


Unravelling bacterial complexity at high resolution with single-cell transcriptomics

Received: 14 December 2024

Accepted: 18 March 2026

Published online: 01 May 2026

 Check for updates


Anne E. Clatworthy^{1,2,3,4,5}, Vincenzo P. DiNatale^{1,2,4,5}, Emanuel Burgos-Robles^{2,4,5},
Fernando Lopes^{1,2,4}, Chris Smillie^{2,4,5}  & Deborah T. Hung ^{1,2,3,4,5} 

Phenotypic heterogeneity, a feature of both bacteria and eukaryotic cells, arises from inherent cell-to-cell variability. In eukaryotes, single-cell RNA sequencing has led to an explosion in understanding how heterogeneity impacts different cell types and states in organs and tissues. While single-cell RNA sequencing analyses in bacteria have lagged behind eukaryotic studies, recent technological advances now enable similar, high-resolution studies to be performed at scale in bacteria, yielding fundamental insights into how heterogeneity influences bacterial physiology, metabolism, antibiotic resistance, pathogenesis and interactions within complex microbial communities. Here we review recent advances in bacterial single-cell RNA sequencing, including the methods developed so far and what has been learned from their application. We also discuss technological and computational challenges going forwards, the need for standardization and how that could be achieved, and how this emerging field is now poised to revolutionize our understanding of bacterial physiology, infection biology and interactions within bacterial communities, such as the microbiota.

Cell-to-cell phenotypic variability is essential for the survival and function of populations of cells, whether organized in the tissues and organs of multicellular organisms or within single- or multi-species microbial communities. Clonal, genetically identical bacterial populations can display cell-to-cell phenotypic variability, termed phenotypic heterogeneity, across many traits (for example, size, metabolism, motility, virulence factor expression and antibiotic tolerance), which is thought to confer multiple possible advantages. Phenotypic heterogeneity may act as an important bet-hedging strategy to ensure that at least a subpopulation of the bacterial population is poised to adapt and survive in fluctuating environments¹, including shifts in nutrient availability² or antibiotic exposure³. Alternatively, phenotypic heterogeneity may enable the division of labour to segregate incompatible metabolic processes into distinct subpopulations, as occurs with nitrogen fixation and photosynthesis in cyanobacteria¹. Finally, it may increase evolutionary potential or better enable niche exploration, as in biofilms, where concentration of flagellar gene expression in bacteria in the periphery

helps motile clones break away from non-motile brethren to explore different, potentially more favourable environments^{1,2,4}.

In many cases, phenotypic heterogeneity is driven by variability in gene expression between cells. Historically, however, methods for characterizing the entire transcriptional programme of individual cells have been lacking, with technologies restricted to examining small numbers of genes using, for example, microscopy or flow cytometry^{3,5,6}. Meanwhile, characterization of an entire bacterial transcriptome was restricted to bulk analyses, which mask cell-to-cell variability by averaging gene expression across entire cell populations and fail to detect and elucidate the behaviours of small, but potentially important, subpopulations of phenotypically distinct cells. Bulk transcriptional measurements also cannot differentiate between shifts in the relative proportions of phenotypically distinct cell types within the population and changes in gene expression within some or all of these subpopulations. Hence, bulk transcriptional methods have limited resolution for many biological processes that require

¹Department of Molecular Biology, Massachusetts General Hospital, Boston, MA, USA. ²Center for Computational and Integrative Biology, Massachusetts General Hospital, Boston, MA, USA. ³Department of Genetics, Harvard Medical School, Boston, MA, USA. ⁴Infectious Disease and Microbiome Program, Broad Institute of MIT and Harvard, Cambridge, MA, USA. ⁵Center for Integrated Solutions for Infectious Diseases, Broad Institute of MIT and Harvard, Cambridge, MA, USA.  e-mail: csmillie@mgh.harvard.edu; hung@molbio.mgh.harvard.edu

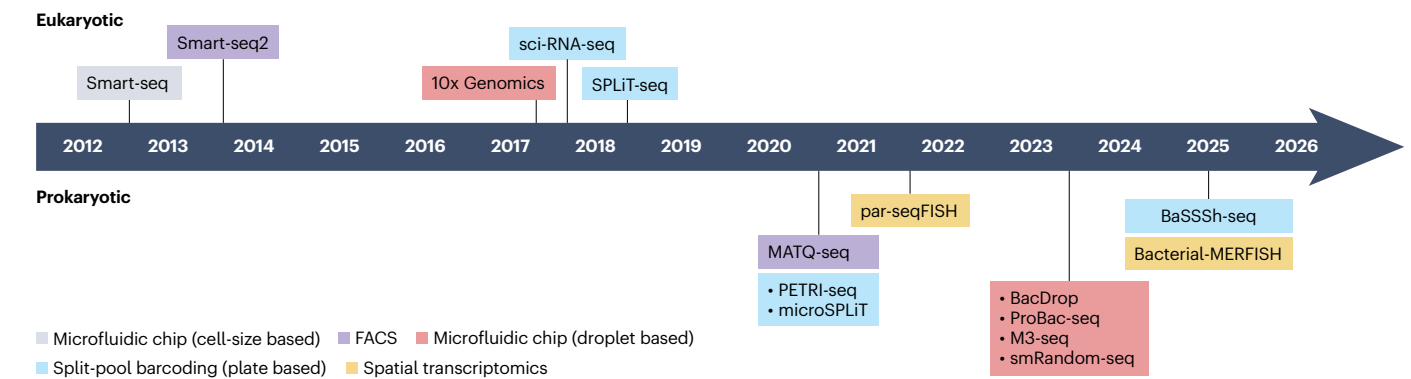


Fig. 1 | Timeline of eukaryotic and prokaryotic single-cell transcriptomic approaches. Bacterial and eukaryotic single-cell transcriptomic approaches are shown below and above the timeline, respectively. Technological frameworks to capture and profile transcriptional programmes in individual cells are indicated by colour: grey, microfluidic chip based on cell size selection; purple, FACS; red, microfluidic chip based on droplet encapsulation; blue, split-pool barcoding

based on plates; yellow, spatial transcriptomic methods. Smart-seq, switch mechanism at the 5' end of RNA template sequencing; Smart-seq2, a more sensitive, optimized version of Smart-seq typically coupled with FACS; sci-RNA-seq, single-cell combinatorial indexing RNA sequencing; SPLiT-seq, split-pool ligation-based transcriptome sequencing.

study at the individual cell level to understand the collective functions of entire cell populations.

More recently, single-cell RNA sequencing (scRNA-seq) and spatial transcriptomics have transformed eukaryotic biology by enabling the discovery of cell types and states, reconstruction of developmental trajectories and mapping of tissue architectures during health and disease⁷. Despite the abundance and importance of bacteria, as approximately -13% of Earth's biomass⁸, scRNA-seq methods for bacteria have lagged far behind such approaches in eukaryotes (Fig. 1). This lag is primarily due to fundamental differences between eukaryotic and prokaryotic cells, including: (1) the relative paucity and rapid half-life of mRNA per cell in bacteria compared with eukaryotes⁹; (2) the lack of a molecular handle (for example, a poly(A) tail) to distinguish bacterial mRNA from other transcripts, such as rRNA; and (3) physiologic differences in morphology, size, nucleotide use (GC content) and cell wall composition and thickness that ultimately affect mRNA capture. Collectively, these differences have stymied scRNA-seq approaches in bacteria for many years compared with eukaryotes (Fig. 1). However, recent efforts to overcome these technical challenges have yielded new biological insights, even as additional technical and computational challenges remain to be solved. As progress continues to advance the ability to study bacteria at high resolution, we anticipate a new, exciting era of discovery.

In this Review, we describe the array of methodologies developed to overcome the challenges inherent to bacterial single-cell transcriptomics and the biology that has been unearthed using these approaches. We also describe current technical and computational challenges, suggest a framework for standardization to facilitate comparisons between methods and studies, and define benchmarks for interpreting the quality, reliability and significance of the data generated. Finally, we highlight the variety of questions these approaches can address, foreshadowing how the ability to examine bacterial cells at this higher resolution will expand our understanding of the diversity of bacterial cell types and states and their impact on infection, the microbiota and the environment.

