer digestion		2 h	
	21–22	RNA digestion		1 h	
	23–25	Poly(C) tailing		45 min	
	26–29	Second-strand synthesis		45 min	
	30–32	PCR amplification		2 h	Store at −20 °C
Day 3, 3–4 h	**cDNA purification**				
	33	Equilibrate beads to room temperature		30 min	
	34	Mix beads with each sample		10 min	
	35	Incubate beads with sample		10 min	
	36–37	Magnetic stand and two washes		15 min	
	38	Dry beads on magnetic stand		8–10 min	
	39–41	Elute cDNA from beads and transfer purified sample		10 min	Store at −20 °C
	QC				
	42	cDNA quantification via Qubit		30 min	
	43	Bioanalyzer		50 min	Store at −20 °C
Day 4, 8–9 h	**Library preparation via Nextera XT**				
	44	Dilution of cDNA samples		30 min	
	45–54	Nextera XT		2 h	
	55–56	Bead purification		1 h	
	57	QC via Qubit flex		20 min	Store at −20 °C
	Ribosomal RNA depletion via DASH				
	58–61	sgRNA–Cas9 complex formation		25 min	
	62–67	DASH reaction and Cas9 inactivation		2.5 h	
	68–70	DASH amplification		35 min	
	71	Library bead purification		1 h	Store at −20 °C
Day 5, 4–8 h	**Pooling, final QC and sequencing**				
	72	Determining library concentration via Qubit		30 min	
	73	Determining average library size via Bioanalyzer		60 min	
	74	Adjusting libraries to final concentration		30 min	Store at −20 °C
	75	Library pooling and QC		1 h	Store at −20 °C
Day 6	**Sequencing**				
	76	Sequencing single-end mode 100 cycles		24 h	
Day 7	**Data analysis**				
	77		Demultiplexing	30 min	
	78		Install Ubuntu-based software	1–2 min	
	79–80		Install Bioconda and bioinformatics tools	5–10 min	
	81		Download bacterial genome and annotation	5 min	
	82–84		Download sequence files	1–1.5 h	

Table 3 (continued) | Timing

Steps		Timing	Stop and store
85	QC of sequenced reads	30–45 min	
86–87	Read trimming, primer and adaptor removal	1 h	
88–89	Alignment and feature counting	2.5 h	
90	Generate sequence report	1–2 min	
91	Post-processing	20 min	

