Cross-species RNA-seq deciphers host-microbe interactions

Alexander J. Westermann^{1,2,*} & Jörg Vogel^{1,2,*}

5 ¹Helmholtz Institute for RNA-based Infection Research (HIRI), Helmholtz Centre for Infection Research (HZI), D-97080 Würzburg, Germany

²Institute of Molecular Infection Biology (IMIB), University of Würzburg, D-97080 Würzburg, Germany

*correspondence: alexander.westermann@uni-wuerzburg.de; joerg.vogel@uni-wuerzburg.de

Abstract:

The human body is constantly exposed to microbes, which entails manifold interactions between human cells and diverse commensal or pathogenic bacteria. The cellular states of the interacting cells are decisive for the outcome of these encounters, such as whether bacterial virulence programs and host defense or tolerance mechanisms are induced. This review summarizes how next-generation RNA sequencing (RNA-seq) has become a primary technology to study host-microbe interactions with high-resolution, improving our understanding of the physiological consequences and the mechanisms at play. We illustrate how the discriminatory power and sensitivity of RNA-seq helps to dissect increasingly complex cellular interactions in time and space, down to the single-cell level. We also outline how future **transcriptomics** may answer currently open questions in host-microbe interactions and inform treatment schemes for microbial disorders.

Introduction

Interactions between **host** and bacterial cells form an integral part of human physiology. Colonizing bacteria may attach to mucosal surfaces or bind specifically to host receptors at epithelial linings, and have to adapt their growth and metabolism to the respective
5 microenvironment. Conversely, to respond adequately to colonizing microbes, our immune system has to continuously distinguish beneficial bacteria from harmful pathogens. In addition, it is increasingly appreciated that the potentially harmful encounter with a bacterial pathogen is not only determined by the presence or absence of virulence traits in that bacterium, but also by when, where, and how it interacts with host cells. What is more, many infectious diseases are
10 polymicrobial in nature and disorders in the composition of the myriads of bacteria, fungi, and viruses that constitute the host-associated microbiota substantially influence susceptibility or resistance to infection.

Complexity extends to the host whose tissues and organs comprise multiple different cell types that **engage in extensive cross talk**. On top of that, spatial aspects determine the outcome
15 of host-microbe encounters. Pathogens occupy specific niches within their host body wherein they persist or wherefrom they disseminate, creating vastly different microenvironments. However, even within a defined host niche, phenotypic heterogeneity between genetically identical cell populations can lead to divergent results with sometimes severe effects for the host. The resultant variability in host-microbe interaction might contribute to therapeutic
20 failures and the establishment of chronic, recurring infections ¹. All in all, to fully understand the major principles of host-microbe interactions in health and disease, high-resolution approaches are needed that can capture their full complexity at different scales, from the level of the whole organism down to its individual single cells.

Experimental infection systems – both, animal models ² and advanced three-dimensional
25 *in vitro* tissues ³ – are getting better at recapitulating the *in vivo* situation within human niches. In parallel, the technologies to measure the biological activities of the involved organisms keep improving. For instance, today's infection biologist can draw on different global approaches to study host-microbe encounters on a systems level (**Box 1**). Among these approaches, transcriptomic methods record **steady-state levels of transcripts and can thereby provide** a
30 snapshot of the cellular physiology.

Thanks to their sensitivity, cost-efficiency, and generic nature, transcriptomics by next-generation RNA sequencing (RNA-seq) has been popular in studying host-microbe interactions ⁴. Since the introduction of RNA-seq to the field of infection biology about a decade ago, there have been three major phases with respect to the analysis of host-microbe interactions (**Fig. 1**). In the

first phase, host and bacterial cells were physically separated from one another and their transcriptomes analyzed individually. The second phase begun with the realization that its high discriminatory power and intrinsic single-nucleotide resolution render RNA-seq an exceptionally powerful technique for the simultaneous detection and quantification of transcripts from different interacting organisms ⁵. This was accompanied by a rethink that noncoding transcripts – long regarded as mere byproducts of RNA-seq profiling – reveal crucial insights into host-microbe interplay ⁶⁻⁸. The present third phase is marked by the increasing use of single-cell RNA-seq (scRNA-seq) to dissect cellular heterogeneity in host-microbe encounters ^{9,10}.

Here, we review cross-species transcriptomics, focusing on the interaction of bacterial pathogens or commensals with mammalian (mostly human) host cells. We first summarize similarities and differences between bacterial and mammalian transcriptomes and transcriptomics, followed by an overview of current cross-species approaches. We will bring the technologies to life with examples of what biological insights they have already provided, including the realization that host-microbe interactions are both context-dependent and highly heterogeneous. We will also discuss current limitations of cross-species transcriptomics and how to overcome them, especially against the backdrop of the current scRNA-seq revolution.

Eukaryotic and bacterial transcriptome features

An organism's transcriptome may provide a snapshot of its physiological state. On the cellular level, this transcriptome is an RNA potpourri of myriads of different coding and noncoding transcripts from several major classes, some of which are being made, others are already present in the cell in their active form, and yet others are undergoing decay. Despite two billions of years of separate evolution, the fundamental RNA classes are shared between prokaryotes and eukaryotes (Fig. 2). Ribosomal RNAs (rRNAs), serving scaffolding and enzymatic functions in the ribosome, are highly abundant, accounting for >80% (and >95% in fast growing cells) of the RNA content of bacterial and eukaryotic cells alike. Transfer RNA (tRNA) molecules translate RNA language (ribonucleobases) into protein information (amino acid chains) and generally contributes another ~10% to the cellular transcriptome.

Messenger RNA (mRNA) instructs the ribosome which protein to make. Making up ~5% of the total cellular transcriptome in both bacteria and eukaryotes, mRNAs are characterized by kingdom-specific features. Bacterial genes are on average 1 kb long; however, they are often transcribed as polycistrons, resulting in long mRNAs that span multiple genes. Primary transcripts typically carry a 5' triphosphate group, while 3' ends of bacterial mRNAs lack an

extended poly(A) stretch (3' polyadenylation does occur, but is limited to a few nucleotides that enable rapid transcript decay¹¹). By contrast, eukaryotic mRNAs carry a 7-methylguanosine cap and a poly(A) tail in the range of 70-250 nucleotides (nt). Unlike bacterial mRNAs, which are made in a translation-ready form, eukaryotic coding transcripts are synthesized as precursor-mRNAs with introns and exons, and undergo splicing (intron removal) in the nucleus prior to export to the cytoplasm. The average length of mature human mRNAs is ~3 kb; since only monocistrons are present in mammals, extremely long coding transcripts (>10 kb) are rare.

