Single-cell RNA-seq reports growth condition-specific global transcriptomes of individual bacteria

Fabian Imdahl[1,*], Ehsan Vafadarnejad[1,*], Christina Homberger[1],

Antoine-Emmanuel Saliba[1,#], Jörg Vogel[1,2,#]

[1]Helmholtz Institute for RNA-based Infection Research (HIRI), Helmholtz-Center for Infection Research (HZI), 97080 Würzburg, Germany

[2]Institute of Molecular Infection Biology, University of Würzburg, Josef-Schneider-Straße 2, D-97080 Würzburg, Germany

**Equal contribution*

[#]Corresponding authors:

emmanuel.saliba@helmholtz-hiri.de, joerg.vogel@uni-wuerzburg.de

Bacteria respond to changes in their environment with specific transcriptional programs but even within genetically identical populations these programs are not homogenously expressed[1]. Such transcriptional heterogeneity between individual bacteria allows genetically clonal communities to develop a complex array of phenotypes[1], examples of which include persisters that resist antibiotic treatment or metabolically specialized cells that emerge under nutrient-limiting conditions[2]. Fluorescent reporter constructs have played a pivotal role in deciphering heterogeneous gene expression within bacterial populations[3] but have been limited to recording the activity of single genes in a few genetically tractable model species whereas the vast majority of bacteria remain difficult to engineer and or even to cultivate. Single-cell transcriptomics is revolutionizing the analysis of phenotypic cell-to-cell variation in eukaryotes but technical hurdles have prevented its robust application to prokaryotes. Here, using an improved poly(A)-independent single-cell RNA-seq protocol, we report the faithful capture of growth-dependent gene expression patterns in individual *Salmonella* and *Pseudomonas* bacteria, across all RNA classes and genomic regions. These transcriptomes provide important reference points for single-cell RNA-seq of other bacterial species, mixed microbial communities and host-pathogen interactions.

Single-cell RNA-seq (scRNA-seq) is becoming routine for eukaryotic cells including small-sized species such yeast[4] and protozoa[5]. This technique has catalyzed the discovery of novel cell types and provided an unprecedented view of tissue anatomy and cellular transitions[6]. Pioneering attempts notwithstanding, however, the application of scRNA-seq to single[7, 8] or few bacteria[9, 10] is yet to prove its ability to infer the physiological state of individual cells within a population[9, 11].

Several important technical barriers exist for bacterial scRNA-seq. Because bacterial cells contain only a femtogram amount of RNA[12], i.e., >100-times less than typical eukaryotic cells, a sensitive cDNA synthesis and amplification protocol is required. In addition, bacterial mRNAs are intrinsically labile (with half-lives in the minute range, as compared to hours in eukaryotes), meaning that perforation of the thick bacterial envelope, cell lysis, and subsequent RNA stabilization must all be done rapidly. Most importantly, the absence of a poly(A) tail on functional bacterial transcripts precludes the straight-forward reverse transcription (RT) strategy that is commonly used in eukaryotes to selectively enrich mRNAs and concomitantly deplete rRNAs. Lastly, whereas most current eukaryotic scRNA-seq protocols have a lower detection limit of 5-10 copies/transcript per cell [13], bacterial scRNA-seq must take into account a much lower average copy-number of mRNAs (0.4 copies/cell; ref. [14]).

To overcome these barriers for bacterial single-cell transcriptomics, here we have applied a highly sensitive random-hexamer priming based scRNA-seq protocol[15] and successfully benchmarked it with the model bacterial pathogen *Salmonella enterica* serovar Typhimurium (henceforth, *Salmonella*), the transcriptome of which is extensively and robustly annotated[16-19]. This has enabled us to obtain high-resolution global transcriptomes of individual *Salmonella* bacteria that faithfully reflect physiologically relevant gene expression profiles established by prior work on bulk populations.