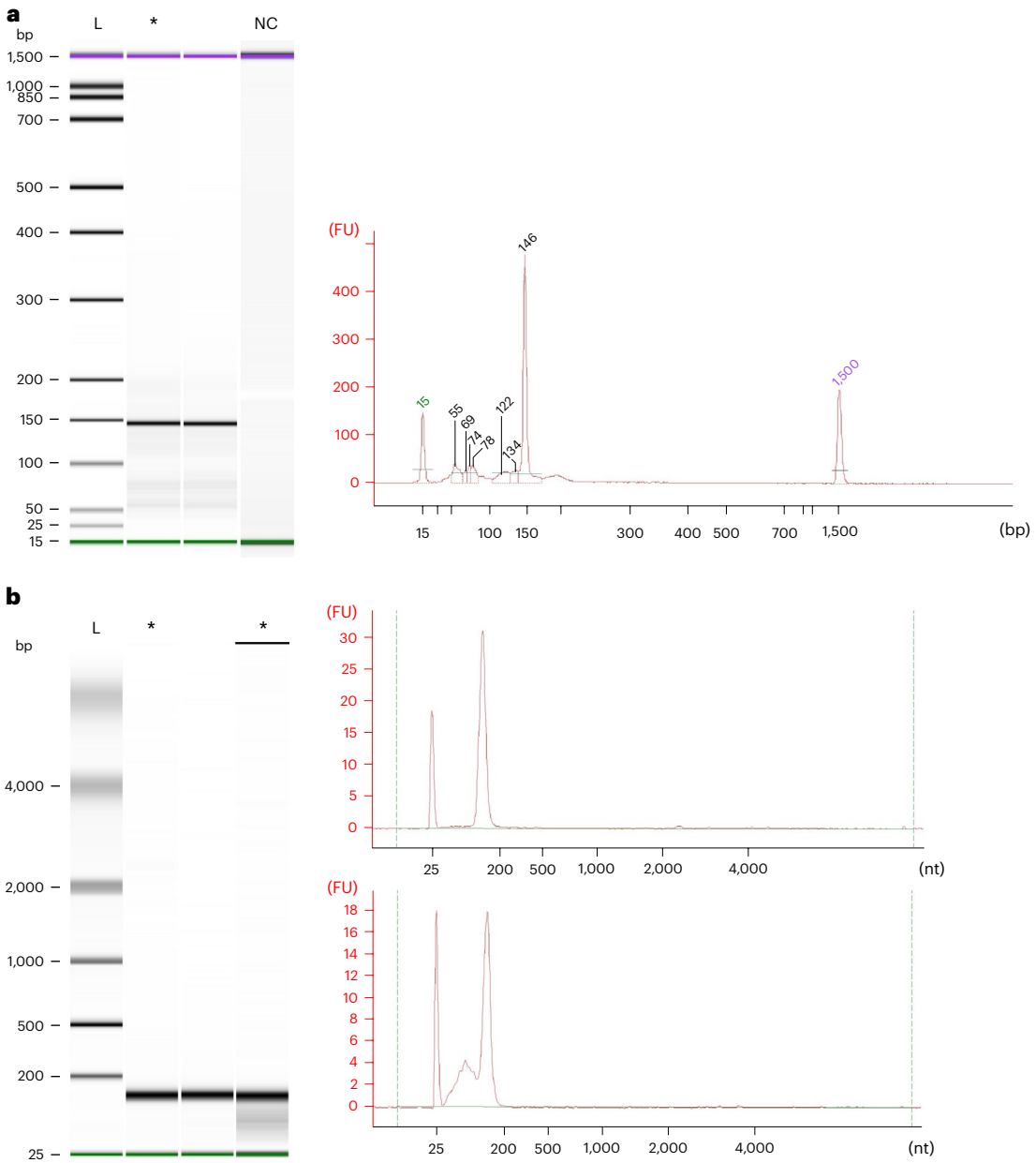

Fig. 4 | QC of sgRNA synthesis. a, Illustrative Bioanalyzer DNA 1,000 profiles depicting the intermediate DNA pool generated during sgRNA synthesis. The data includes two representative DNA pools, along with a negative control exhibiting no detectable peaks. The electropherogram of the pool marked by an asterisk with a primary peak at 146 bp is presented on the right. **b**, QC of the sgRNA pool after IVT using the Bioanalyzer RNA Pico kit. Two desired products are shown along with one exhibiting an undesirable profile. The additional peak in the third pool represents an example of incomplete purification or a product of artificial amplification. sgRNA pools showing such additional peaks should be excluded for subsequent use in the DASH reaction. Electropherograms of pools marked by asterisks are shown on the right: the desired one is displayed above, the undesired one below.