Certain additional, specialized RNA molecules exist in both kingdoms, such as the Y-RNAs, which are stable noncoding transcripts that associate with the Ro60 autoantigen in mammals¹² and Ro60-related proteins in bacteria¹³, or the RNA component of RNase P (M1 RNA)¹⁴. Most noncoding RNA species, however, are exclusively found in either bacteria or eukaryotes, where they carry out regulatory, enzymatic or scaffolding functions. In their sum, these heterogeneous transcripts typically contribute another ~5-10% to the bacterial or eukaryotic transcriptome, although the relative proportion of bacterial noncoding RNAs seem to strongly increase under certain conditions¹⁵. Most bacteria express dozens to hundreds of different small regulatory RNAs (sRNAs)¹⁶. An sRNA may be transcribed from an intergenic noncoding gene or cleaved off the end of an mRNA. Most sRNAs regulate target genes post-transcriptionally by short base-pair interactions with the respective mRNAs.

Noncoding RNAs specific to eukaryotic transcriptomes include small nuclear and small nucleolar RNAs (snRNAs and snoRNAs) that are involved in the splicing of precursor mRNAs or the chemical modification of rRNAs and tRNAs, respectively, and localize to nuclear compartments. PIWI-interacting RNAs (piRNAs) are best known for silencing transposable elements in the germline, while noncoding RNAs in somatic cells are generally divided into microRNAs (miRNAs) and long noncoding RNAs (lncRNAs). The homogeneous class of miRNAs, with their mature form being ~22 nt long, localize to the cytosol, where they regulate target mRNAs in a sequence-dependent manner reminiscent of bacterial sRNAs. By contrast, lncRNAs are more heterogeneous and are operationally defined as noncoding transcripts >200 nt in length (note however, that there is an ongoing debate as to how many of them encode short peptides¹⁷). Some lncRNAs carry a 3' poly(A) tail while others are not polyadenylated. LncRNAs may be retained in the nucleus where they help to regulate chromatin state, or localize to the cytoplasm and fulfill diverse functions, e.g., maintaining subcellular compartments, serving as scaffolds for ribonucleoparticle assembly, or sequestering miRNAs, to name but a few¹⁸.

The differences between transcript repertoires are accompanied by a quantitative disparity between bacterial and eukaryotic cellular RNA contents (Fig. 2). The *Escherichia coli* genome has a size of ~5 Mbp, whereas the entire human genome is ~3,000 Mbp. The RNA

content **also** differs dramatically: an average *E. coli* cell contains ~100 fg of RNA, whereas an average human cell contains up to 30 pg RNA, i.e., several hundred times more ¹⁹. Overall, these **qualitative and quantitative** transcriptome differences get reflected in the current standard RNA-seq protocols for bacterial and mammalian samples.

5

Bacterial versus eukaryotic RNA-seq work flows

The standard steps in an RNA-seq experiment are purification of nucleic acids from a biological sample, enzymatic digestion of contaminating genomic DNA, depletion of abundant, but less informative rRNA, conversion of the remaining transcripts into complementary DNA (cDNA),
10 high-throughput sequencing of the cDNA fragments, alignment of the resulting sequencing reads to a reference genome sequence, and quantification of the reads mapped to individual genetic features. RNA-seq provides various readouts, **as follows**. Counting all sequencing reads from the same transcript reveals the relative expression of the corresponding gene. Additionally, read coverage distributions **reveal** general transcript features, such as 5' and 3' boundaries,
15 processing sites, and operon (**for** bacteria) or intron-exon (**for** eukaryotes) structures. Researchers can choose between different library preparation pipelines and RNA-seq platforms that all come with specific strengths and weaknesses (see **Box 2**). For more details of the sequencing technologies available, we refer the reader to recent specialized reviews ^{20,21}. Suffice to say here that the current standard for both bacterial and eukaryotic expression profiling is
20 'short-read' sequencing-by-synthesis using Illumina technology. Here, we restrict ourselves to discussing commonalities and differences in the most commonly used bacterial and eukaryotic Illumina-based RNA-seq protocols (**Fig. 3a**), as these are closest to being developed into multi-species RNA-seq approaches.

25 **Transcriptome fixation, cellular lysis, and RNA extraction.** Sample processing may be time-consuming and can involve multiple handling steps; especially so in scenarios where specific sub-populations of bacterial or host cells are collected. Given that transcriptomes are notoriously unstable, in such cases the RNA composition should be preserved with the help of fixatives that block *de novo* transcription and RNA decay. A variety of RNA-preserving reagents
30 exist – each with specific strengths and weaknesses ²² – and are increasingly used for robust microbe-microbe and host-microbe transcriptomics ²³⁻²⁷. A systematic evaluation of several standard fixation methods showed unexpected substantial fragmentation of RNA after isolation from infection samples, with consequences for the coverage of different RNA classes in subsequent RNA-seq ²⁸. **As a general guideline, since both formaldehyde- and alcohol-based**

fixation can have adverse effects on RNA integrity ²⁴, ammonium sulfate-containing stabilizing reagents, such as the commercial *RNAlater* (Sigma-Aldrich) or *RNAprotect* (QIAGEN), are often preferred for whole-transcriptome studies. In experimental setups that rely on sort-enrichment, however, careful evaluation of these RNA-preserving reagents is recommended since they might
5 quench the fluorescent signals of marker proteins.

The physicochemical properties of prokaryotic and eukaryotic cells differ vastly, as they do among bacterial species themselves (e.g. Gram-negatives versus Gram-positives). For instance, while Gram-negative cells can easily be cracked, also allowing joint lysis with mammalian host cells, efficient disruption of Gram-positive cell walls often depends on
10 enzymatic and/or mechanical treatment. It is thus advisable to carefully evaluate lysis efficiencies for cross-species transcriptomics. That is, lysis conditions need to be empirically established that are sufficiently harsh to effectively and homogeneously break up all cell types in the sample, while still being mild enough to not degrade cellular RNAs.

Once the lysates are prepared, downstream RNA isolation techniques are typically
15 interchangeable between organisms, but may bias toward individual RNA classes dependent on their size, secondary structure, and degree of modification. Further information on these topics can be obtained from Ref²⁹.

Ribodepletion. Ribosomes are the most abundant ribonucleoprotein particles in any living cell, and although their number fluctuates with growth rate ^{30,31}, the rRNA they contain will dominate
20 the cellular RNA pool. While rRNA reads in RNA-seq data are occasionally used to infer bacterial replication rates ^{32,33}, this transcript class is commonly depleted in both bacterial and eukaryotic RNA-seq to indirectly increase coverage of mRNAs and other informative RNA classes. However, ribodepletion strategies partly differ between bacterial and eukaryotic studies.