Technological advances enabling bacterial scRNA-seq

To address the technical and computational challenges inherent in bacterial scRNA-seq, investigators have drawn on a range of methodologies to capture and quantify mRNA signal in individual bacterial cells^{10–13}. Technological frameworks to measure transcriptional changes in individual bacterial cells have largely mirrored

approaches previously developed for scRNA-seq in eukaryotic cells, including fluorescence-activated cell sorting (FACS)^{14,15}, droplet encapsulation^{16–19}, split-pooling^{20–22} and microscopy^{23,24} (Table 1). Without a molecular handle to separate mRNA from other RNA in prokaryotic cells, mRNA-capture efforts have mostly involved random priming or hybridization of the mRNA to predesigned probe sets^{10–13}. Finally, given the scarcity in transcripts and genes captured for individual cells, lessons from eukaryotic scRNA-seq studies have revealed that large numbers of cells are often required to discern statistically significant differences among subpopulations²⁵, meaning that efficient, scalable methods with minimal loss are required to enable profiling of thousands of cells in a single experiment^{10,16,18,26}.

mRNA transcript sparsity poses the greatest challenge for single-cell transcriptomics in bacteria. Previous estimates suggest that most bacterial mRNAs have a half-life on the order of minutes⁹, rather than hours compared with eukaryotes, and are present at approximately 0.05–5 copies per gene per cell⁹—nearly 2 orders of magnitude lower than mRNA levels in eukaryotic cells. Similar mRNA levels have been confirmed more recently through spatial transcriptomic analysis of *Pseudomonas aeruginosa* during a growth phase when mRNA is most abundant: the total number of mRNAs averaged fewer than 150 molecules per cell over a 105-gene probe set²³. Hence, the vast majority of bacterial scRNA-seq methods capture only 2–6% of the bacterial genome in each individual cell, which is on par with the percentage of genes captured in eukaryotic scRNA-seq^{10,15–21,27}. A critical lesson learned from eukaryotic scRNA-seq is that mRNA scarcity can be compensated by characterizing large numbers of cells²⁵, with robust conclusions drawn even when only 1–2% of the transcriptome is detected per cell when scale is leveraged^{28–35}. Importantly, scRNA-seq does not directly reveal the behaviour of individual cells, no matter how deep the coverage; it is a statistical sampling strategy that enables characterization of the transcriptional landscape of the population²⁵. Large cell numbers can enable the clustering and identification of distinct subpopulations representing different cell states or cell types, with the behaviours of individual cells interpreted only in the context of this landscape. High-throughput approaches are more critical if studying very rare bacterial cell populations, such as persister or heteroresistant subpopulations^{1,9,36,37} or the phylogenetically diverse microbiome. Thus, bacterial scRNA-seq methods must integrate experimental technologies that are able to efficiently capture the limited mRNA signal across large numbers of cells using computational approaches that can then extract meaningful insight from biologically and technically constrained sparse datasets.

Table 1 | Bacterial single-cell transcriptomic approaches

Method		Maximum throughput	Species agnostic	Instrumentation	Species	rRNA depletion or mRNA enrichment	Advantages and disadvantages
scRNA-seq							
FACS based	MATQ-seq ^{14,15}	10 ²	Yes	FACS	<i>Salmonella enterica</i> <i>P. aeruginosa</i>	CRISPR-Cas9	Suitable for low-input samples; lower throughput than other methods; requires specialized instrumentation (FACS)
Split-pool barcoding based	PETRI-seq ^{20,43,48}	10 ⁴	Yes	None	<i>E. coli</i> <i>S. aureus</i>	CRISPR-Cas9	Lower cost relative to other scRNA-seq methods; no requirement for specialized equipment; high initial upfront investment in barcoded oligonucleotides
	microSPLIT ^{10,21}		Yes	None	<i>E. coli</i> <i>B. subtilis</i> <i>P. putida</i>	PAP	
	BaSSSh-seq ²²		Yes	None	<i>S. aureus</i>	Subtractive hybridization	
Droplet based	BacDrop ¹⁶	10 ⁴ –10 ⁵	Yes	10x Genomics (Chromium)	<i>E. coli</i> <i>P. aeruginosa</i> <i>K. pneumoniae</i> <i>Enterococcus faecium</i>	RNaseH	Higher throughput relative to other scRNA-seq methods; 10x Chromium droplet microfluidics (BacDrop and M3-seq) increases cost relative to split-pool barcoding-based methods; custom microfluidics (smRandom-seq) not yet widely available
	M3-seq ¹⁸		Yes	10x Genomics (Chromium)	<i>E. coli</i> <i>B. subtilis</i>	RNaseH	
	smRandom-seq ^{19,26,38–40}		Yes	Custom microfluidics	<i>E. coli</i> <i>B. subtilis</i> <i>Acinetobacter baumannii</i> <i>K. pneumoniae</i> <i>P. aeruginosa</i> <i>S. aureus</i> <i>Serratia marcescens</i> Human stool and bovine rumen microbiota	CRISPR-Cas9	
Probe-based hybridization	ProBac-seq ¹⁷	10 ³ –10 ⁴	No	10x Genomics (Chromium)	<i>E. coli</i> <i>B. subtilis</i> <i>Clostridium perfringens</i>	None	Avoids the need for rRNA depletion; requires separate probe sets for each species
Spatial transcriptomics							
	par-seqFISH ²³ Bacterial-MERFISH ²⁴	10 ⁵ –10 ⁶	No	Advanced microscope	<i>P. aeruginosa</i> <i>E. coli</i> <i>B. thtaiotaomicron</i>	None	Subcellular and single molecule resolution; avoids the need for rRNA depletion; requires specialized instrumentation and separate probe sets for each species

Cellular isolation and mRNA capture

Recent bacterial scRNA-seq approaches have integrated diverse methodologies to isolate single bacterial cells, capture mRNA signal and increase the throughput of profiled cells^{10–13}. Given bacterial mRNA lacks a poly(A) tail, most methods rely on random priming to capture mRNA from individual cells. For example, multiple annealing and dC-tailing-based quantitative scRNA-seq (MATQ-seq) uses FACS to isolate single bacterial cells into individual microtiter plate wells followed by random priming for mRNA capture and library construction in each well^{14,15} (Fig. 2a). Following the recognition that bacterial cells can themselves be used as reaction ‘vessels’, several plate-based, split-pooling methods such as prokaryotic expression-profiling by tagging RNA in situ and sequencing (PETRI-seq)²⁰, microbial split-pool ligation transcriptomics (microSPLIT)²¹ and bacterial scRNA-seq with split-pool barcoding, second strand synthesis and sub-tractive hybridization (BaSSSh-seq)²² were developed. These approaches first capture mRNA in individual fixed, permeabilized cells via reverse transcription and then uniquely label the transcriptional content of single cells by

adding combinatorial indices through multiple rounds of ligation via splitting and pooling^{20–22} (Fig. 2b). While droplet-encapsulation-based methods such as bacterial droplet-based scRNA-seq (BacDrop)¹⁶, massively-parallel, multiplexed, microbial sequencing (M3-seq)¹⁸ and single-microbe randomly primed RNA-seq (smRandom-seq)^{19,26,38–40} use either commercial (10x; BacDrop and M3-seq)^{16,18} or custom (smRandom-seq)^{19,26,38–40} microfluidic devices, they also use the cell as a reaction vessel to increase throughput^{16,18,26,38–40} (Fig. 2c). Here mRNA in individual fixed, permeabilized cells is first captured through random priming. This step also introduces a single-step barcoding index before droplet encapsulation that enables multiple cells to be encapsulated in the same droplet, thus increasing throughput, in principle, to hundreds of thousands of cells^{16,18,26}. These methods generate scRNA-seq libraries in which the transcriptional content of individual cells is measured by next-generation sequencing.

An alternative method for capturing mRNA from individual cells is hybridization to predesigned probe sets for a specific species. For probe-based bacterial sequencing (ProBac-seq), these probes are