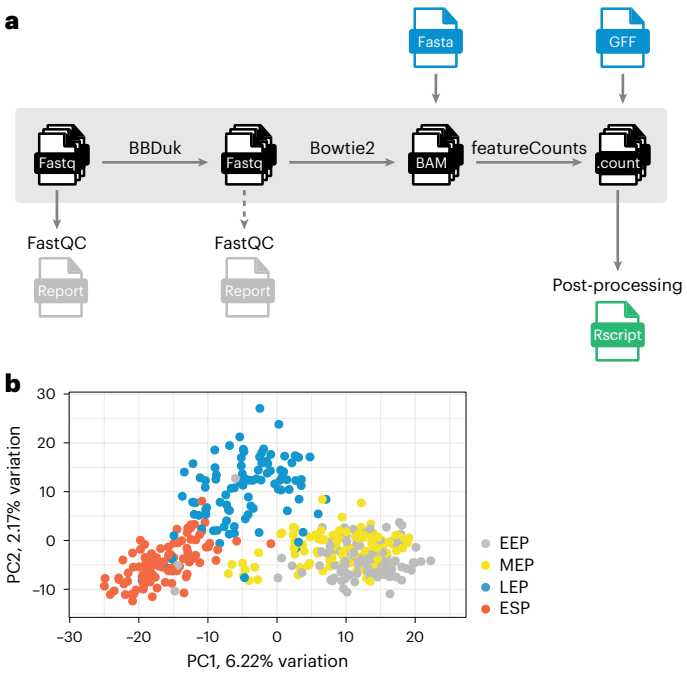

Fig. 5 | Bioinformatic processing and analysis. a, Computational analysis pipeline. Once FASTQ files are downloaded, basic QC can be performed using FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Adapters and MATQ-seq primers are removed using BBDuk (https://archive.jgi.doe.gov/data-and-tools/software-tools/bbtools/bb-tools-user-guide/bbduk-guide/). Bowtie2[23] is used to map reads to the bacterial genome and featureCounts[23] is used to count reads assigned to genomic features (genes). Post-processing using the resulting count matrix is performed using R. **b**, An example PCA plot as output from computational analysis showing the recovery of growth phase clusters from scRNA-seq. The four conditions are: early exponential phase (EEP) at OD 0.1, mid exponential phase (MEP) at OD 0.3, late exponential phase (LEP) at OD 1.0 and early stationary phase (ESP) at OD 2.0. The R script to generate this PCA plot from the featureCounts output is available as Supplementary Code. Partially adapted with permission from ref. 8, ASM.

Anticipated results

The first opportunity for QC in the MATQ-seq protocol is after PCR amplification because the RNA and cDNA concentrations before this step are too low for detection. After cDNA purification, a concentration of 3–10 ng/µl can be expected for individual cells. The Bioanalyzer traces should show characteristic rRNA peaks and match the positive control in this case for *S. enterica* (Fig. 3a–c). Of note, these peaks are species-specific. Therefore, running positive controls using a spike-in of 50 pg of total RNA of the bacterial species under investigation is highly recommended in each experiment. In addition, the Bioanalyzer profile of a negative control should always be checked to identify possible contamination leading to unspecific amplification (Fig. 3a–c). When first implementing the MATQ-seq protocol, it is advisable to begin with sorted gradients ranging from 100, 50, 10 cells down to single cells to estimate workflow efficiency and ensure proper cell lysis.

For library preparation, the initial QC is conducted after the index PCR. Library concentrations are anticipated to fall within the range of 5–20 ng/µl, with the Bioanalyzer pattern displaying a broad distribution and an average library size of 400–500 bp (Fig. 3d). Following DASH, the targeted cleavage of rRNA-derived cDNA results in a concentration drop to 1–3 ng/µl. On the Bioanalyzer, an additional smaller peak emerges from cleaved products (Fig. 3e). After reamplification of non-cleaved products, the original library peak reappears with an average size of 400–500 bp, and concentrations between 10 and 30 ng/µl are expected (Fig. 3f).

Using DASH in *S. enterica*, the proportion of remaining rRNA reads was 75% on average[8]. Within replicates and under different growth conditions, the remaining rRNA reads varied between 65% and 85% compared with 97% of rRNA reads without depletion[8]. For other organisms, these efficiencies are expected to vary due to changes in the sgRNA pool and the overall amount of rRNA.