In bacteria, rRNA is typically removed actively before or during cDNA library generation. Most commonly, it is pulled out from the pool of total RNA by sequence-specific ‘capture oligonucleotides’ that can be biotinylated (e.g. riboPOOL [siTOOLS]) or coupled to magnetic
25 beads (e.g. RiboCop [Lexogen]), available as commercial or “do-it-yourself” kits ³⁴, reaching depletion efficiencies of >95% ^{35,36}. Alternatively, upon incubation with sequence-specific DNA
30 oligonucleotides, the resulting rRNA-DNA hybrids are digested with RNase H (a principle incorporated in some NEBNext protocols and in Illumina’s new Ribo-Zero Plus technology). Recently, we have introduced bacterial rRNA depletion on the cDNA level with programmed Cas9 nuclease cleavage ³⁷. This technique, which is generally known as DASH (depletion of
abundant sequences by hybridization; ³⁸), removes rRNA less efficiently but seems particularly

suitable for low-input samples ³⁷. What remains after bacterial ribodepletion are mRNAs, sRNAs, and tRNAs. The latter require specialized cDNA protocols for efficient conversion ³⁹⁻⁴¹, so it is primarily mRNAs and sRNAs that will eventually make it into conventional sequencing libraries.

In eukaryotic RNA-seq, rRNA is generally depleted indirectly by selective cDNA priming on polyadenylated transcripts, i.e., primarily mRNAs and lncRNAs (poly(A)-selection is an integral part of the TruSeq Stranded mRNA kit from Illumina). Alternatively, if non-polyadenylated transcripts (e.g. poly(A)-negative lncRNAs) are of interest as well, active rRNA removal strategies such as the commercial riboPOOL, RiboCop or Ribo-Zero Plus technologies or customized DASH ³⁸ may be used. These latter methods will also be the top choice when analyzing mixtures of bacteria and eukaryotic host cells by so-called dual RNA-seq ²⁸, which will be covered further below.

cDNA library preparation and sequencing. The dominating Illumina technology offers a maximum read length of 300 bp, necessitating fragmentation of total RNA samples prior to cDNA conversion for bacteria and eukaryotes alike. In bacterial RNA-seq pipelines, RNA adapters are then often ligated to the 3' end of the obtained RNA fragments and subsequently used as anchors for reverse transcription (RT) via adapter-specific primers (as in the popular NEBNext Small RNA Library Prep Set; New England Biolabs). Similar adapter-based library generation protocols can also be used in combination with size-selected, unfragmented eukaryotic short RNAs, e.g., for miRNA profiling. By contrast, profiling of eukaryotic mRNA expression generally involves enrichment of polyadenylated transcripts followed by RNA fragmentation and RT by random priming (e.g. TruSeq Stranded mRNA kit; Illumina). Dependent on the sample type and the scope of the RNA-seq experiment, alternative, custom-tailored library preparation protocols exist, some of which are described in **Box 2**.

Sequencing depth requirements, too, depend on the desired outcome of an RNA-seq experiment. **Although** accurate mapping of transcript boundaries, operon or alternative splicing structures, and quantification of low-abundance transcripts demand high sequencing depth, most RNA-seq studies in bacteria and eukaryotes are, in essence, differential expression analyses and as such, much less sequencing-intensive. Simulation studies proposed ballpark figures of 5-10 million non-rRNA reads for exhaustive expression profiling in bacteria ⁴² and ~20 million non-rRNA reads in mammals ^{43,44}.

Read alignment, normalization, quantification, and functional analyses. The concepts of read mapping and quantification are shared between bacterial and eukaryotic RNA-seq, with differences in the details ²⁹. Since splicing is a molecular process exclusive to eukaryotes, standard read mappers for prokaryotic RNA-seq generally save both computing power and time by blanking it out, whereas eukaryotic mapping pipelines strictly depend on spliced aligners ⁴⁵. Normalization is similar for bacterial and eukaryotic RNA-seq data, with the most common methods assuming that the expression of a substantial fraction of genes is not altered between two distinct conditions ⁴⁶. However, this assumption fails for pairwise comparison of extreme conditions, particularly likely in bacteria that can change a considerable part of their transcriptome, which has fostered the development of normalization tools for bacteria ⁴⁷. On the experimental side, the inclusion of artificial RNA spike-ins in the RNA-seq protocol provides an opportunity to normalize to cell numbers ⁴⁸⁻⁵⁰.

For differential expression analysis, the DESeq2 ⁵¹, edgeR ⁵², and limma/voom ⁵³ algorithms have been most popular. Since these tools cannot analyze differential isoform usage, RNA-seq studies addressing alternative splicing in eukaryotes rely on different algorithms ⁵⁴. Once differentially expressed genes (or isoforms) have been identified, these changes are usually subjected to global pathway analysis. Both bacterial and eukaryotic transcriptomics rely on well-curated databases such as Gene Ontology (GO) ⁵⁵ or Kyoto Encyclopedia of Genes and Genomes (KEGG) ⁵⁶. Additionally, for the analysis of mouse or human RNA-seq data, more refined databases exist that specialize in, for example, metabolic ⁵⁷ or immune-related gene sets ⁵⁸.

In summary, despite the basic steps being shared between conventional prokaryotic and eukaryotic RNA-seq protocols, experimental and analytical pipelines diverged over the past decade to satisfy the specific needs of the respective target transcriptome. As a consequence, RNA-seq profiling of bacterial gene expression *in vivo* was long one-sided and entailed the physical enrichment of bacteria from host tissue, as will be reviewed in the next section.

Bacterial RNA-seq for *in vivo* transcriptomics

In order to fully exploit global expression data of host and microbe, high-resolution transcriptome maps are required. Whereas the mouse and human transcriptomes have been well annotated over the years, fine-grained transcriptome maps are not a given for most bacteria. Rather, bacterial transcriptomes are often annotated based on computational searches for potential open reading frames (ORFs) with a length of >100 amino acids, but short ORFs, untranslated regions (UTRs) or noncoding RNA genes may be missing. Consequently, a truly comprehensive global expression analysis usually requires prior transcriptome refinement

using RNA-seq-based techniques such as differential RNA-seq (dRNA-seq; initially applied to Gram-negative species, such as *Helicobacter pylori*⁵⁹ and *Salmonella enterica*⁶⁰, and later to Gram-positives, e.g. *Streptococcus pneumoniae*⁶¹) to globally map transcriptional start sites, or terminator-sequencing (Term-seq; established in *Listeria monocytogenes*⁶²) or end-enriched RNA-seq (Rend-seq; established in *E. coli* and *Bacillus subtilis*)⁶³ to determine transcription termination sites. Such reference transcriptomes are now available for numerous aerobic^{60,64-69} and anaerobic bacteria⁷⁰. In addition, *Salmonella* research has greatly benefitted from a gene expression atlas called SalCom, compiling RNA-seq profiles under 22 defined *in vitro* conditions, each mimicking specific phases of the *Salmonella* infection cycle¹⁵. Similarly, the recently established PATHOgenex database⁷¹, which features host stress-related *in vitro* transcriptome signatures of 32 human pathogens, provides an excellent resource for infection biologists.