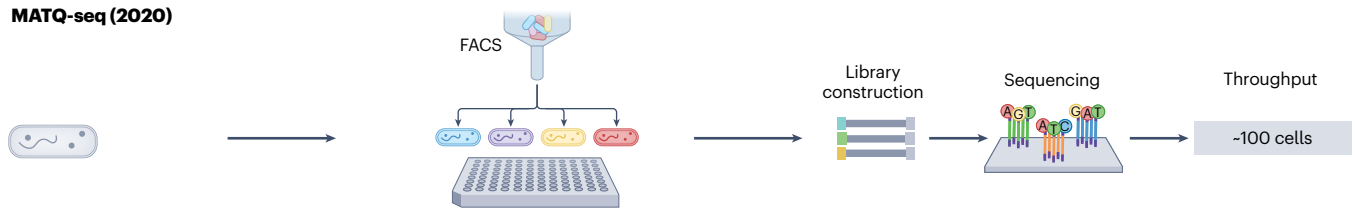
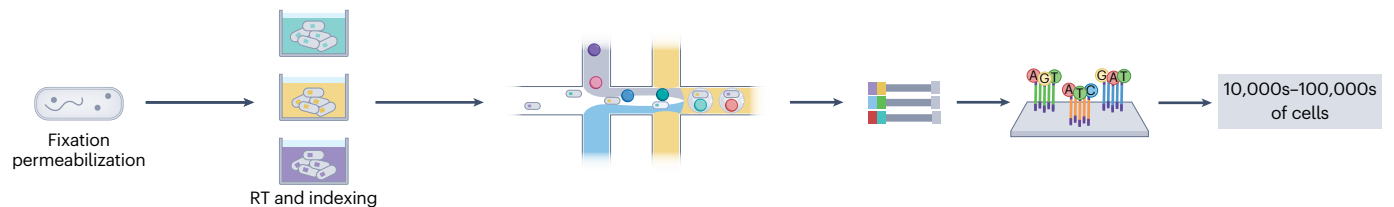
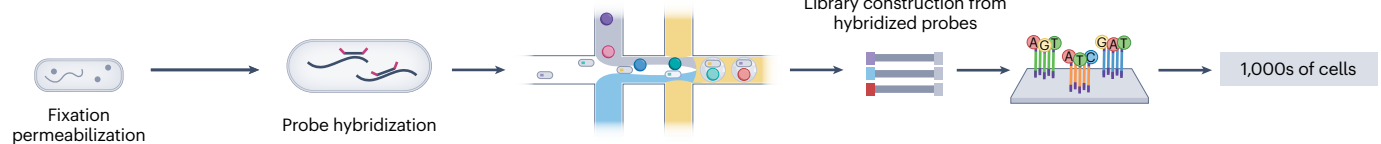
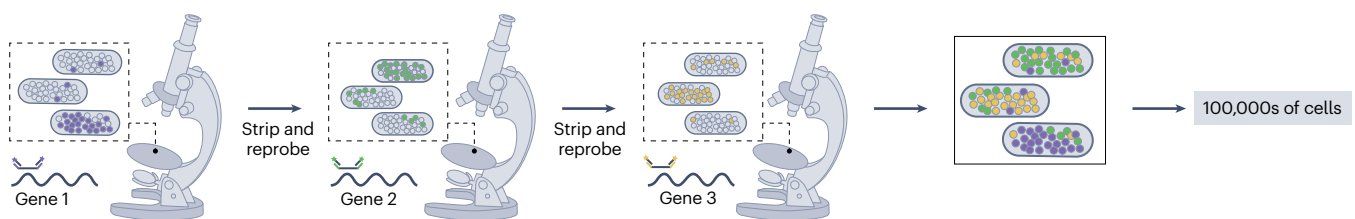
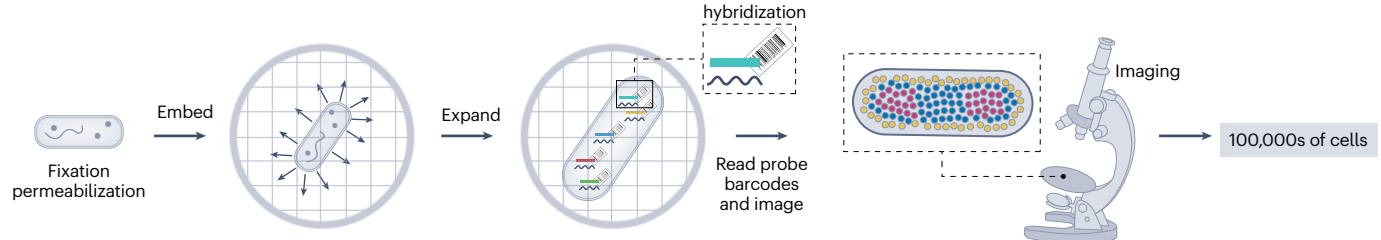
a MATQ-seq (2020)**b PETRI-seq (2020), microSPLIT (2021) and BaSSSh-seq (2024)****c BacDrop, M3-seq and smRandom-seq (2023)****d ProBac-seq (2023)****e par-seqFISH (2021)****f Bacterial-MERFISH (2025)**

Fig. 2 | Bacterial single-cell transcriptomic methods. a–d, Transcriptomes from individual bacterial cells can be captured and profiled by sequencing using: FACS (MATQ-seq) (a), plate-based split-pooling (PETRI-seq, microSPLIT and BaSSSh-seq) (b), microfluidic droplet encapsulation (BacDrop, M3-seq and smRandom-seq) (c) or by sequencing libraries constructed from hybridized probes specific for an organism of interest (ProBac-seq) (d). **e, f**, Alternatively, spatial transcriptomic approaches rely on imaging to profile in situ hybridization

patterns for probe sets specific to an organism of interest, preserving both the spatial and subcellular architecture of gene expression in individual cells. **e**, par-seqFISH uses iterative rounds of in situ hybridization and imaging followed by stripping and re-probing. **f**, Bacterial-MERFISH relies on physical expansion of individual bacterial cells followed by probe hybridization and probe barcode pattern enumeration through imaging. Figure created in BioRender; Lab, H. <https://biorender.com/cq18jhs> (2026).

hybridized to mRNA in bacterial cells *in situ*¹⁷. Cells are then encapsulated in microfluidic droplets, where scRNA-seq libraries are prepared from hybridized probe sets, followed by next-generation sequencing¹⁷ (Fig. 2d). In contrast, spatial transcriptomic methods are based on fluorescence *in situ* hybridization (FISH), in which gene expression changes are determined by hybridization to predesigned probes in fixed, permeabilized cells and captured by imaging rather than sequencing. However, transcriptome-wide changes in expression at single-cell resolution are challenged by the limited number of distinct fluorophores available to detect hundreds to thousands of different probes, and by the ability to then optically resolve and quantify them in a cell the size of a bacterium, given that mRNA density is estimated to be ~8,000 transcripts per 3 μm^3 cell⁴¹. Parallel sequential FISH (par-seqFISH) circumvents this technical challenge through iterative rounds of non-combinatorial probe hybridization, detection and stripping over a select set of 105 marker genes²³ (Fig. 2e). Bacterial-multiplexed error-robust FISH (MERFISH), in contrast, overcomes these problems by using combinatorial barcodes to uniquely label each probe with a distinct pattern of fluorophores and by physically expanding the cell itself over 1,000-fold, thereby decreasing the mRNA density per cell and facilitating the detection of 1,930 *Escherichia coli* operons (~80% of the transcriptome) at single-molecule resolution²⁴ (Fig. 2f). Importantly, both par-seqFISH and bacterial-MERFISH enable hundreds of thousands of cells to be analysed in a single experiment while preserving the spatial architecture of transcriptional changes with cellular and subcellular resolution^{23,24}.

Both random-priming and probe-based methods of mRNA capture have strengths and weaknesses (Table 1). Random priming can, in principle, be used to capture mRNA from any organism without the need for specialized probe sets for each species or a priori knowledge of genomes for probe-set design. However, it also captures rRNA transcripts, which constitute the vast majority (80–95%) of total RNA, so technical and computational methods for rRNA removal are required to prevent their predominance in sequencing reads. Further, random priming can be subject to sequence bias in binding affinity, due to GC content or secondary structure, which may result in misrepresentation of transcript abundance. When different combinatorial indices are fused to the random primer sequences, the sequence of the indices can themselves introduce additional bias to priming⁴². Conversely, probe-based hybridization approaches can be very sensitive and avoid rRNA signal; however, obvious drawbacks are that prior knowledge of gene expression patterns for a species of interest is required for probe-set design, creating a potential for bias if only a subset of genes can be queried, and a separate probe set is required for each species. Probe-based hybridization approaches also rely on accurate open reading frame annotations, which are often incomplete or inconsistent in non-model species and complicated by substantial strain-to-strain variation in gene content. These methods also do not necessarily scale well to complex communities because probes may hybridize non-specifically to homologues rather than to their intended targets.

Maximizing mRNA signal

Without efficient methods to enrich mRNA or deplete rRNA for random-priming approaches, most reads in the final library map to rRNA (80–95%), obscuring important variation in protein-coding genes and increasing overall sequencing costs. Preferential mRNA capture efficiency can be enzymatically enhanced through the use of *E. coli* poly(A) polymerase (PAP), which preferentially polyadenylates mRNA, followed by the use of oligo(dT) for mRNA capture, resulting in a modest 2.5-fold enrichment²¹ (Fig. 3f). Alternatively, rRNA molecules can be depleted using RNase H^{16,18}, CRISPR–Cas9^{14,19,43} or subtractive hybridization²² to increase mRNA recovery from 5% to 50–90% (RNase H)¹⁶, 16–63% (CRISPR–Cas9)^{14,19} or ~50–92% (subtractive hybridization)^{22,44} at various stages (Fig. 3). For example, RNase H can be used to digest rRNA before mRNA capture (BacDrop)¹⁶ (Fig. 3a) or following

cDNA amplification (M3-seq)¹⁸, although this latter approach requires an additional step, transcribing the double-stranded cDNA sequence into single-stranded RNA for RNase H digestion (Fig. 3b). CRISPR–Cas9 and subtractive hybridization-based rRNA depletion have also been used post cDNA amplification^{14,19,22,44} (Fig. 3c,e) or after final library construction⁴³ (Fig. 3d). Given the importance of minimizing loss of mRNA signal, rRNA-depletion approaches that rely on multi-enzymatic steps or physical separation may be prone to more mRNA signal loss, whereas *in situ* rRNA depletion may decrease mRNA capture efficiency^{18,21}. While these approaches have similar upfront costs, most methods have not been compared head to head, hence the optimal method that maximizes rRNA depletion while preserving mRNA signal remains to be determined.