To achieve highly sensitive bacterial scRNA-seq, we developed a generic workflow (**Fig. 1a**), which starts with the isolation of single cells from a culture, followed by a one-tube procedure of cell lysis, cDNA synthesis and amplification. To obtain sufficient cDNA from single bacteria, we used the poly(A)-independent MATQ-

seq (multiple annealing and dC-tailing-based quantitative scRNA-seq)[15] protocol in which RT primers are hybridized to internal transcript regions at low temperature, thus enabling the detection of theoretically all transcripts, including those with low abundance[15]. We analyzed *Salmonella* grown from three different common growth conditions: (i) 'late stationary phase' reflecting mainly resting cells; (ii) 'anaerobic shock' mimicking the intestinal environment; and (iii) 'salt (NaCl) shock' caused by increased sodium chloride concentration in the medium (**Fig. 1b, Methods**). Samples of ten (10-pooled) or single bacteria were systematically isolated using Fluorescence Activated Flow Cytometry (FACS). The accuracy of single-bacteria sorting was >97% as judged by the formation of single colonies (**Extended Data Figure 1**), whereas control sorting of 2, 10 and 100 bacteria resulted in multiple colonies.

In order to minimize transcriptional changes during sorting, the bacteria were treated with *RNAlater*, a stabilization solution that maintains RNA integrity[20, 21]. In addition, cells were sorted directly into PCR plates preloaded with lysis buffer (**Fig. 1a**, **Methods**). The cell wall was digested with 5U of lysozyme targeting the peptidoglycan layer, using an amount of the enzyme that does not inhibit the subsequent RT step; we found that amounts >20U abolished cDNA synthesis. After cell lysis, RT was performed using multiple rounds of annealing (**Fig. 1**). Libraries were amplified with a limited number of 23 PCR cycles (comparable to eukaryotic scRNA-seq[15]), indexed and pooled before each was sequenced to a depth of 62.4 ± 20.9 million reads per 10-pooled and single bacterial cells.

Analyzing 60 of 10-pooled bacteria and 71 single-bacterium samples (**Supplementary Table 1**), we consistently captured all classes of bacterial RNAs (**Fig. 2a**). As expected for priming on total RNA, the cDNA libraries were dominated by rRNA

and tRNA which on average accounted for ~93% of all mapped reads (**Fig. 2a, Supplementary Tables 2 and 3**). Transcripts from mRNA and small non-coding RNA (sRNA) genes represented by ~5% and ~1.2%, respectively, of the total aligned reads, without significant differences between the pooled and single-bacterium samples (**Supplementary Tables 2 and 3**). The relative proportions of the RNA classes are in good agreement with RNA-seq data of *Salmonella* populations without prior ribosomal depletion[20] (**Fig. 2a**). After removal of unmapped reads as well as rRNA and tRNA reads, we proceeded with an average of ~1,500,000 (for 10-pooled) and ~800,000 (single-bacteria) uniquely mapped reads per cell (**Extended Data Figure 2a**); in eukaryotic scRNA-seq such read numbers are usually sufficient to assign physiological states[22].

Removing rRNA and tRNA reads, MATQ-seq detected 413±237 and 170±81 (Av±Std; with five reads) genes in the 10-pooled and single-bacterium libraries, respectively (**Fig. 2b**). Interestingly, we obtained ~2.5-fold more reads for *Salmonella* grown under anaerobic or salt (NaCl) shock, as compared to cells in late stationary phase (**Fig. 2b, Extended Data Figure 2b**). A saturation analysis showed that the number of detected genes marginally increased after 500,000 uniquely aligned reads per bacterium (**Fig. 2c**).