The emergence of the field of cellular microbiology three decades ago⁷² brought about a need to study gene expression in bacterial pathogens within the context of their host organism. The first such approaches relied on physical enrichment of bacteria from infection samples, followed by transcriptome analysis of the purified bacteria alone. For example, infected eukaryotic cells were selectively lysed with a detergent and the released intracellular bacteria were collected by differential centrifugation^{73,74}. Strikingly, the gene expression profile of *S. enterica* recovered from macrophages this way revealed several activated genes that were silent in any of the 22 *in vitro* conditions in the SalCom atlas. This highlights the difficulty in reconstructing host-like conditions in a culture flask and underscores the importance of true cellular microbiology approaches. Similar studies have been performed in animal infection models, for example, using anti-*Staphylococcus aureus* antibodies coupled to magnetic beads to enrich these bacteria from a mouse model of osteomyelitis⁷⁵; engineering bioluminescent strains of *Citrobacter rodentium* that enabled their recovery from mouse colon tissue by bioluminescence imaging²⁷; and *Pseudomonas aeruginosa* in the sputum of cystic fibrosis patients⁷⁶. Collectively, these studies revealed new virulence strategies and metabolic adaptations of these pathogens within their host niches.

Even if bacteria are enriched, the resulting transcriptome data may still be dominated by host sequences, especially in *in vivo* studies. To minimize host RNA/cDNA background, bacterial transcripts may be enriched using commercial kits (e.g., MicrobEnrich; Ambion) that deplete eukaryotic rRNA and polyadenylated transcripts. This was used for *in vivo* expression studies of *Vibrio cholerae* in an infant rabbit and mouse model⁷⁷; of *Yersinia pseudotuberculosis* in mouse cecal tissue biopsies²⁶; and of *C. rodentium* recovered upon infection of mouse colon tissue²⁷.

Bacterial transcripts can also be enriched directly. 'Hybrid-selection' refers to the incubation of complex cDNA samples with species-specific, biotinylated probes to capture and

enrich the cDNAs of interest away from the host cDNA background. Hybrid-selection recently enabled niche-specific RNA-seq of the gut commensal *Bacteroides fragilis* within the proximal colon ⁷⁸. This revealed, for the first time, spatially distinct expression profiles of a microbiota member between its luminal, mucus-associated, and epithelial niches. Although the capture probes comprised bait sequences for the complete *B. fragilis* genome, prior knowledge of a bacterium's transcriptome structure can refine probe design. For example, biotinylated probes specific for bacterial mRNAs and sRNAs (omitting rRNA, tRNA, or nontranscribed loci) were employed in a technical study looking at intracellular expression profiles of *P. aeruginosa* or *Mycobacterium tuberculosis* ⁷⁹. Of note, given that enrichment occurs at the cDNA level – i.e. after sample amplification – hybrid-selection is extremely sensitive and works down to the single-cell level ⁷⁹. So far, hybrid-selection of bacterial cDNA has been applied to host cell cultures or gnotobiotic mice after monocolonization with the target bacterium. However, it should also permit the enrichment of target sequences from multi-species samples from conventional mice with a complex microbiota, to reveal potentially protective effects of commensal species.

Cross-species RNA-seq approaches

Steady improvements in RNA-seq technology have allowed researchers to advance beyond one-sided RNA-seq studies and investigate microbes together with other microbes or with host cells, in unprecedented detail, and under physiological conditions (Fig. 4). The following sections will focus on metatranscriptomics within host-associated bacterial consortia, and on dual RNA-seq to read out bacterial and host transcriptomes simultaneously. We conclude the section with an outline of envisaged extensions of cross-species RNA-seq.

Metatranscriptomics. Sequencing-based analysis of the structure of bacterial communities began with metagenomics, i.e., analysis of 16S rDNA sequences (16S profiling), before more sensitive shotgun-sequencing of genomic DNA became feasible. However, the abundance of a certain microbial species or a certain gene does not necessarily correlate with that species' or gene's functional contribution in the consortium. In this regard, metatranscriptomics has been gaining popularity for functional insights into interactions within microbiomes ⁸⁰.

To minimize the risk of RNA degradation during sample preparation, an RNA preserving reagent is often used to 'freeze' metatranscriptomes ^{23,81}. The standardized usage of such agents should facilitate cross-comparisons between different studies. As for conventional RNA-seq, metatranscriptomics typically involves rRNA depletion using the same techniques as used for single-species RNA-seq ³⁵. It needs to be stressed, however, that because most commercial rRNA

removal kits target model bacteria, they might be less efficient for the many non-model species of, say, our gut microbiota. Cas9-based ribosomal cDNA removal ³⁷ on metatranscriptomic samples is yet to be tested; again, it should be particularly suitable for samples with low starting material, e.g., when wanting to profile microbiome activity in an insect gut on the RNA level ⁸².

5 Accurate quantification of gene expression in a microbiome requires calibration of RNA for the DNA content of the same sample, as done in hybrid DNA/RNA-seq analyses ^{23,83-85}. To this end, DNA and RNA samples are **each** sequenced to a depth of 20-40 million Illumina reads. Briefly, the resulting DNA and RNA sequencing data are processed separately, with the metagenomic data then being used to normalize the metatranscriptomic reads. Unsurprisingly,
10 the computational steps are far from trivial and two orthogonal approaches have been developed to this end, namely taxonomic classification involving either reference-based or *de novo* metagenome-assembled genomes (MAGs). Reference-based classifiers such as HUMAnN2 ⁸⁶, Kaiju ⁸⁷ or Kraken2 ⁸⁸ work well when community member species are largely predetermined and only their relative abundances unknown, such as in the human intestinal microbiota. If
15 bacterial species within a sampled community are not known *a priori*, *de novo* taxonomic classification that assembles MAGs from metagenome reads is preferred ⁸⁹.

Comparative studies of healthy human gut microbiomes revealed metatranscriptomic profiles to be generally less individual than their metagenomic counterparts ^{23,83,85}. This argues for subject-specific whole-community regulation, wherein a core set of housekeeping transcripts
20 is universally expressed over individual subjects, but by different microbial species. By contrast, longitudinal profiling showed the metatranscriptome of an individual to be more variable over time than that person's metagenome, suggesting active regulation of microbial gene expression in response to environmental and nutritional fluctuations ^{83,85}. Compositional (metagenomics) and functional (metatranscriptomics) changes in the human intestinal microbiota have also been
25 linked with the outcome of infectious diseases. In a human challenge model, a gut metatranscriptome dominated by mRNAs encoding antioxidants and metal ion homeostasis proteins was linked with a reduced risk of developing symptoms of typhoid fever upon infection with *S. enterica* serovar Typhi ⁹⁰.