Computational approaches for cell state discovery

Although mRNA recovery is low due to both intrinsic biology and technical loss, the fraction of the bacterial transcriptome that is recovered is similar to that of eukaryotic systems. Bacterial scRNA-seq analyses have adopted computational workflows originally designed for eukaryotes, ranging from standard clustering pipelines (for example, Scanpy⁴⁵ and Seurat⁴⁶) to deep generative models that denoise gene expression and account for technical variation, such as scVI⁴⁷. These latter models, which learn low-dimensional representations that capture gene–gene dependencies, have shown the most potential, recovering reproducible transcriptional states even from highly sparse bacterial single-cell data⁴⁷. Crucially, these methods must be run with parameters tuned for sparse bacterial data. This includes cell-quality thresholds far below those typically used in eukaryotic analyses (Fig. 4), which in many cases involves retaining cells with only 10 detected transcripts^{16,18,48}. Nevertheless, these computational methods successfully extract signals from sparse datasets by capitalizing on the core innovation of single-cell biology: leveraging scale. Even with ultra-sparse measurements, pooling thousands of transcriptionally similar cells can reveal reproducible gene expression programmes and cell state distributions.

A major unresolved challenge is identifying shared transcriptional states across samples and experimental conditions. In practice, bacterial cells often cluster by experimental condition rather than underlying state, making it difficult to identify corresponding subpopulations across datasets. While standard clustering approaches have uncovered a limited number of distinct cellular states—for example, mobile genetic element induction and stress responses^{16,18,19,21}—more subtle transcriptional states frequently fail to align across experiments, limiting cross-condition comparisons.

Single-cell harmonization frameworks^{47,49,50} have been widely adopted in eukaryotic systems to address this problem by aligning cells from multiple samples, conditions or tissues. While they could serve as a model for bacterial systems, they remain more challenging to benchmark for bacteria without established ground-truth cell states. Recent approaches have aligned bacterial cells by learning a shared latent space, either by training a reference model and projecting new data into that space^{43,48} or by joint embedding single cells across all samples⁴⁸. By doing so, they can recover high-resolution insights into the gene regulatory architecture of bacteria⁴⁸. However, careful calibration is required as they may struggle to find shared axes or they may be too aggressive, potentially forcing distinct transcriptional states to merge artificially.

Biological insights enabled by scRNA-seq

While studies^{14–24,26,43,48,51} have mostly focused on a relatively limited number of bacterial species so far, scRNA-seq analyses^{14–21,23,43,51} have already revealed important insights into bacterial physiology, metabolism, virulence, phage induction, plasmid carriage and antibiotic resistance. These studies have characterized bacterial heterogeneity within static populations, in response to perturbations, across variable microenvironments, in the context of polymicrobial microbiota

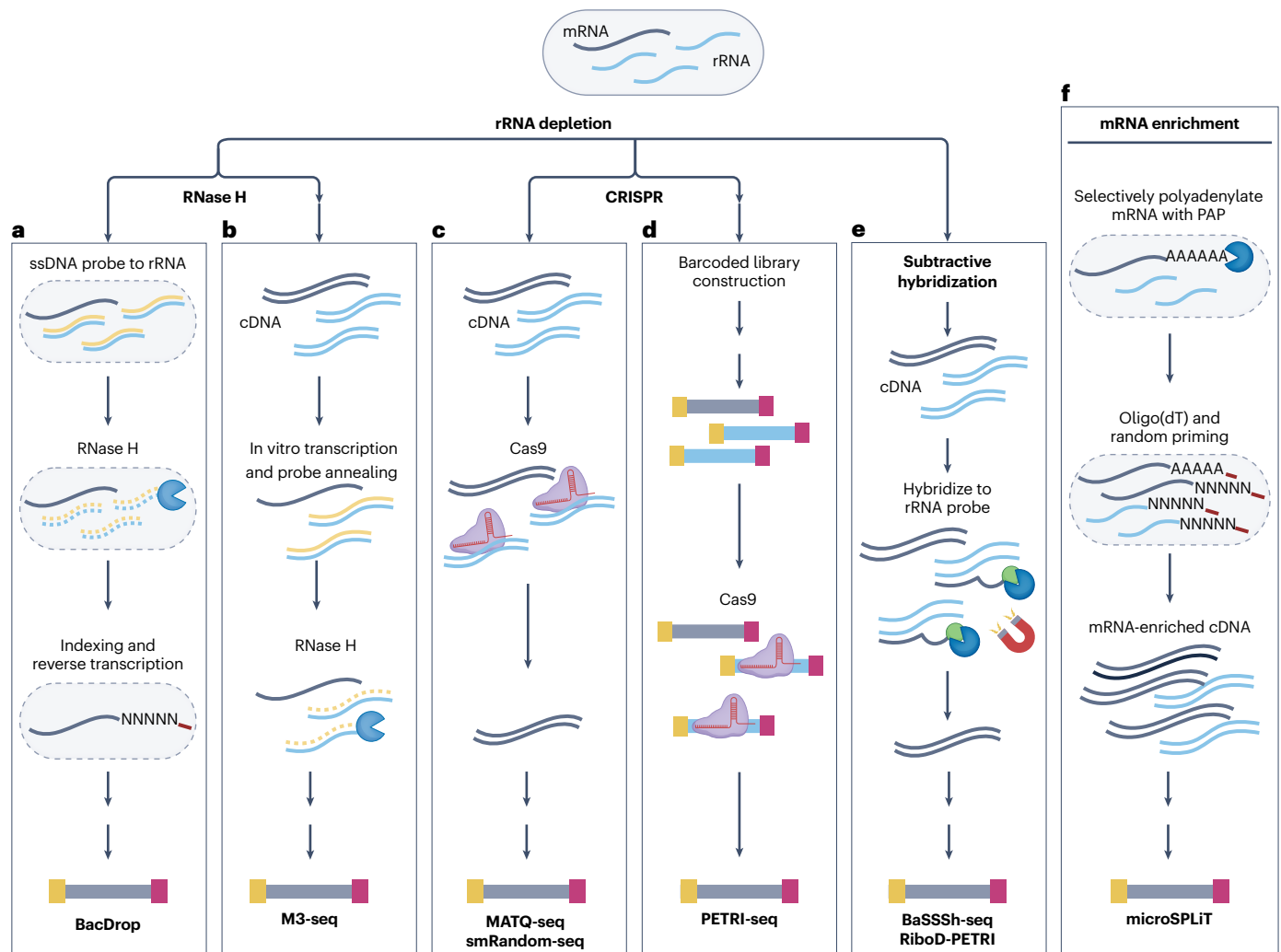


Fig. 3 | rRNA depletion and mRNA enrichment strategies. a, b, rRNA or cDNA derived from rRNA can be depleted from samples or libraries using RNase H by digesting rRNA in permeabilized cells in situ before mRNA target capture (BacDrop) (a) or following library construction via addition of a T7 in vitro transcription step and subsequent digestion of rRNA–DNA hybrids with RNase H (M3-seq) before final library construction (b). **c, d,** Alternatively, rRNA-derived cDNA (MATQ-seq and smRandom-seq) (c) or final library sequences (PETRI-seq) (d) can be cleaved using CRISPR–Cas9 and single-guide RNAs targeting rRNA. **e,** Finally, rRNA-derived cDNA species can be removed by subtractive

hybridization using biotinylated rRNA probes that, following hybridization to rRNA-derived cDNA, are separated from the rest of the sample using magnetic streptavidin beads (BaSSSh-seq and ribosomal RNA-derived cDNA depletion PETRI-seq (RiboD-PETRI)). **f,** As an alternative to rRNA depletion, mRNA species can be enriched during target capture by first selectively polyadenylating mRNA in permeabilized bacterial cells in situ using PAP followed by priming with oligo(dT) in addition to random hexamers during reverse transcription (microSPLIT). ssDNA, single-stranded DNA. Figure created in BioRender; Lab, H. <https://biorender.com/wcdtbpz> (2026).

communities^{26,38} and during bacterial interactions with mammalian cells^{22,24}. For example, these studies have shown that static populations (cultures) of bacteria previously thought to be homogeneous are in fact heterogeneous in terms of competence²¹, virulence factor expression^{17,23}, pro-phage induction^{16,18,20,21} and carbon and arginine metabolism^{17,21}. Further, small differences between closely related strains can also impact transcriptional heterogeneity. In *Pseudomonas putida*, for example, cells carrying a broad-host-range plasmid partitioned into distinct clusters based on conjugation gene expression, suggesting that plasmid carriage itself can drive heterogeneity within a population⁵¹.

scRNA-seq analyses have revealed aspects of bacterial physiology that were not possible from bulk measurements. Recent work⁴⁸ showed that gene dosage, as driven by DNA replication, has a predictable impact on gene expression patterns on a genome-wide scale for many genes in rapidly growing bacteria (that is, expression of many genes on the chromosome is related to their replication). Further, scRNA-seq data and chromosomal gene position can be integrated to

infer the replication stage of individual cells, the origin and terminus of replication, and biophysical parameters such as the speed of DNA and RNA polymerases in different species⁴⁸. The finding that gene expression changes are influenced by DNA replication stage, at least during rapid growth, establishes a critical baseline to understand when individual genes depart from this trend⁴⁸. In *Staphylococcus aureus*, for example, genes of mobile genetic elements fall into this latter class⁴⁸, while in static populations of *Klebsiella pneumoniae*, heterogeneity is driven by transposable elements¹⁶.