The read distributions along detected genes further supported that we were indeed sequencing cDNA derived from RNA and not genomic DNA: as shown for the *ssrA* (tmRNA), *fliC* (flagellin) and several ribosomal protein genes, the obtained reads covered only their transcribed regions (**Fig. 2d**, **Extended Data Figure 3a-c**). Similar to patterns in *Salmonella* population RNA-seq[16], we observed uneven mapping across the transcript (**Fig. 2d**, **Extended Data Figure 3a-c**). Finally, we plotted for each gene the coefficient of variation of the average read counts and obtained a similar profile as for

conventional single-cell eukaryotic transcriptomes[23] (**Extended Data Figure 4a,b**). 10-pooled and single-bacterium showed a Spearman's correlation ρ=0.5 between Anaerobic and Salt (NaCl) shock and ρ=0.42 for late stationary phase and similar values are obtained when compared with bulk derived RNA-seq (**Extended Data Figure 5a,b**). This correlation analysis shows a large proportion of close to zero values that might stem from dropouts and/or an averaging effect, given the fact that the average bacterial mRNA is present in <1 copy/cell (ref. [14]).

To determine whether the obtained transcriptomes reflected cells in different states, we performed an unbiased clustering of the 10-pooled or single-bacterium libraries using principal-component analysis (PCA). Along the first two components, the clustering robustly delineated cells according to the original growth condition (**Fig. 3a, Extended Data Figure 6**). We also sought to identify differentially expressed genes between conditions using DESeq2 analysis. For the 10-pooled libraries, we retrieved 101 anaerobic shock and 274 salt shock specific genes; the single-bacterium libraries yielded 63 and 131 such genes, respectively (**Fig. 3b, Supplementary Tables 4 and 5**). Among these genes, we found characteristic genes for metabolic shift under anaerobic condition (e.g. *glpA*, *tdcC*) and under salt condition (e.g. *ygdL, yadF, sodA*) (**Fig. 3b**). More globally, a gene ontology analysis of differentially expressed genes revealed typical signatures associated with stationary phase metabolic processes while under anaerobic conditions we observed a global shift to anabolism and catabolism (**Supplementary Table 6**).

We compared our single-cell RNA-seq results with previous bulk RNA-seq results. To this end, we correlated our anaerobic shock and salt (NaCl) shock expression patterns with *Salmonella* RNA-seq profiles provided by the SalCom database[16] (**Fig. 3c,**

Extended Data Figure 5b, Supplementary Tables 4 and 5). Importantly, we observed that ~75% genes upregulated under anaerobic shock and salt shock matched upregulated genes in the SalCom database (**Fig. 3c**), ascertaining the robustness gene expression signatures obtained from single bacteria. However, we emphasize that even if an RNA-seq compendium such as SalCom is available, we recommend a bulk RNA-seq sample be prepared in parallel for benchmarking of the scRNA-seq data.

Finally, to establish proof-of-principle for generalizability of our bacterial scRNA-seq approach, we applied MATQ-seq to another species, the clinically important pathogen *Pseudomonas aeruginosa* whose genome has a much higher GC content than that of *Salmonella* (>67% versus ~50%). Generating libraries from 10-pooled and single *P. aeruginosa* bacteria (**Extended Data Figure 7, Supplementary Tables 7 and 8**), MATQ-seq captured on average 147 genes at the single-bacterium level, a number that compares favourably with the *Salmonella* scRNA-seq above.

In conclusion, by adopting MATQ-seq to capture low abundant transcripts, we have been able to perform scRNA-seq on minimal microbial samples, down to single bacteria, and subsequently infer different growth conditions. Some previous reports of prokaryotic scRNA-seq detected thousands of genes expressed in a single bacterium, which is hard to reconcile with the fact that most bacterial mRNAs are predicted to be present in <1 copy per cell, on average[7, 8]; or the scRNA-seq libraries had to be pooled back to infer expression programmes[9]. Here, by choosing well-characterized growth conditions and independent datasets[16], we could for the first time assign robust transcriptomic signatures to single bacteria. While the number of expressed genes detected here by MATQ-seq is in the hundreds, and so allows transcription of individual bacterial cells to be studied globally, this number is still lower than what is required to

reflect a full bacterial transcriptome[24]. Nonetheless, our mRNA capture rates seem realistic in light of the fact that a recently published PETRI-seq protocol achieved capture in the range of 200 mRNAs per cell for exponentially growing *Escherichia coli* [25].