The synthesis of the sgRNA pool involves two primary stages, each followed by a QC step. Following the fill-in reaction, a DNA concentration ranging from 100 to 200 ng/µl is anticipated. The negative control should be below 50 ng/µl. The Bioanalyzer profile should exhibit a distinct peak between 140 and 150 bp (Fig. 4a). In the subsequent in vitro transcription (IVT) phase,

the final sgRNA pool is expected to have a concentration exceeding 800 ng/µl and should display a sharp peak on the Bioanalyzer profile within the range of 150–160 bp (Fig. 4b). Samples showing additional smaller peaks can be repurified using the Monarch RNA Cleanup kit. If the peak persists, the sample should be excluded and not used for DASH.

After sequencing of the library, a computational pipeline is used to analyze the scRNA-seq dataset (Fig. 5a). This pipeline generates QC reports and output files that are used for post-processing analysis such as determining the number of detected genes per cell, RNA class distribution, gene expression profiles and PCA[7,8] (Fig. 5b).

Reporting summary

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

Data availability

The sequencing data is available from the Sequence Read Archive under BioProject number PRJNA904727. The annotated genome of *Salmonella enterica* subsp. *enterica* serovar Typhimurium str. SL1344 is available from NCBI at https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_000210855.2/.

Code availability

All code and files used in this protocol are available at https://github.com/BarquistLab/MATQ-seq_2023/. The code in this protocol has been peer reviewed.

Received: 19 April 2024; Accepted: 5 February 2025;
Published online: 09 April 2025

References

1. Milo, R. & Phillips, R. *Cell Biology by the Numbers*. *Cell Biology by the Numbers* (Garland Science, 2015); https://doi.org/10.1201/9780429258770
2. Jenniches, L. et al. Improved RNA stability estimation through Bayesian modeling reveals most Salmonella transcripts have subminute half-lives. *Proc. Natl Acad. Sci. USA* **121**, e2308814121 (2024).
3. Gaisser, K. D. et al. High-throughput single-cell transcriptomics of bacteria using combinatorial barcoding. *Nat. Protoc.* https://doi.org/10.1038/s41596-024-01007-w (2024).
4. Imdahl, F. & Saliba, A.-E. Advances and challenges in single-cell RNA-seq of microbial communities. *Curr. Opin. Microbiol.* **57**, 102–110 (2020).
5. Homberger, C., Barquist, L. & Vogel, J. Ushering in a new era of single-cell transcriptomics in bacteria. *microLife* **3**, uqac020 (2022).
6. Sheng, K., Cao, W., Niu, Y., Deng, Q. & Zong, C. Effective detection of variation in single-cell transcriptomes using MATQ-seq. *Nat. Methods* https://doi.org/10.1038/nmeth.4145 (2017).
7. Imdahl, F., Vafadarnejad, E., Homberger, C., Saliba, A.-E. & Vogel, J. Single-cell RNA-sequencing reports growth-condition-specific global transcriptomes of individual bacteria. *Nat. Microbiol.* **5**, 1202–1206 (2020).
8. Homberger, C., Hayward, R. J., Barquist, L. & Vogel, J. Improved bacterial single-cell RNA-Seq through automated MATQ-Seq and Cas9-based removal of rRNA reads. *MBio* **14**, e0355722 (2023).