30 **Dual RNA-seq.** Transcriptomics is increasingly being harnessed to study interactions across kingdoms, foremost between bacterial pathogens and mammalian hosts ^{5,29}. Dual RNA-seq measures gene expression simultaneously in (intra- or extracellular) bacteria and infected host cells or tissues. As eukaryotic and bacterial material is processed together without prior physical separation (**Fig. 3b**), dual RNA-seq promises the discovery of novel interdependencies in host-

microbe interactions. Technically, dual RNA-seq benefits from commercial kits optimized for parallel depletion of prokaryotic and eukaryotic rRNA; when it comes to the analysis of clinical samples that often offer only minute amounts of RNA, rRNA depletion on the cDNA level (by DASH, see above) should be an attractive strategy. Sequencing depth overall is not significantly different from conventional mammalian RNA-seq, provided that the bacterial RNA proportion in the sample is in the range of 0.5-10%²⁹.

Dual RNA-seq has been extensively used to unveil previously hidden aspects of *Salmonella*-host cell interactions. It revealed how the activities of PinT sRNA, which is highly induced in intracellular *Salmonella*, alters epithelial STAT3 signaling²⁸; linked the global activities of *Salmonella* RNA-binding protein ProQ with MAPK signaling in epithelial host cells⁹¹; and showed that *Salmonella* persists arising during macrophage infection maintain a metabolically active state that allows them to reprogram their host cells⁹² (Fig. 5a). Dual RNA-seq studies have also been performed with cell culture models of infection with *S. pneumoniae*⁹³, uropathogenic *Escherichia coli*⁹⁴, and *Haemophilus influenza*⁹⁵. Dual RNA-seq is particularly suitable to study obligate intracellular pathogens, as demonstrated with *Chlamydia trachomatis*⁹⁶, *M. tuberculosis*^{97,98} (Fig. 5b), and *Orientia tsutsugamushi*⁹⁹ (Fig. 5c). Many of these studies were designed as temporal studies, to follow the kinetics of bacterial and host gene expression during their interaction, or as comparative studies, comparing expression patterns between infections with different bacterial strains. The power of these designs is well-illustrated with the PinT sRNA, for which dual RNA-seq profiling over the course of infection with wild-type versus $\Delta pinT$ bacteria revealed that PinT functions as a post-transcriptional timer in the transition of *Salmonella*'s two major virulence programs with a massive impact on gene expression in infected host cells²⁸.

The application of dual RNA-seq is increasingly shifting from host cell monoculture toward *in vivo*-like infection models. For example, a new three-dimensional intestinal tissue model recapitulates aspects of human gastroenteritis in response to *S. Typhimurium* infection¹⁰⁰. Dual RNA-seq of individual cell types isolated from this model helped to parse out direct and indirect effects of bacterial infection in the epithelial and endothelial compartments. While it corroborated a previously observed connection between *Salmonella*'s type-III secretion systems and the STAT3-dependent inflammatory response¹⁰¹, it revealed that this response stays local in the epithelial lining¹⁰⁰.

Mouse infection models that can closely recapitulate human disease exist for many bacterial pathogens. Generally, bulk RNA-seq data from infected host tissues or organs are difficult to analyze, as altered transcript abundances represent the sum of differentially expressed genes and alterations in host cell type composition. This notwithstanding, dual RNA-

seq has been successfully used with some of these **infection models**, analyzing Peyer's patches, kidneys or lungs of mice infected with the extracellular pathogens *Y. pseudotuberculosis* (Fig. 5d), *S. aureus*, and *P. aeruginosa*, respectively¹⁰²⁻¹⁰⁴. Comparative dual RNA-seq of two different strains of *S. pneumoniae* in **mouse** lung tissue linked a single nucleotide polymorphism in the bacterial raffinose utilization pathway to differential neutrophil recruitment as a potential cause of the divergent disease outcomes¹⁰⁵. Examples of dual RNA-seq revealing how bacteria adapt to and even exploit the human immune response include a human challenge model infected with the extracellular pathogen *Haemophilus ducreyi*¹⁰⁶ or samples from human patients with intracellular *Mycobacterium leprae*³². Others used the technique on human skin biopsies to dissect the pathophysiology of mono- and polymicrobial necrotizing soft tissue infections¹⁰⁷, revealing *in vivo* virulence expression patterns, including an upregulation of invasion genes of *Streptococcus pyogenes* during mono-infection and a pronounced interferon response in the surrounding host tissue.

By contrast, a metatranscriptomic-like analysis of the bacterial expression data from polymicrobial biopsies revealed an elevated expression of **lipopolysaccharide (LPS)** biosynthetic genes and adhesion (but not invasion) factors as compared to mono-infections¹⁰⁷. The data further indicated functional specialization between the co-occurring species, with key metabolic pathways only expressed by a fraction of the community members, thereby suggesting bacterial synergy at the heart of the pathogenicity of polymicrobial soft tissue infections. The host response to polymicrobial infection was dominated by a strong pro-inflammatory profile – potentially as a result of elevated LPS levels – and genes for extracellular matrix components, indicative of activated fibroblasts. Eventually, these divergent host response patterns were exploited to pinpoint biomarkers specific for mono- or polymicrobial infection for early therapeutic intervention.

When sequencing bulk tissue samples, cell-type specific gene expression **may be** lost in the average expression profiles. Resolution can, however, be increased in two ways: pre-sequencing, by dissociating an infected tissue and isolating pre-defined cell types; or post-sequencing, by utilizing cell-type-specific markers to computationally deconvolute the heterogeneous data. A proof-of-principle study of mouse colon tissue infected with *S. Typhimurium* demonstrated that cell type enrichment is generally feasible²⁵. Using a fluorescent *Salmonella* strain (expressing GFP) in combination with selective antibody staining, enabled separation of infected and bystander colonocytes. Another dual RNA-seq study used cell-type-specific antibody staining and enrichment of alveolar and interstitial macrophages from **mouse** lungs infected with *M. tuberculosis*¹⁰⁸. However, since the mycobacteria were not fluorescent, the enriched macrophages were a mixture of infected and uninfected cells. To not out-dilute the

bacterial reads, the sorted host macrophages were selectively lysed, followed by enrichment and mechanical lysis of the mycobacterial cells and subsequent pooling of the host and bacterial lysates. This strategy yielded sufficient numbers of bacterial reads to observe different mycobacterial expression profiles in alveolar versus interstitial macrophages, indicating **that alveolar macrophages represent a conducive environment, whereas interstitial macrophages are** a more hostile environment. However, there are considerable technical challenges posed by lengthy tissue dissociation, antibody staining, and cell sorting, all of which threaten transcriptome integrity. Therefore, samples should ideally be fixed as early in the procedure as technically feasible ²⁵.