Our workflow should pave the way for gene activity profiling in complex ecological microbial niches such as the microbiome on the level of its many different single bacteria and to monitor the drug susceptibility of pathogens based on RNA signatures from clinical samples[26]. At present, bacterial single-cell transcriptomics by MATQ-seq comes with considerable cost (**Supplementary Table 9**) but we are positive that making the protocol compatible with microfluidics devices and implementing rRNA read removal prior to sequencing [27, 28] will permit to dramatically scale up throughput while also reducing the cost per analyzed cell.

Figure legends

Fig. 1: Overview of bacterial single-cell RNA-seq workflow.

a, Single bacteria are sorted from a culture into a 96-wells plate using a Fluorescent Activated Cell Sorter (FACS). The individual bacteria are lysed enzymatically and multiple annealing and dC-tailing-based quantitative single-cell RNA-seq (MATQ-seq) is applied to reverse-transcribe the bacterial RNA and to tail first-strand cDNAs with poly(C). After PCR amplification, cDNA is processed with tagmentation via Tn5 transposase to generate sequencing-ready libraries. **b**, Growth conditions used for scRNA-seq of *Salmonella*.

Fig. 2: Characterization of transcriptomes down to the single bacterium level across different growth conditions. **a**, Representative proportions of transcript categories after removal of unmapped reads obtained for 10-pooled and single bacteria in this study and from *Salmonella* bulk RNA-seq generated with random hexamers (data from Ref [20]). IGR, intergenic region; ncRNA: non-coding RNA; sRNA: small RNA; other: all other RNA classes (**Supplementary Tables 2 and 4;** Single bacterium : Late Stationary Phase n=19, NaCl Shock n=28, Anaerobic Shock n=19; 10-pooled bacteria: n=20 in each condition). **b**, Violin plot displaying the number of detected genes across all libraries (Single bacterium : Late Stationary Phase n=19, NaCl Shock n=28, Anaerobic Shock n=19; 10-pooled bacteria: n=20 in each condition). **c**, Number of genes detected per cell (counts > 5) when downsampling total read counts to the indicated depths. Dashed line for single bacteria (late stationary phase condition) data represents an extrapolated asymptotic fit. **d**, Representative raw reads found to align against the

reference sequence of tmRNA-encoding *ssrA* gene (±300 bp upstream and downstream the CDS) across 10-pooled bacteria and single-bacteria.

Fig. 3: Single-bacteria RNA-seq reveals specific transcriptional signatures associated with growth conditions.

a, Principal component analysis (PCA) allowed to delineate the three studied growth conditions. Each dot represents 10-pooled-bacteria (upper part) or a single bacterium (lower part). Colors indicate the growth conditions. **b**, Heatmaps displaying the transcriptional profiles specific for each growth condition. A total of 431 and 209 genes were differentially expressed in 10-pooled (upper heatmap) and single bacteria (lower heatmap), as identified by DEseq2 (**Supplementary Tables 4 and 5**). **c**, For each gene found to be differentially expressed in our study between Salt (NaCl) and Anaerobic (Ana) shocks (heatmap, panel **b**), the $\log_{10}$ transformed ratio of the expression values independently measured under both shock conditions in the bulk RNA-seq benchmark study (in TPM (Transcripts per Million), Kröger et al[16]) has been calculated and represented in histograms (left; color code represents whether the log transformed ratio is >1 or <1) and boxplots (right; median, first and third quartile, 5–95 percentile and outliers are shown).

Reporting Summary

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

Correspondence and requests for materials

Correspondence and requests for materials should be addressed to Antoine-Emmanuel Saliba and Jörg Vogel.