These initial studies have provided reference points from which to understand how population dynamics change in response to perturbations or upon interactions between species. For example, whereas bulk RNA-seq measurements have shown little evidence of heterogeneity in response to β -lactam antibiotic treatment^{16,19}, scRNA-seq suggests otherwise. Treatment of *K. pneumoniae* with the β -lactam, meropenem, results in four distinct subpopulations characterized by the induction of genes involved in stress responses, cell wall membrane synthesis, DNA replication and *cspD*—a toxin that inhibits DNA replication and

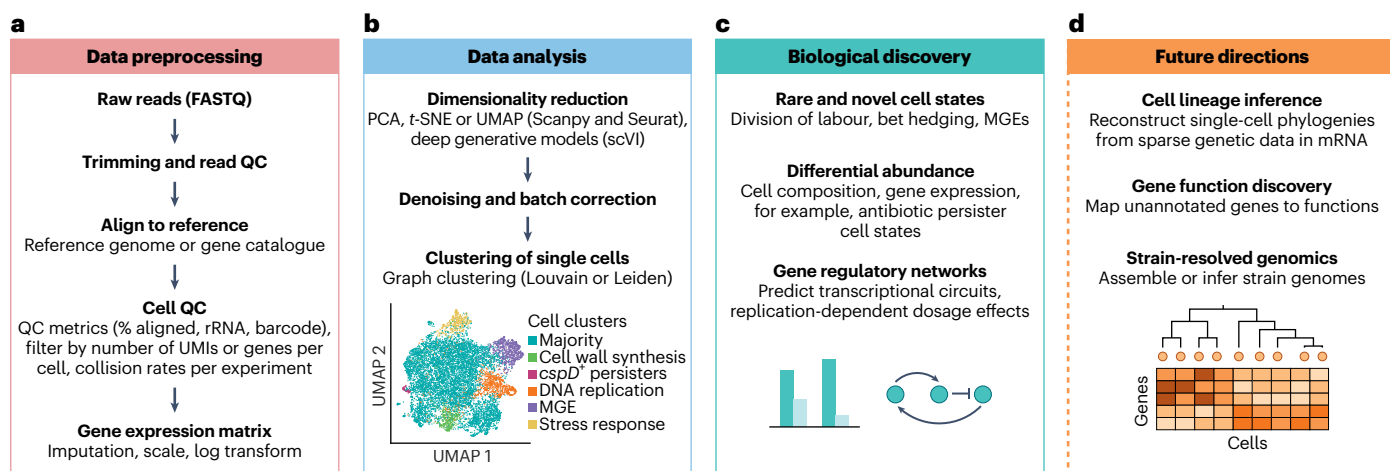


Fig. 4 | Computational analysis pipeline of bacterial scRNA-seq. **a**, Data first undergo preprocessing, in which sequenced reads are trimmed, quality filtered and aligned to a reference genome or a metagenomic gene catalogue. Quality metrics are used to assess data quality, estimate barcode collision rates and filter cells to generate a digital gene expression matrix from high-quality cells that pass the defined quality metrics. The gene expression matrix is then imputed, scaled and log transformed. **b**, These high-dimensional gene expression data are then analysed by embedding into a lower dimensional space, where they are denoised, batch corrected and ultimately clustered into cell types and cell states. **c**, Cell-centric clustering can reveal new cell types or cell states, while

differential abundance tests can reveal changes in these populations or in their gene expression levels that are associated with a condition, leading to biological discovery. Gene-centric analyses of their regulatory networks can also reveal bacterial transcriptional circuits. **d**, Future directions for computational analyses may include the inference of single-cell lineages, the discovery of new gene functions and leveraging cell barcodes to resolve strain genomes. MGE, mobile genetic element; PCA, principal component analysis; *t*-SNE, *t*-distributed stochastic neighbour embedding; UMAP, uniform manifold approximation and projection; UMI, unique molecular identifier.

induces meropenem tolerance (persistence)¹⁶. scRNA-seq is also beginning to reveal heterogeneity of responses in more complex settings, such as in interactions between bacteria and different host cells or within complex microbial communities. For example, when *S. aureus* biofilm cells are co-cultured with macrophages, genes associated with adaptation to the presence of reactive oxygen species are induced, whereas when they are co-cultured with polymorphonuclear phagocytes, cell wall maintenance and virulence genes are induced²². In the bovine rumen, scRNA-seq identified a role for *Basifia succiniciproducens* in carbohydrate metabolism²⁶, foreshadowing the impact scRNA-seq may have in deconvolving heterogeneity in responses within complex microbial communities.

The studies described earlier were conducted on dissociated cells and thus did not address the spatial organization of cellular heterogeneity within a population or distribution of transcripts within a single cell. Spatial transcriptomic approaches such as par-seqFISH and bacterial-MERFISH are beginning to reveal both the intracellular spatial organization of bacterial transcriptomes and heterogeneity within different microenvironments and niches^{23,24}. For example, in *E. coli*, transcript subcellular localization for many operons was shown to be governed either by protein function, particularly for operons that encode inner membrane proteins, or by chromosomal position, where transcripts from operons positioned near the replication terminus localize near the terminus itself²⁴. At the cellular level, par-seqFISH revealed subpopulations of cells expressing genes associated with anaerobic metabolism, oxidative stress, denitrification and heat-shock proteases co-existing alongside but mutually excluded from adjacent subpopulations characterized by expression of tricarboxylic acid cycle genes²³. These subpopulations suggest differences in oxygen tension within local microenvironments and evidence of distinct physiological states and metabolic heterogeneity within single-species biofilms²³. It has also shown how microenvironment affects cell states in host–bacteria interactions. For example, bacterial-MERFISH revealed that *Bacteroides thetaiotaomicron*, a gut commensal, shows spatial heterogeneity in polysaccharide utilization loci (PUL) expression: cells proximal to the mucus layer express PULs

targeting host-mucus polysaccharides, whereas cells occupying the lumen express PULs targeting dietary polysaccharides²⁴.

Current challenges

Despite the technological advances and resulting biological insights previously discussed, much work is still needed for bacterial scRNA-seq to reach its full potential. Methods need to increase sensitivity and throughput, and solutions are needed for the computational challenges of managing the unique biology and genetic organization of bacteria compared with eukaryotes. Standardization is also needed, including quality control (QC) metrics similar to those established for eukaryotic scRNA-seq. Without such metrics, it is difficult to compare bacterial scRNA-seq methods, evaluate gains from technical and computational advances and, perhaps most importantly, assess the quality and rigour of data from which new biological insights and conclusions are drawn. A standardization framework is thus needed, in parallel to current technological and computational efforts, to improve and democratize this currently complex technology while reducing costs to enable its more widespread application.

Technical challenges

Further improvements are needed to increase the sensitivity and throughput of sequence-based scRNA-seq, both in terms of the number of detected genes and cells analysed, to more fully characterize distinct and potentially rare cell types and states. Progress in throughput has already been achieved through both combinatorial indexing and droplet microfluidics, which capture tens of thousands and even hundreds of thousands of cells^{16,18,20,21,26}; however, further improvements to mRNA capture are still required. Each step of an scRNA-seq protocol is open to further optimization to increase efficiency, including the steps of reverse transcription, mRNA enrichment and/or rRNA depletion, and sample handling, barcoding, library construction and amplification. These methodologies are elaborate multi-step protocols in which cells or RNA can be lost at each centrifugation or washing step and the molecular complexity of the library diminishes with each amplification cycle. Therefore, optimizing and streamlining protocols