As a post-sequencing alternative to physical enrichment of distinct cell types, *in silico* dissection of host expression signatures upon mouse infection with *Y. pseudotuberculosis* revealed extensive neutrophil infiltration into infected Peyer's patches ¹⁰³. Likewise, a cell-type-specific expression signature tool was employed to dissect alterations in the cellular composition of leprosy skin lesions ³². Although not yet combined with **infection studies**, more sophisticated cell-type deconvolution analyses fed with scRNA-seq data ¹⁰⁹⁻¹¹¹ bear great potential to increase the resolution of *in vivo* dual RNA-seq analysis in the future.

Triple RNA-seq. Conceptually speaking, it is a small step from dual RNA-seq to any type of 'multi' RNA-seq to investigate polymicrobial infections, especially with different types of pathogens. The latter is illustrated by a recent triple RNA-seq study in an *ex vivo* dendritic cell model co-infected with human cytomegalovirus and the pathogenic fungus *Aspergillus fumigatus* ¹¹²; this combination of pathogens is a serious medical threat in organ and stem cell transplantation. The very different viral, fungal, and human genomes facilitated the unequivocal assignment of the triple RNA-seq reads with negligible cross-mapping, revealing synergistic pathogen strategies during co-infection and inter-species expression network analyses pinpointing host genes with biomarker potential. The same approach should facilitate a better understanding of co-infections with a bacterium and a virus, which are common and usually lead to a more severe outcome than the respective mono-infections. Promising models to understand how one infection makes the host more susceptible for another include non-typhoidal *Salmonella* and human immunodeficiency virus ^{113,114}, *Streptococcus* spp. and influenza virus ¹¹⁵ or *Chlamydia* spp. and human herpes virus ¹¹⁶.

Toward ‘omni RNA-seq’ approaches. Increasing read lengths in combination with better computational algorithms to resolve RNA-seq data at the taxonomic level enable RNA-seq studies of eukaryotic host cells interacting with more than one bacterium, and ideally, with hundreds of different bacterial species as in the case of the gastrointestinal microbiota (Fig. 4).

Although not done in a single-step experiment yet, some pioneering work towards such ‘omni’ RNA-seq includes transcriptomics of the nasal epithelium of asthmatic children, complemented with independently obtained metatranscriptomics of the nasal microbiome¹¹⁷. Similarly, the pulmonary microbiome was profiled on the RNA level and integrated with host lung transcriptome data in patients with chronic obstructive pulmonary disease (COPD)¹¹⁸. Despite donor variability, the number of patients (8 and 25) and control subjects (6 and 9) in these studies were sufficient to define the pediatric asthma- or COPD-associated microbiome signatures, correlating with a modulation of the host’s immune system (interleukin 1-alpha was associated with asthmatic patients; Th17 response correlated with COPD exacerbation).

Together, the past five years have seen an enormous rise in the popularity of multi-organism RNA-seq approaches (Figs. 1, 4). While metatranscriptomics and dual RNA-seq are on the way to becoming routine techniques in modern infection biology, the development of cross-species RNA-seq will continue, by considering ever more interaction partners within an infected ‘holobiont’¹¹⁹ and trace their complex interaction networks. **Appropriate analysis of the resulting data mass demands for sophisticated bioinformatic tools and should be followed up with experiments seeking to probe causality and functional consequences of the predicted multi-species interaction nodes.** In parallel to these efforts, the field of infection research is currently shaken up by the introduction of single-cell transcriptomics, which will be the next section’s focus.

Single-cell RNA-seq

The recent advances in scRNA-seq have ushered in a new era of transcriptomics, which has already witnessed profound new discoveries, ranging from previously unknown cell types or physiological states to new principles of stochastic gene expression¹²⁰. At the same time, cellular heterogeneity is increasingly understood to play important roles in host-pathogen interactions, e.g., pathogens exploit pre-existing and/or induced cellular heterogeneity to establish an infection niche. Likewise, pathogens themselves can present with substantial phenotypic diversification, e.g., *S. Typhimurium* populations are characterized by a bistable expression of invasion genes already before host cell contact¹²¹, and gene expression variability even intensifies over the course of the infection¹²². Physiologically, such a bet-hedging strategy primes a subpopulation of *Salmonella* for invasion, inducing epithelial inflammation that in turn

benefits the luminal population^{123,124}. New transcriptomic methods going beyond bulk analysis are needed to capture and understand how cellular heterogeneity determines the outcome of host-microbe interactions.

5 **Eukaryotic scRNA-seq.** At this point in time, eukaryotic scRNA-seq has become a routine technique, with a wide variety of experimental and computational pipelines to choose from¹²⁵, and several commercial platforms available. **Although** transcriptome fixation is less common than in bulk sequencing approaches, the first RNA-preserving techniques are now compatible with single-cell applications¹²⁶. Unsurprisingly, scRNA-seq is also making its way into infection
10 biology¹⁰, with the pioneering work again using *S. Typhimurium* to study heterogeneity in populations of infected host cells^{127,128}. This bacterium had been known to display heterogenic replication rates inside macrophages¹²⁹ and *in vivo* tissues¹³⁰.

Using a macrophage infection model, scRNA-seq uncovered a rapid polarization of *Salmonella*-infected host cells that would have been impossible to detect by bulk RNA-seq (**Fig. 5a**). Moreover, it also showed that actively replicating *Salmonella* reside predominantly in anti-inflammatory, infection-permissive M2-like macrophages, non-replicating bacteria primarily dwell in pro-inflammatory M1 macrophages¹²⁸, with bacterial replication rates correlating with the magnitude of the host's interferon response¹²⁷. Subsequent integration of the scRNA-seq data with corresponding (population-level) dual RNA-seq established a link between the
20 expression of a specific *Salmonella* virulence factor and macrophage polarization⁹². In other words, *Salmonella* actively manipulate their host cells to commit development toward a replication-permissive state¹³¹; **a virulence strategy this pathogen seems to share with *M. tuberculosis***¹⁰⁸ (**Fig. 5b**). scRNA-seq data have also become available for human primary immune cells infected with either *S. Typhimurium* **or** *M. tuberculosis*¹³². Moreover, there is a
25 comprehensive atlas of cell-type-specific gene expression signatures and cellular compositional changes within the gut epithelium of a mouse model after *S. Typhimurium* infection¹³³. These single-cell data provide a rich resource for studying heterogeneity in diverse host cell types.

A severe limitation of most *in vivo* scRNA-seq protocols is the loss of spatial information in the process of tissue dissociation. Spatial information, however, is pivotal to interpreting
30 scRNA-seq results from infected samples, because many bacterial pathogens colonize vastly different local microenvironments. By combining photoactivatable reporters and scRNA-seq, NICHE-seq overcomes this limitation and reconstructs the spatial organization of infection niches¹³⁴. Specifically, transgenic mice ubiquitously expressing a photoactivatable GFP variant were infected with lymphocytic choriomeningitis virus and – upon photoactivation of
35 subregions in B-cell follicles or the T-cell area of inguinal lymph nodes with two-photon laser

scanning microscopy – the tissue was dissociated and GFP-positive cells were collected by fluorescence-activated cell sorting and analyzed by massively parallel scRNAseq.