Acknowledgements

The authors thank Stan Gorski for critical comments on the manuscript, Kerim Secener and Xavier Hartmann for performing preliminary experiments, Charlotte Michaux for advices on bacterial culture, Lars Barquist and Florian Erhard for advices on statistical analysis and Panagiota Arampatzi, Tobias Heckel (SysMed Würzburg) and Robert Geffers (Genome Analytics, HZI) for conducting the sequencing.

Competing interests

The authors declare no competing interests.

Contributions

F.I. and C.H. conducted experiments. E.V. performed data analysis. A.-E.S. designed the research. J.V and A.-E. S. directed research. A.-E.S. and J.V. wrote the manuscript.

Material and Methods

Bacterial growth conditions.

Wild-type (WT) *Salmonella* Typhimurium strain SL1344 were cultured in 5 ml Lennox broth (LB) at 37°C under constant agitation at 220 rpm (New Brunswick Innova 44) to reach the optical density (OD, λ=600 nm) of 2.0 followed by a further 6hr growth. This growth condition is defined as "late stationary phase (LSP)" in ref. [16]. A 1:100 dilution of the latter bacterial culture was grown at 37°C under 220 rpm agitation to reach an OD (λ=600 nm) of 0.3 and subjected to a "Salt shock" (or "Nacl Shock") or an "Anaerobic shock" (**Fig. 1 and Extended Data 1a**) as described previously [16] (**Extended Data 1a**). 2 mL of the culture were pelleted, washed twice with 1 mL of Dulbecco's phosphate-buffered saline (DPBS, Gibco) and mild fixed by resuspending the pellet with 1 mL of a 1:1 v/v RNAlater (Ambion, Thermo Fischer) – DPBS solution to prevent advert transcriptome changes following the vendor instructions. *Pseudomonas aeruginosa* strain PAO1 was cultured overnight (O/N) in 5 ml LB at 37°C under constant agitation at 220 rpm (New Brunswick Innova 44). A 1:100 dilution of the later-culture was then grown to an OD of 0.3. Anaerobic shock condition was prepared as previously described [16]. Further processing corresponds to the *Salmonella* treatment described above.

Single bacteria isolation.

Salmonella were sorted using a BD FACS Aria III (70 µm nozzle; single-cell precision) into 48-well plates (Brand) pre-filled with a lysis buffer (0.26 µL 10x Lysis buffer (Takara), 0.03 µL RNase Inhibitor (100U/uL, Takara), 0.26 µL DPBS (Gibco), 0.1 µL Lysozyme (50 Units/µL, Epicentre) and 1.95 µL nuclease free water (Ambion) to reach 2.6 µL). For *Pseudomonas aeruginosa* 0.026 µl EDTA (50mM, Thermo) was added to the

previously described lysis composition. Afterwards, the sorted cells were sonicated for 10 seconds (Sonorex Digitec DT 52, Bendelin). Plates were kept on ice after sorting and stored at - 80°C for long term storage covered by an aluminium-based microfilm (Biorad seal F).

10-pooled and single bacteria RNA-seq using MATQ-seq.

Cycling and incubation steps were performed in a Bio-Rad T-100 thermocycler and primer (IDT) list can be found in **Supplementary Table 10** as described in Sheng et al. [15]. After sorting, every well was supplemented with a primer mix containing 0.05 µL DTT (Life Technologies), 0.4 µL primer mix 100 µM containing GAT27dT and two random hexamers GAT27 5N3G and GAT27 5N3T and 0.12 µL 10 mM dNTP. Sealed plates (Microseal B, Biorad) were incubated 3 min at 72°C and transferred on ice for at least 1 min. Then, MATQ-seq protocol was performed following the initial report[15] .