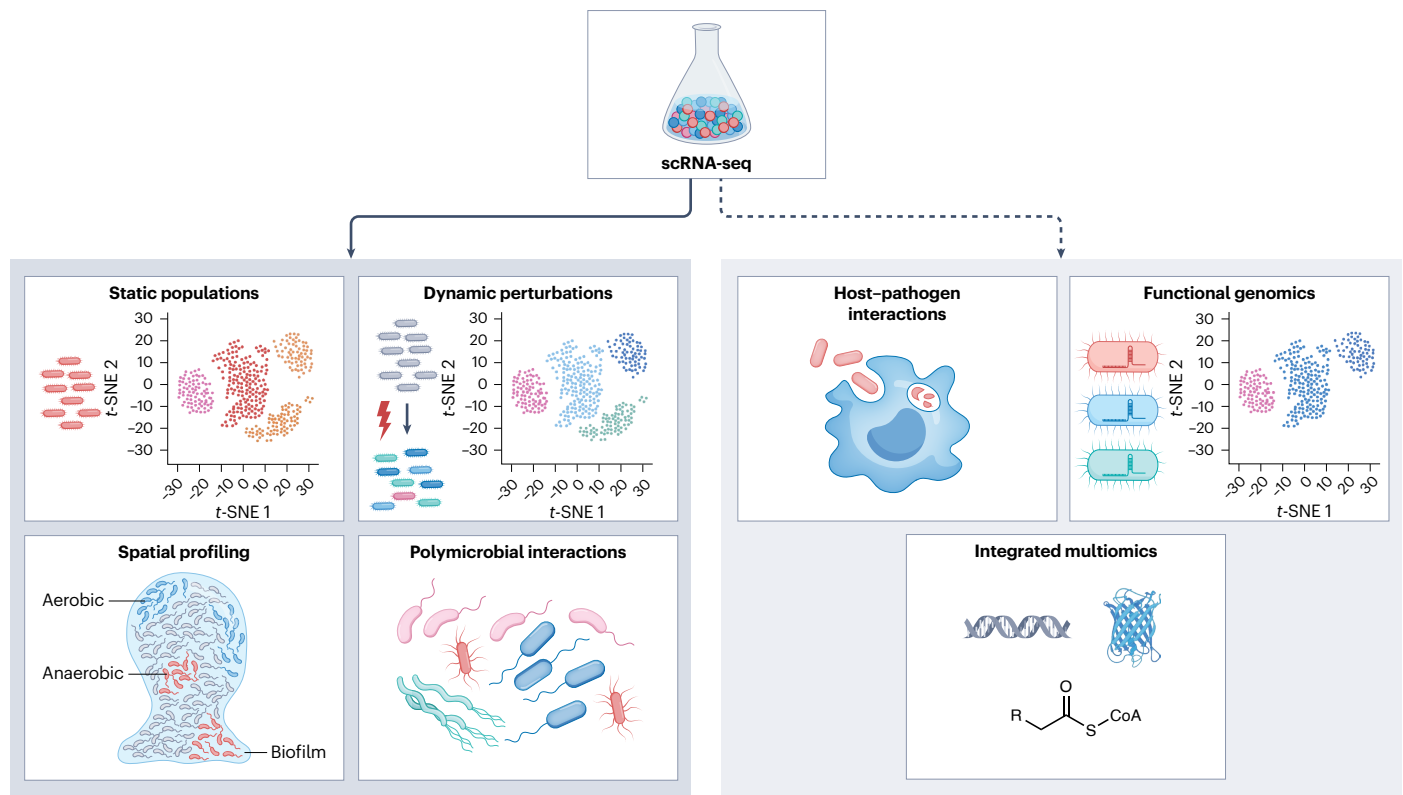


Fig. 5 | Current and future applications of bacterial scRNA-seq. Bacterial scRNA-seq can currently be used to determine subpopulation differences in cell states in both static and dynamically perturbed populations of cells, the three-dimensional architecture of cell types and states (spatial transcriptomics) and interspecies polymicrobial interactions. Future directions for these technologies

could include more functional genomic studies (for example, Perturb-seq), host–pathogen interactions at single-cell resolution and integrated multi-omic approaches. Figure created in BioRender; Lab, H. <https://biorender.com/e4uh5zv> (2026).

to minimize RNA loss is important and may necessitate automation rather than manual handling, as was recently done for MATQ-seq¹⁴. In particular, increasing the percentage of reads mapping to mRNA in the final library, as opposed to mapping to rRNA, will help to democratize this technology by reducing sequencing costs, which have also decreased dramatically.

Another major challenge unique to bacteria is the variable access to mRNA content in each cell due to differences in cell wall composition, which has implications for cell permeabilization efficiency. When the cell itself is used as a reaction vessel, all cells in a population must be permeabilized in a way that leaves the vessel intact but gives free access to all reagents without allowing RNA to be degraded or to diffuse away, resulting in the mixing of RNA between cells or loss of signal. Hence, permeabilization is critical as variability can lead to a biased representation of the community or enhanced ambient RNA contamination. Currently, this step must be tailored to each individual experiment due to species- and strain-specific variations in cell wall components, such as LPS, peptidoglycan, teichoic acids or polysaccharide capsule^{52,53}. Even growth conditions and environment can complicate matters further, for example, L-form bacteria, which have no cell wall, can develop under antibiotic or osmotic stress^{54,55} and may be more susceptible to detergents but more resistant to permeabilization reagents such as lysozyme^{54,55}. Most concerning is variability in cell wall composition, which may result in permeabilization differences across different subpopulations within the same sample, as this could yield a distorted view of the phenotypic landscape. While this could occur in single-species analyses, it certainly does occur in complex populations, such as the microbiota or the environment, thus necessitating solutions to overcome this challenge.

Finally, for complex, mixed populations of bacterial cells, even uniform isolation and capture of different strains in a mixed population can be challenging. For most scRNA-seq methods, cellular isolation relies on multiple centrifugation steps; however, species differ in their sedimentation coefficients and therefore do not pellet under the same centrifugal conditions. In a mixed population of bacteria, this can result in both sample loss¹⁶ and bias. Solutions for these issues will be critical to enable application of bacterial scRNA-seq to complex, mixed populations such as the microbiota. Nevertheless, although these methods do have clear technological challenges and caveats, they have enabled higher-resolution transcriptional profiling of different cell types and states in single and mixed communities.

Computational challenges and potential solutions

Data analysis is a critical step for all scRNA-seq approaches. However, bacterial data pose particular challenges that current computational methods struggle to handle, such as mRNA sparsity, phylogenetic diversity within complex communities and the unique organization of gene expression programmes, which are less amenable to cell clustering-based analyses. These challenges may require us to devise new computational approaches specifically tailored to bacteria.

The issue of mRNA sparsity is mitigated in eukaryotic analyses by pooling thousands of transcriptionally similar cells, enabling inference across populations of cells. However, the even greater sparsity for bacterial scRNA-seq pushes the limits of these approaches. One possible solution is to impute ‘missing’ gene expression values, for example, by leveraging the structure of operons that encode genes as polycistronic RNA. Another strategy is to exploit the vast datasets now attainable with combinatorial indexing or droplet barcoding and use deep generative models to learn latent representations that

capture important transcriptional features (for example, cell types and regulatory networks) while discarding unwanted technical variation, such as noise, library size and batch effects. This approach was used to uncover the subtle impact of gene dosage on bacterial transcription and to identify operons and gene regulatory networks⁴⁸.

Even with reliable estimates of the transcriptional similarities among cells, most analyses focus on clustering these cells into transcriptionally distinct cell types or cell states. However, this framework may not directly translate to bacteria due to their distinct regulatory architectures. Unlike multicellular organisms, where tightly orchestrated developmental programmes often give rise to recurrent and discrete cellular identities, bacteria often execute overlapping responses to local environmental conditions, growth state, stress conditions and stochastic gene expression. As a result, cell clustering-based analyses that do not incorporate these decentralized features of bacterial cells may fail to resolve their cellular states. Thus, the objective for bacterial analyses may extend beyond 'cell-centric' analyses to direct modelling of their gene regulatory networks.

Gene-centric models, such as consensus non-negative matrix factorization, represent each single-cell transcriptome as a mixture of distinct gene expression programmes. However, in extremely sparse bacterial datasets, where cells contain tens of detected genes, it may prove to be challenging to reliably estimate these programmes and their proportions across cells. Future developments could leverage the unique features of bacterial regulation, such as operons, to impute missing expression, providing a more complete view of the bacterial transcriptome. In addition, rather than focusing on ultra-sparse single-cell transcriptomes, future work could focus on gene–gene relationships, which can be robustly inferred by pooling data across cells.

Finally, bacteria evolve more rapidly than eukaryotes and acquire new genes via horizontal gene transfer, with the result that closely related strains of *E. coli* can vary by over 50% of their gene content⁵⁶. While previous work has focused on isogenic populations of bacteria, applications to natural isolates or complex communities such as the gut microbiota may lack suitable reference transcriptomes, which will be needed for the alignment and quantification of transcripts within individual cells. One solution is to map reads onto catalogues of genomes or metagenome-assembled genomes for the human and environmental microbiota, as was recently done with the bovine rumen²⁶. However, even this approach may fail to fully capture reference genes, particularly for disease-associated strains that are not well represented in reference databases. One powerful alternative approach is to perform de novo assembly of genomes and transcriptomes, leveraging the cellular barcodes to link single cells to their genes, transcripts and single nucleotide polymorphisms.