Another common limitation of scRNA-seq protocols is their dependence on oligo(dT)-primed RT, thereby losing information for many non-polyadenylated RNA classes, foremost, small RNAs. To overcome this limitation, a 3' adapter ligation-based method termed “Small-seq” was developed which allows for the profiling of miRNAs, tRNA fragments and snoRNAs in individual mammalian cells^{135,136}. Another method, Holo-seq, captures both small RNA species and mRNAs from single cells by the addition of *in vitro*-transcribed carrier RNA, which allows conventional library construction (as for bulk RNA-seq), and involves enzymatic digestion at the cDNA level prior to sequencing¹³⁷. These developments reflect previous advancements in bulk RNA-seq, emanating from mRNA-centric approaches toward a ‘sequence-all’ strategy. Expression of several miRNAs and lncRNAs responds remarkably rapidly to a pathogenic stimulus, proposing them as suitable biomarkers for diagnostics. Poly(A)-independent scRNA-seq may assess whether some of these noncoding RNAs are heterogeneously expressed and could serve as biomarkers also for specific cellular sub-populations during infection.

Bacterial scRNA-seq. While single-cell transcriptomics began to revolutionize eukaryotic biology, technical hurdles prevented its robust application to bacteria until recently. Since bacterial cells contain only a femtogram amount of RNA¹⁹, i.e., >100-times less than the typical eukaryotic cell, a sensitive cDNA synthesis and amplification protocol is required. Whereas most current eukaryotic scRNA-seq protocols have a lower detection limit of 10 copies of a transcript per cell, bacterial scRNA-seq must take into account a much lower average copy-number of mRNAs (0.4 copies/cell;¹³⁸). The intrinsic lability of bacterial mRNAs (with half-lives in the minute range, as compared to hours in eukaryotes), is another issue, and requires perforation of the bacterial envelope, cell lysis, and subsequent RNA stabilization to be done rapidly. Above all, however, the absence of a poly(A) tail on functional bacterial transcripts precludes oligo(dT)-based RT priming.

With the recent publication of robust bacterial scRNA-seq protocols^{139,140}, single-bacterium transcriptomics have now become a reality for major human pathogens. One study used Multiple Annealing and Tailing-based Quantitative scRNA-seq (MATQ-seq) – originally developed for eukaryotic scRNA-seq¹⁴¹ – to profile individual *S. Typhimurium* and *P. aeruginosa* cells after their isolation by cell sorting¹⁴⁰. Benchmarking with an established bulk RNA-seq compendium for *Salmonella* revealed these single-bacterium transcriptomes to faithfully capture growth-dependent gene expression patterns. The other study introduced so-called Prokaryotic Expression profiling by Tagging RNA *In situ* and sequencing (PETRI-seq) to study

individual Gram-negative (*E. coli*) and Gram-positive (*S. aureus*) bacteria ¹³⁹. PETRI-seq also builds on a protocol previously developed for eukaryotic scRNA-seq; it uses combinatorial indexing to barcode transcripts *in situ*, a method originally known as ‘SPLiT-seq’ for **Split-Pool Ligation-based Transcriptome sequencing** ¹⁴². Both MATQ-seq and PETRI-seq capture in the range of 200-300 different mRNAs per single bacterium, i.e. close to ~5% of all mRNAs in a typical bacterial cell ¹⁴³, but two orders of magnitude up from previous approaches to studying heterogeneity of gene expression in single bacteria with the help of fluorescent reporter genes. Both methods have strengths and weaknesses: MATQ-seq seems to have a lower dropout rate than PETRI-seq, whereas the latter offers much higher throughput. **On a general note, while there are estimates of how many cells (and sequencing reads per cell) are typically required for an informative eukaryotic scRNA-seq study ^{144,145}, bacterial scRNA-seq is too much in its infancy for such general guidelines. In the MATQ- and PETRI-seq studies, 19-27 or 204-875 bacteria, respectively, per condition were sequenced to ~60 million or ~40 million reads per library (without rRNA removal), but it remains unclear if this was sufficient to resolve cellular variability in these samples to saturation. The rule of thumb is that the complexity of the sample under investigation and scope of the experiment will ultimately dictate the required cell numbers and sequencing depths.**

The ramifications of single-bacterium RNA-seq for infection biology **are manifold**. To give just one example, persisters are a dangerous threat in diverse infectious diseases as the respective bacteria withstand antibiotic exposure and cause infection relapse ¹⁴⁶. Within their host niche, however, persisters seem far from being transcriptionally inert ⁹². Bacterial scRNA-seq of pathogenic persisters isolated from infected patients could help study their *in vivo* activities and provide functional insights into when, how, and why they reactivate. We expect technical refinement and automation of steps in the MATQ-seq and PETRI-seq protocols to help bacterial scRNA-seq to become a widely used method for studying microbial pathogens in infection settings.

Perspective

RNA-seq has become a central methodology in the quest to understand the gene expression changes that ensue from host-microbe interactions. As described above, there has been a steady increase in the scope and sensitivity of RNA-seq-based methods. In regards to scope, it is fair to say that while **the protein-centric history of infection biology tends RNA-seq studies** to focus on mRNA expression levels, there is an increasing appreciation of the many noncoding transcripts that may change in either infection partner, i.e., the microbe and the host. For example, in human

cells infected with *Salmonella*, lncRNA profiles changed faster than mRNAs; while in the bacteria themselves, a number of highly conserved sRNAs can be used as proxy for important regulons, inferring which type of stress the pathogen is experiencing²⁸. Beyond their role as putative biomarkers, there is accumulating evidence that individual noncoding RNAs impact the outcome of host-microbe interactions, again with *Salmonella* infection models taking the lead. Certain lncRNAs, for example, decrease susceptibility to *Salmonella* infection^{147,148} and individual miRNAs contribute to host defense against this pathogen¹⁴⁹⁻¹⁵², while others are actively exploited by *Salmonella* to favor pathogenesis¹⁵³. In turn, *Salmonella* itself dispatches an arsenal of its own noncoding RNA elements in the tug-of-war between infection and clearance^{28,154-157}. We expect the noncoding branch of host-microbe interaction studies to further expand in the future, particularly in the context of obligate anaerobic pathogens and commensals that have so far escaped in-depth analysis.