9. Gu, W. et al. Depletion of Abundant Sequences by Hybridization (DASH): using Cas9 to remove unwanted high-abundance species in sequencing libraries and molecular counting applications. *Genome Biol.* **17**, 41 (2016).
10. Prezza, G. et al. Improved bacterial RNA-seq by Cas9-based depletion of ribosomal RNA reads. *RNA* **26**, 1069–1078 (2020).
11. Zong, C., Lu, S., Chapman, A. R. & Xie, X. S. Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. *Science* https://doi.org/10.1126/science.1229164 (2012).
12. Cichocki, N. et al. Bacterial mock communities as standards for reproducible cytometric microbiome analysis. *Nat. Protoc.* **15**, 2788–2812 (2020).
13. Roche, B. & Bumann, D. Single-cell reporters for pathogen responses to antimicrobial host attacks. *Curr. Opin. Microbiol.* **59**, 16–23 (2021).
14. Blattman, S. B., Jiang, W., Oikonomou, P. & Tavazoie, S. Prokaryotic single-cell RNA sequencing by in situ combinatorial indexing. *Nat. Microbiol.* **5**, 1192–1201 (2020).
15. Kuchina, A. et al. Microbial single-cell RNA sequencing by split-pool barcoding. *Science* **371**, eaba5257 (2021).
16. Rosenberg, A. B. et al. Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science* https://doi.org/10.1126/science.aam8999 (2018).
17. Xu, Z. et al. Droplet-based high-throughput single microbe RNA sequencing by smRandom-seq. *Nat. Commun.* **14**, 5130 (2023).
18. Wang, B. et al. Single-cell massively-parallel multiplexed microbial sequencing (M3-seq) identifies rare bacterial populations and profiles phage infection. *Nat. Microbiol.* **8**, 1846–1862 (2023).
19. Ma, P. et al. Bacterial droplet-based single-cell RNA-seq reveals antibiotic-associated heterogeneous cellular states. *Cell* **186**, 877–891.e14 (2023).
20. Samanta, P., Cooke, S. F., McNulty, R., Hormoz, S. & Rosenthal, A. ProBac-seq, a bacterial single-cell RNA sequencing methodology using droplet microfluidics and large oligonucleotide probe sets. *Nat. Protoc.* 1–28 https://doi.org/10.1038/s41596-024-01002-1 (2024).
21. He, J., Du, S., Tan, X., Arefin, A. & Han, C. S. Improved lysis of single bacterial cells by a modified alkaline-thermal shock procedure. *Biotechniques* **60**, 129–135 (2016).
22. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (2016).
23. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
24. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
25. Westermann, A. J. et al. Dual RNA-seq unveils noncoding RNA functions in host–pathogen interactions. *Nature* **529**, 496–501 (2016).
26. Grüning, B. et al. Bioconda: sustainable and comprehensive software distribution for the life sciences. *Nat. Methods* **15**, 475–476 (2018).
27. Leinonen, R. et al. The European Nucleotide Archive. *Nucleic Acids Res.* **39**, D28–D31 (2011).
28. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021).
29. Papenfort, K. et al. Specific and pleiotropic patterns of mRNA regulation by ArcZ, a conserved, Hfq-aependent small RNA. *Mol. Microbiol.* **74**, 139–158 (2009).
30. Homberger, C., Saliba, A.-E. & Vogel, J. A MATQ-seq-based protocol for single-cell RNA-seq in bacteria. *Methods Mol. Biol.* **2584**, 105–121 (2023).