RNA was reverse transcribed using Superscript III (0.15 µl, 200 U/µl; Life Technologies) with a temperature ramping between 8°C and 50°C and repeated for 10 cycles. Residual primers were digested with T4 DNA Polymerase (NEB). RNase H (0.2 µl, 5,000 U/ml) and RNase If (0.2 µl, 50,000 U/µl) (NEB) were used for enzymatic RNA digestion and TdT Terminal Transferase (NEB) and dCTP were provided for the tailing step. Subsequently second strand was synthesized using GAT21 6N3G MALBAC primer and deepvent exo- DNA polymerase (0.4 µl, 2,000 U/ml; NEB). Amplification was performed with a PCR mix containing GAT27 PCR primer and deepvent exo- DNA polymerase (3 µl, 2,000 U/ml, NEB) besides Thermopol-Buffer (NEB) and dNTP with a 24 cycle PCR program.

Amplified cDNA was purified using 40 µL of Ampure XP beads (Beckman Coulter) with a 1:1 v/v beads: amplicons. cDNA concentration was measured using Qubit HS Assay (Life

Technologies) and fragment size distribution was tracked using a high sensitivity DNA kit Bioanalyzer (Agilent).

Library preparation and sequencing.

Libraries were prepared using Nextera XT (Illumina) with the minor modifications as described previously[29]. Briefly, an input of 1 ng of cDNA was tagmented using one-quarter of the recommended volumes and a longer fragmentation time (10 min), and resuspended in 15 µL Elution Buffer (EB; Qiagen). Pooled libraries were sequenced using an Illumina Nextseq 500 (2*75 bp) mid-output (1x) and high output (1x) for *Pseudomonas* and with an Illumina Novaseq6000 (2xS1 and 1xS2 full cartridges 2*50bp PE) for *Salmonella*. Libraries identities and description can be found in **Supplementary Table 1**.

Data analysis and statistics.

After demultiplexing, quality of data was examined using FastQC (v 0.11.7). Illumina and MATQ-seq adapters were removed using cutadapt (v 1.9). Trimmed reads were mapped to *Salmonella* enterica SL1344 (NCBI ASM21085v2) and *Pseudomonas aeruginosa* strain PAO1 genome using STAR aligner (v 2.5.4b) with default settings. Read counts of each gene was determined using the featureCounts program. Genes with more than 5 aligned reads were considered as detected genes. Afterwards, data was subjected to a PCA; the top 300 genes with the highest variance were selected to perform dimension reduction. Five libraries named M2G3, M3C4, M2H7, M2B7 and M2C10, which were clear outliers on the PCA plot were excluded from further analysis (**Supplementary Table 1**). The differentially expressed genes between different culture conditions were analyzed using DESeq2 package[30] and genes with adjusted *p*-value below 0.01 were considered

statistically significant. Gene expression values from the reference database were extracted from SalCom (http://bioinf.gen.tcd.ie/cgi-bin/salcom.pl?_HL). To quantify the relationship between the number of detected genes and sequencing depth, we downsampled reads to varying depths, and we estimated to what extent libraries were sequenced to saturation.

Data availability

All RNA-seq data has been deposited in NCBI's Gene Expression Omnibus under accession number GSE119888. Source data for Figs. 2-3 and Extended Data Figs. 2, and 4-7 are provided with the paper.

Code Availability

All codes to perform the analysis and reproduce the figures are available on GitHub:

https://github.com/saliba-lab/single-bacterium-transcriptome-profiling

References

1. Ackermann, M. A functional perspective on phenotypic heterogeneity in microorganisms. *Nat Rev Microbiol* **13**, 497-508 (2015).
2. Gollan, B., Grabe, G., Michaux, C. & Helaine, S. Bacterial Persisters and Infection: Past, Present, and Progressing. *Annu Rev Microbiol* **73**, 359-385 (2019).
3. Kreibich, S. & Hardt, W.D. Experimental approaches to phenotypic diversity in infection. *Curr Opin Microbiol* **27**, 25-36 (2015).
4. Gasch, A.P. et al. Single-cell RNA sequencing reveals intrinsic and extrinsic regulatory heterogeneity in yeast responding to stress. *PLoS Biol* **15**, e2004050 (2017).
5. Muller, L.S.M. et al. Genome organization and DNA accessibility control antigenic variation in trypanosomes. *Nature* **563**, 121-125 (2018).
6. Stubbington, M.J.T., Rozenblatt-Rosen, O., Regev, A. & Teichmann, S.A. Single-cell transcriptomics to explore the immune system in health and disease. *Science* **358**, 58-63 (2017).
7. Kang, Y. et al. Transcript amplification from single bacterium for transcriptome analysis. *Genome Res* **21**, 925-935 (2011).
8. Wang, J., Chen, L., Chen, Z. & Zhang, W. RNA-seq based transcriptomic analysis of single bacterial cells. *Integr Biol (Camb)* **7**, 1466-1476 (2015).
9. 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).
10. Betin, V. et al. Hybridization-based capture of pathogen mRNA enables paired host-pathogen transcriptional analysis. *Sci Rep* **9**, 19244 (2019).
11. Penaranda, C. & Hung, D.T. Single-Cell RNA Sequencing to Understand Host-Pathogen Interactions. *ACS Infect Dis* (2019).
12. Milo, R. & Phillips, R. Cell biology by the numbers. (Garland Science, 2015).
13. Bagnoli, J.W. et al. Sensitive and powerful single-cell RNA sequencing using mcSCRB-seq. *Nat Commun* **9**, 2937 (2018).
14. Taniguchi, Y. et al. Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells. *Science* **329**, 533-538 (2010).
15. Sheng, K., Cao, W., Niu, Y., Deng, Q. & Zong, C. Effective detection of variation in single-cell transcriptomes using MATQ-seq. *Nat Methods* **14**, 267-270 (2017).
16. Kroger, C. et al. An infection-relevant transcriptomic compendium for Salmonella enterica Serovar Typhimurium. *Cell Host Microbe* **14**, 683-695 (2013).
17. Smirnov, A. et al. Grad-seq guides the discovery of ProQ as a major small RNA-binding protein. *Proc Natl Acad Sci U S A* (2016).
18. Chao, Y. et al. In Vivo Cleavage Map Illuminates the Central Role of RNase E in Coding and Non-coding RNA Pathways. *Mol Cell* **65**, 39-51 (2017).
19. Hor, J., Matera, G., Vogel, J., Gottesman, S. & Storz, G. Trans-Acting Small RNAs and Their Effects on Gene Expression in Escherichia coli and Salmonella enterica. *EcoSal Plus* **9** (2020).
20. Westermann, A.J. et al. Dual RNA-seq unveils noncoding RNA functions in host-pathogen interactions. *Nature* **529**, 496-501 (2016).
21. Westermann, A.J. & Vogel, J. Host-Pathogen Transcriptomics by Dual RNA-Seq. *Methods Mol Biol* **1737**, 59-75 (2018).
22. Picelli, S. et al. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat Methods* **10**, 1096-1098 (2013).
23. Brennecke, P. et al. Accounting for technical noise in single-cell RNA-seq experiments. *Nat Methods* **10**, 1093-1095 (2013).
24. Haas, B.J., Chin, M., Nusbaum, C., Birren, B.W. & Livny, J. How deep is deep enough for RNA-Seq profiling of bacterial transcriptomes? *BMC Genomics* **13**, 734 (2012).
25. Blattman, S.B., Jiang, W., Oikonomou, P. & Tavazoie, S. Prokaryotic single-cell RNA sequencing by in situ combinatorial indexing. *Nat Microbiol* (2020).

26. Barczak, A.K. et al. RNA signatures allow rapid identification of pathogens and antibiotic susceptibilities. *Proc Natl Acad Sci U S A* **109**, 6217-6222 (2012).
27. 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).
28. Prezza, G. et al. Improved bacterial RNA-seq by Cas9-based depletion of ribosomal RNA reads. *RNA* (2020).
29. Patel, A.P. et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* **344**, 1396-1401 (2014).
30. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).