Standardization moving forwards

Given the breadth of existing approaches, as the field of bacterial scRNA-seq grows and becomes more widely applied, we propose that standardization and transparency will be critical. Although a range of bacterial scRNA-seq technologies have now been reported, it is currently difficult to know the advantages and disadvantages of each. For now, methods remain fragmented, operating in isolation within individual laboratories, preventing comparison across technologies in controlled settings. The field must adopt standardized methodologies to enable comparisons of established methods. Coordinated efforts to implement and compare methods across institutes would better identify strengths and weaknesses of methods, following examples set by the successful consortium models established for the Human Microbiome Project and the Human Cell Atlas. In the current landscape, a major challenge is the use of disparate model organisms, ranging from *E. coli* to *P. aeruginosa* and *Bacillus subtilis*, across a wide range of growth conditions with no way to efficiently assess performance across methodologies without costly experiments. Ideally, the field would establish consensus model systems (for example, *E. coli* and

B. subtilis) and even specific strains, including phylogenetically diverse species that display abundant, well-defined cell types (for example, motile and sessile populations of *B. subtilis*⁵⁷) detectable with only minimal sequencing coverage. Eukaryotic methods succeeded by benchmarking against standardized tissues and cell types (for example, peripheral blood mononuclear cells for single-cell methods and brain for single-nucleus RNA-seq methods). Unlike eukaryotes, however, bacteria lack distinct cell types present at defined frequencies. Going forwards, it will be critical to establish this ground truth for bacteria in several model organisms grown under standard conditions that elicit consistently predictable cell populations at defined frequencies for benchmarking. From there, one potential approach is to mix different model species, grown under conditions that produce these well-defined populations, at varying proportions. Such samples, if standardized in the field, should enable the rapid assessment of single-cell purity, taxonomic coverage and transcriptional state recovery, while providing a standardized framework for platform evaluation. Finally, we suggest that studies that help to empirically determine such benchmarking standards should be viewed as foundational rather than incremental.

Given the immense variability of bacterial species and growth conditions examined between studies and the current lack of benchmarking standards by which to evaluate assay performance, we suggest studies should incorporate within-experiment controls to assess single-cell quality per experiment. We propose collision rate determinations to ensure single-cell resolution in individual experiments or microscopy, all the more important for bacteria that are prone to clumping, and the inclusion of biological replicates to remove technical artefacts. We also suggest that heterogeneity between subpopulations determined using scRNA-seq should be validated using orthologous methods such as RNA-FISH or flow cytometry as standard practice.

Currently, even with standardized methods, it remains difficult to directly compare the different technologies because QC metrics have not yet been standardized for this rapidly evolving field. For example, cell quality (that is, numbers of cells and numbers of detected transcripts per cell) is highly dependent on (1) sequencing depth, (2) cell-quality cut-offs, (3) bacterial species studied and their growth phases, (4) rRNA depletion, and (5) throughput. Moreover, confounding factors such as genomic DNA contamination, which is challenging to assess in bacterial cells due to their high coding densities, can erroneously inflate such metrics.

To address these challenges, reporting of standardized QC metrics that are measured across multiple quality thresholds (for example, 10, 25, 50, 100 and 250 detected genes per cell) should be required. Essential reported metrics for each study should include: cell recovery rates, transcripts and unique molecular identifiers detected per cell, multiplet rates, RNA collision rates, rRNA percentage and discarded reads. These should be reported both as raw totals and with sequencing depth normalized via downsampling to enable fair cross-study comparisons. This standardized framework would enable effective comparisons across single-cell technologies, while highlighting unique advantages (for example, cell quality, cell recovery and taxonomic capture) and an assessment of the quality of the data from which biological conclusions are drawn.

Future directions

While innovation and refinement are still clearly needed, we now have the tools to begin defining cell types and states in static populations and to follow dynamic perturbations, with spatial resolution and in mixed communities (Fig. 5). Moving forwards, given the broad range of applications of eukaryotic single-cell transcriptomics, it is easy to imagine that bacterial single-cell transcriptomics will similarly yield numerous insights into bacterial behaviour. One obvious new direction is to leverage bacterial single-cell transcriptomics for more functional studies (Fig. 5), in an equivalent of Perturb-seq⁵⁸, where the impact of knockdown or knockout of specific genes on gene expression can be

measured in pools, using CRISPR or transposon-mediated systems. With this type of technology, one can imagine mapping the role of individual genes in various cellular processes, mapping regulatory networks and even examining how knockdown affects subpopulation dynamics at the genome-wide scale.

While dual RNA-seq enabled the simultaneous profiling of interacting host and pathogen^{39,60}, averaging over bulk bacterial populations masks the heterogeneity underlying bacterial behaviours during infection and the role of specific subpopulations in infection phenotypes. scRNA-seq could enable the definition of bacterial subpopulations that differentially contribute to infection severity, invasion, latency or persistence in response to antibiotic treatment. This heterogeneity may be particularly critical during infection where a pathogen is likely to encounter fluctuating environments and host responses throughout the course of infection. One current limitation is the relatively large number of input bacteria required to manage losses from cell collection to final library preparation to yield sufficient numbers at the end of the pipeline for analysis¹⁰. While some experiments of this type may be possible with infected cells in culture or from experimental animals where material is less limiting, not all patient samples or infection models may yet be amenable. The hope is that with further optimization and technological improvements, such as has been recently reported using hydrogel beads⁶¹, experiments of this type and resolution will be possible. Then, single-cell dual RNA-seq to characterize the heterogeneous responses occurring during individual encounters between a bacterium and a host cell^{40,62–64} will be achievable, including spatial transcriptomic characterization of the host–pathogen and host–microbiota interactions⁶⁵, for example, in tissues, biofilms or granulomas.

Finally, a future with more integrated multi-omic analysis including bacterial scRNA-seq will enable a more comprehensive understanding of cell states and functions. For example, by simultaneously analysing both RNA and protein expression—particularly of cell surface proteins using DNA-barcoded antibodies—analogue to CITE-seq (cellular indexing of transcriptomes and epitopes by sequencing)⁶⁶, we may gain further insight into bacterial phenotypes such as immune evasion, virulence, environmental sensing and antibiotic responses. Further coupling of spatial single-cell transcriptional methods with advances in imaging mass spectrometry to characterize metabolites or peptides and proteins in microbiological samples⁶⁷ portends a future of understanding microbiology at a resolution that transcends currently observable phenotypes.

The early studies described here have laid foundations to both establish methodologies and begin to define the depth and breadth of bacterial behaviour. How this heterogeneity contributes to complex phenotypes—such the evolution of antibiotic resistance, the establishment of chronic infection, infection outcome, colonization and maintenance of the microbiota, or the formation of complex microbial communities in the environment—is an open area now ripe for further exploration. Thus, this early work on bacterial single-cell transcriptomics has brought us to an inflection point in bacterial research.

References

- Ackermann, M. A functional perspective on phenotypic heterogeneity in microorganisms. *Nat. Rev. Microbiol.* **13**, 497–508 (2015).
- Solopova, A. et al. Bet-hedging during bacterial diauxic shift. *Proc. Natl Acad. Sci. USA* **111**, 7427–7432 (2014).
- Balaban, N. Q., Merrin, J., Chait, R., Kowalik, L. & Leibler, S. Bacterial persistence as a phenotypic switch. *Science* **305**, 1622–1625 (2004).
- Beaumont, H. J., Gallie, J., Kost, C., Ferguson, G. C. & Rainey, P. B. Experimental evolution of bet hedging. *Nature* **462**, 90–93 (2009).
- Freed, N. E. et al. A simple screen to identify promoters conferring high levels of phenotypic noise. *PLoS Genet.* **4**, e1000307 (2008).
- Norman, T. M., Lord, N. D., Paulsson, J. & Losick, R. Memory and modularity in cell-fate decision making. *Nature* **503**, 481–486 (2013).
- Regev, A. et al. The Human Cell Atlas. *eLife* **6**, e27041 (2017).
- Bar-On, Y. M., Phillips, R. & Milo, R. The biomass distribution on Earth. *Proc. Natl Acad. Sci. USA* **115**, 6506–6511 (2018).
- Taniguchi, Y. et al. Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single cells. *Science* **329**, 533–538 (2010).
- Gaiser, K. D. et al. High-throughput single-cell transcriptomics of bacteria using combinatorial barcoding. *Nat. Protoc.* **19**, 3048–3084 (2024).
- Homberger, C., Barquist, L. & Vogel, J. Ushering in a new era of single-cell transcriptomics in bacteria. *Microlife* **3**, uqac020 (2022).
- Walls, A. W. & Rosenthal, A. Z. Bacterial phenotypic heterogeneity through the lens of single-cell RNA sequencing. *Transcription* **15**, 48–62 (2024).
- Pountain, A. W. & Yanai, I. Dissecting microbial communities with single-cell transcriptome analysis. *Science* **389**, eadp6252 (2025).
- 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).
- 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).
- Ma, P. et al. Bacterial droplet-based single-cell RNA-seq reveals antibiotic-associated heterogeneous cellular states. *Cell* **186**, 877–891.e814 (2023).
- McNulty, R. et al. Probe-based bacterial single-cell RNA sequencing predicts toxin regulation. *Nat. Microbiol.* **8**, 934–945 (2023).
- 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).
- Xu, Z. et al. Droplet-based high-throughput single microbe RNA sequencing by smRandom-seq. *Nat. Commun.* **14**, 5130 (2023).
- 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).
- Kuchina, A. et al. Microbial single-cell RNA sequencing by split-pool barcoding. *Science* **371**, eaba5257 (2021).
- Korshoj, L. E. & Kielian, T. Bacterial single-cell RNA sequencing captures biofilm transcriptional heterogeneity and differential responses to immune pressure. *Nat. Commun.* **15**, 10184 (2024).
- Dar, D., Dar, N., Cai, L. & Newman, D. K. Spatial transcriptomics of planktonic and sessile bacterial populations at single-cell resolution. *Science* **373**, eabi4882 (2021).
- Sarfatis, A., Wang, Y., Twumasi-Ankrah, N. & Moffitt, J. R. Highly multiplexed spatial transcriptomics in bacteria. *Science* **387**, eadr0932 (2025).
- Tanay, A. & Regev, A. Scaling single-cell genomics from phenomenology to mechanism. *Nature* **541**, 331–338 (2017).
- Jia, M. et al. Single-cell transcriptomics across 2,534 microbial species reveals functional heterogeneity in the rumen microbiome. *Nat. Microbiol.* **9**, 1884–1898 (2024).
- Kharchenko, P. V. The triumphs and limitations of computational methods for scRNA-seq. *Nat. Methods* **18**, 723–732 (2021).
- Wang, X., He, Y., Zhang, Q., Ren, X. & Zhang, Z. Direct comparative analyses of 10X Genomics Chromium and Smart-seq2. *Genom. Proteom. Bioinform.* **19**, 253–266 (2021).
- Zhang, X., Xu, C. & Yosef, N. Simulating multiple faceted variability in single cell RNA sequencing. *Nat. Commun.* **10**, 2611 (2019).