Acknowledgements
We thank A. Sparmann for editing the manuscript and E. Hauschild for excellent technical support on MATQ-seq protocol and library preparation. This work was supported by DFG grant INST 93/1068-1 (MICROSEQ, Centre for Microbial Single-Cell RNA-seq) and a DFG Gottfried Wilhelm Leibniz award (Vo875-18), both to J.V. A.-E.S. thanks the DFG SPP2389 priority program 'Emergent Functions of Bacterial Multicellularity' (project number 503267551).

Author contributions
C.H. and F.I. performed the MATQ-seq protocol. F.I. established sorting strategy and initial datasets. C.H. implemented enhancements and integrated DASH and automation features into the MATQ-seq protocol. R.J.H. performed data analysis and developed code. C.H., F.I. and R.J.H. designed the figures. J.V., A.-E.S. and L.B. supervised the project. All authors were involved in manuscript preparation.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41596-025-01157-5.

Correspondence and requests for materials should be addressed to Jörg Vogel.

Peer review information *Nature Protocols* thanks Adam Rosenthal and the other, anonymous, reviewer for their contribution to the peer review of this work.

nature portfolio

Corresponding author(s): Jörg Vogel

Last updated by author(s): 19.12.2024

Reporting Summary

Nature Portfolio 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 Portfolio policies, see our Editorial Policies 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	Sequencing data was collected from Illumina sequencer platform and demultiplexed using the bcl2fastq2 Conversion Software from Illumina (v20.2) generating FASTQ files used as input for data analysis.
Data analysis	For the data analysis the following software has been used: Bioconda, FastQC, BBDuk, Bowtie2, featureCounts, MultiQC Code availability: https://github.com/BarquistLab/MATQ-seq_2023/

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 and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio 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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The sequencing data is available from the Sequence Read Archive under BioProject number PRJNA904727. The annotated genome of Salmonella enterica subsp. enterica serovar Typhimurium str. SL1344 is available from NCBI at https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_000210855.2/

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	*Use the terms sex (biological attribute) and gender (shaped by social and cultural circumstances) carefully in order to avoid confusing both terms. Indicate if findings apply to only one sex or gender; describe whether sex and gender were considered in study design whether sex and/or gender was determined based on self-reporting or assigned and methods used. Provide in the source data disaggregated sex and gender data where this information has been collected, and consent has been obtained for sharing of individual-level data; provide overall numbers in this Reporting Summary. Please state if this information has not been collected. Report sex- and gender-based analyses where performed, justify reasons for lack of sex- and gender-based analysis.*
Population characteristics	*Describe the covariate-relevant population characteristics of the human research participants (e.g. age, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."*
Recruitment	*Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.*
Ethics oversight	*Identify the organization(s) that approved the study protocol.*

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

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

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

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

Sample size	Please see initial publications - Homberger, C., Hayward, R. J., Barquist, L. & Vogel, J. Improved Bacterial Single-Cell RNA-Seq through Automated MATQ-Seq and Cas9-Based Removal of rRNA Reads. MBio 14, (2023) DOI: 10.1128/mbio.03557-22. Imdahl, F., Vafadarnejad, E., Homberger, C., Saliba, A.-E. & Vogel, J. Single-cell RNA-sequencing reports growth-condition-specific global transcriptomes of individual bacteria. Nat. Microbiol. 5, 1202–1206 (2020) DOI: 10.1038/s41564-020-0774-1.
Data exclusions	*Describe any data exclusions. If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.*
Replication	*Describe the measures taken to verify the reproducibility of the experimental findings. If all attempts at replication were successful, confirm this OR if there are any findings that were not replicated or cannot be reproduced, note this and describe why.*
Randomization	*Describe how samples/organisms/participants were allocated into experimental groups. If allocation was not random, describe how covariates were controlled OR if this is not relevant to your study, explain why.*
Blinding	*Describe whether the investigators were blinded to group allocation during data collection and/or analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.*

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 and archaeology
☐	☒ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

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

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	This project only involves bacterial organsims and did not involve any laboratory anaimals.
Wild animals	This project only involves bacterial organsims and did not involve any wild anaimals.
Reporting on sex	*Indicate if findings apply to only one sex; describe whether sex was considered in study design, methods used for assigning sex. Provide data disaggregated for sex where this information has been collected in the source data as appropriate; provide overall numbers in this Reporting Summary. Please state if this information has not been collected. Report sex-based analyses where performed, justify reasons for lack of sex-based analysis.*
Field-collected samples	This project did not involve any field-collected samples.
Ethics oversight	No ethical approval was required since only bacterial samples were used.

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

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☐ All plots are contour plots with outliers or pseudocolor plots.

☐ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Bacterial cultures are grown, centrifuged and optionally washed with 1 x PBS or RNAlater and finally resuspended in RNAlater. Right before sorting, samples are diluted in 1x PBS to achieve an event rate of 1000 - 1500 events/second.
Instrument	BD FACS aria III
Software	FACS DIVA
Cell population abundance	Cultured bacteria. No population discrimination.
Gating strategy	Gating to discriminate bacteria vs. debris and dust using a GFP expressing bacterial strain and applied the sampe gating for wt bacteria.

☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.