30. Wu, A. R. et al. Quantitative assessment of single-cell RNA-sequencing methods. *Nat. Methods* **11**, 41–46 (2014).
31. Pollen, A. A. et al. Low-coverage single-cell mRNA sequencing reveals cellular heterogeneity and activated signaling pathways in developing cerebral cortex. *Nat. Biotechnol.* **32**, 1053–1058 (2014).
32. Olah, M. et al. Single cell RNA sequencing of human microglia uncovers a subset associated with Alzheimer's disease. *Nat. Commun.* **11**, 6129 (2020).
33. Mathys, H. et al. Single-cell transcriptomic analysis of Alzheimer's disease. *Nature* **570**, 332–337 (2019).
34. Del-Aguila, J. L. et al. A single-nuclei RNA sequencing study of Mendelian and sporadic AD in the human brain. *Alzheimers Res. Ther.* **11**, 71 (2019).
35. Habib, N. et al. Massively parallel single-nucleus RNA-seq with DroNc-seq. *Nat. Methods* **14**, 955–958 (2017).
36. Dewachter, L., Fauvart, M. & Michiels, J. Bacterial heterogeneity and antibiotic survival: understanding and combatting persistence and heteroresistance. *Mol. Cell* **76**, 255–267 (2019).
37. Andersson, D. I., Nicoloff, H. & Hjort, K. Mechanisms and clinical relevance of bacterial heteroresistance. *Nat. Rev. Microbiol.* **17**, 479–496 (2019).
38. Shen, Y. et al. High-throughput single-microbe RNA sequencing reveals adaptive state heterogeneity and host-phage activity associations in human gut microbiome. *Protein Cell* **16**, 211–226 (2025).
39. Wang, Z. et al. Hosts manipulate lifestyle switch and pathogenicity heterogeneity of opportunistic pathogens in the single-cell resolution. *eLife* **13**, RP96789 (2024).
40. Meng, H. et al. High-throughput host-microbe single-cell RNA sequencing reveals ferroptosis-associated heterogeneity during *Acinetobacter baumannii* infection. *Angew. Chem. Int. Ed.* **63**, e202400538 (2024).
41. Bartholomaeus, A. et al. Bacteria differently regulate mRNA abundance to specifically respond to various stresses. *Philos. Trans. A* **374**, 20150069 (2016).
42. O'Donnell, J. L., Kelly, R. P., Lowell, N. C. & Port, J. A. Indexed PCR primers induce template-specific bias in large-scale DNA sequencing studies. *PLoS ONE* **11**, e0148698 (2016).
43. Blattman, S. B. et al. Identification and genetic dissection of convergent persister cell states. *Nature* **636**, 438–446 (2024).
44. Yan, X. et al. An improved bacterial single-cell RNA-seq reveals biofilm heterogeneity. *eLife* **13**, RP97543 (2024).
45. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).
46. Stuart, T. et al. Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902.e1821 (2019).
47. Lopez, R., Regier, J., Cole, M. B., Jordan, M. I. & Yosef, N. Deep generative modeling for single-cell transcriptomics. *Nat. Methods* **15**, 1053–1058 (2018).
48. Pountain, A. W. et al. Transcription-replication interactions reveal bacterial genome regulation. *Nature* **626**, 661–669 (2024).
49. Korsunsky, I. et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
50. Luecken, M. D. et al. Benchmarking atlas-level data integration in single-cell genomics. *Nat. Methods* **19**, 41–50 (2022).
51. Cyriaque, V. et al. Single-cell RNA sequencing reveals plasmid constrains bacterial population heterogeneity and identifies a non-conjugating subpopulation. *Nat. Commun.* **15**, 5853 (2024).
52. Silhavy, T. J., Kahne, D. & Walker, S. The bacterial cell envelope. *Cold Spring Harb. Perspect. Biol.* **2**, a000414 (2010).
53. Ernst, C. M. et al. Adaptive evolution of virulence and persistence in carbapenem-resistant *Klebsiella pneumoniae*. *Nat. Med.* **26**, 705–711 (2020).
54. Claessen, D. & Errington, J. Cell wall deficiency as a coping strategy for stress. *Trends Microbiol.* **27**, 1025–1033 (2019).
55. Kawai, Y., Mickiewicz, K. & Errington, J. Lysozyme counteracts β -lactam antibiotics by promoting the emergence of L-form bacteria. *Cell* **172**, 1038–1049.e1010 (2018).
56. Lukjancenko, O., Wassenaar, T. M. & Ussery, D. W. Comparison of 61 sequenced *Escherichia coli* genomes. *Microb. Ecol.* **60**, 708–720 (2010).
57. Kearns, D. B. & Losick, R. Cell population heterogeneity during growth of *Bacillus subtilis*. *Genes Dev.* **19**, 3083–3094 (2005).
58. Dixit, A. et al. Perturb-Seq: dissecting molecular circuits with scalable single-cell RNA profiling of pooled genetic screens. *Cell* **167**, 1853–1866.e1817 (2016).
59. Westermann, A. J. & Vogel, J. Cross-species RNA-seq for deciphering host-microbe interactions. *Nat. Rev. Genet.* **22**, 361–378 (2021).
60. Westermann, A. J. et al. Dual RNA-seq unveils noncoding RNA functions in host-pathogen interactions. *Nature* **529**, 496–501 (2016).
61. Wang, Y. et al. Fast encapsulation of microbes into dissolvable hydrogel beads enables high-throughput microbial single-cell RNA sequencing of clinical microbiome samples. *Adv. Mater.* **37**, e2500481 (2025).
62. Avraham, R. et al. Pathogen cell-to-cell variability drives heterogeneity in host immune responses. *Cell* **162**, 1309–1321 (2015).
63. Betin, V. et al. Hybridization-based capture of pathogen mRNA enables paired host-pathogen transcriptional analysis. *Sci Rep.* **9**, 19244 (2019).
64. Avital, G. et al. scDual-Seq: mapping the gene regulatory program of *Salmonella* infection by host and pathogen single-cell RNA-sequencing. *Genome Biol.* **18**, 200 (2017).
65. Ntekas, I. et al. Spatial transcriptomics maps host-gut microbiome biogeography at high resolution. *Nat. Microbiol.* <https://doi.org/10.1038/s41564-026-02286-7> (2026).
66. Stoeckius, M. et al. Simultaneous epitope and transcriptome measurement in single cells. *Nat. Methods* **14**, 865–868 (2017).
67. Zuffa, S. et al. microbeMASST: a taxonomically informed mass spectrometry search tool for microbial metabolomics data. *Nat. Microbiol.* **9**, 336–345 (2024).

Acknowledgements

We thank P. Ma and J. Livny for critical reading of this paper. This work was supported by a generous gift from A. and J. Bekenstein (D.T.H.).

Author contributions

A.E.C., F.L., C.S. and D.T.H. contributed to the conception of this work. A.E.C., E.B.-R., F.L., C.S. and D.T.H. contributed to the writing of this paper. A.E.C., C.S. and D.T.H. contributed to editing. A.E.C., V.P.D. and C.S. contributed to figure generation.

Competing interests

D.T.H. has filed a US patent (application number 17/819,034) for BacDrop. The other authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Chris Smillie or Deborah T. Hung.

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

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

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with

the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© Springer Nature Limited 2026