RNA modification is another aspect for expanding the scope of RNA-seq analysis of interspecies interaction. Traditionally known in rRNA and tRNA molecules, RNA modifications are now reported in many other cellular transcript classes, including mRNAs and regulatory noncoding RNAs in both bacteria¹⁵⁸ and mammals¹⁵⁹. Modifications influence base-pairing, conformation and protein-binding properties of RNA molecules, and have been shown to impact bacterial virulence and host-microbe interactions^{160,161}. There is no shortage of protocols to profile RNA modifications in a transcriptome-wide manner for single-sided bacterial or eukaryotic studies^{162,163}. These work flows are typically based on chemical reactions of specific base modifications in the substrate RNA samples, resulting in differential reverse-transcription efficiencies during cDNA synthesis. Besides, direct RNA sequencing using nanopores¹⁶⁴⁻¹⁶⁶ (Box 2) bears the capability to infer modified ribonucleotides from kinetic variation of the electrostatic potential in the pore. It is thus generic and may offer a possibility to simultaneously detect the epitranscriptomes of host and microbe during their interplay.

With respect to sensitivity, ongoing progress in the acquisition^{78,79} and analysis⁸⁶ of bacterial *in vivo* transcriptomes paves the way for an improved appreciation of bacterial gene expression within complex, host-associated consortia. Inasmuch as scRNA-seq may look in its infancy, it has already provided important insights into cellular heterogeneity arising during infection, and it now works for both eukaryotic and microbial cells. It is important to stress that the protocols available could also fill the gap between bulk RNA-seq protocols, which commonly require input RNA in the nanogram to microgram range, and true single-cell studies, enabling robust analysis of biopsy samples from an infected organ or the small populations of bacteria in the gut of an insect infection model. Another application might be the profiling of the RNA content of extracellular vesicles, which are known to be produced by eukaryotic and bacterial

cells alike ^{167,168}. The Extracellular RNA Communication Consortium (ERCC) investigates the roles these RNAs play in intercellular communication and their potential as biomarkers and therapeutic targets ¹⁶⁹. There is, however, also an active debate about the putative function of vesicular RNA in interspecies communication ¹⁷⁰, which would greatly benefit from having experimentally determined RNA cargo profiles from single vesicles, under physiological conditions.

Temporal and spatial resolution is bound to increase further as well. When it comes to deciphering expression kinetics, methods using metabolic labeling allow for discrimination of *de novo*-transcribed from pre-existing RNA molecules, and thus, a better understanding of the order of gene expression ¹⁷¹. This type of tracking *de novo* transcription has been combined with scRNA-seq to provide a fine-grained temporal picture of the early cellular response to a viral attack ¹⁷²; it should be straightforward to apply it to cells infected with bacterial cells as well. In parallel, spatial transcriptomics ¹⁷³, including near-genome-wide fluorescence *in situ* hybridization ¹⁷⁴ and *in situ* RNA-seq approaches that sequence nucleic acids directly in preserved tissue ^{175,176}, are on the rise. The introduction of NICHE-seq ¹³⁴ represented a breakthrough for studying spatial aspects of host-microbe interactions at single-cell resolution. While NICHE-seq is limited to local microenvironments, integrating spatial/temporal expression profiling at single-cell resolution with imaging and bioinformatic tools for higher-order tissue reconstruction has been achieved (reviewed in ^{177,178}). Currently, technical constraints and high costs limit such analyses to small organs from mouse models. However, the potential to reconstruct infection processes in four dimensions at tissue or organ scale is fascinating and would promise new avenues for interdisciplinary infection research.

The ultimate goal for RNA-seq in the present context is to simultaneously profile and correlate gene expression changes in single infected host cells, together with the pathogen. Dual scRNA-seq is in the starting blocks ^{79,179}, but needs to be brought to a genome-wide scale. This should be in reach, given that the now available bacterial scRNA-seq protocols ^{139,140} are poly(A)-independent, thus should be able to capture the full complement of both eukaryotic and bacterial RNA, as does bulk dual RNA-seq ²⁸. Improved by targeted cDNA removal ³⁷ and enrichment ⁷⁹, dual scRNA-seq promises dissection of host-microbe interactions at an unprecedented resolution.

A crucial inherent caveat of any transcriptomic approach is the fact that mRNA levels may not necessarily be correlated with abundance of the respectively encoded proteins, especially so under steady-state conditions ¹³⁸. This problem became very apparent in dual RNA-seq of the obligate intracellular pathogen *O. tsutsugamushi* and human endothelial cells, for which complementary proteomics data were obtained ⁹⁹. Mathematical modeling concluded that

the mRNA expression levels alone resulted in predictions of protein abundance that were slightly better than chance. *Orientia* may be an extreme case, for its high content of transposable elements being associated with pervasive antisense transcription. Indeed, taking the ratio of sense-to-antisense reads into account enhanced the predictability of protein levels from RNA-seq data substantially.

Nevertheless, such examples are a good reminder of the importance of validating RNA-seq-derived findings at the protein level, **through** methods for protein quantification **that** range from low-throughput (for selected key factors by Western blot or flow cytometry) to mid-throughput (multiplex ELISA or ‘immuno-PCR’ with oligonucleotide-conjugated antibodies¹⁸⁰) to high-throughput (see **Box 1**). And yet, even the presence of a protein does not necessarily guarantee its functional importance under the given condition. In this respect, although outside the focus of this review, it is worth mentioning that sequencing-based perturbation screens are gaining traction for high-throughput studies of gene functionality within host-microbe interaction settings¹⁸¹ (**Box 3**). Integrating host-microbe transcriptomics with these functional genomics data will eventually result in testable hypotheses and propose RNA molecules that could serve as future biomarkers or drug targets to combat infectious diseases and microbial disorders.

Figure 1: The history of RNA-seq-based infection research. The timeline can be broadly subdivided into three phases: phase I is characterized by one-sided RNA-seq studies of either host or microbe; phase II is associated with the concept of sequencing multiple transcriptomes together; and phase III was heralded with the introduction of single-cell RNA-seq technology into infection biology. For each phase and each technology, select milestone studies are cited, without claim to completeness. References: 5,23,26,28,59,62,75,83,85,93-96,112,127,128,134,139,140,149,182-192

Figure 2: Comparison of cellular RNA content between bacteria and mammals. Pie charts of the fraction of different gene classes in the genome (left; from RefSeq) or RNA molecules in the transcriptome (from ⁵) in *S. Typhimurium* (upper) and humans (lower). The area of the bacterial pie charts was magnified by the indicated factors; the unmagnified pie charts reflect the **relative size** of the *Salmonella* compared with the human genome/transcriptome. Informative transcript classes (mRNAs, regulatory noncoding RNAs) to deduce cellular states are highlighted (*).

Figure 3

a

b

Figure 3: The basic steps in commonly used protocols for strand-specific bacterial, mammalian, or dual expression profiling.

a, Single-species RNA-seq. The bacterial RNA-seq protocol (left) is based on the NEBNext Small RNA Library Prep Set workflow (New England Biolabs) and the mammalian protocol (right) is based on the TruSeq Stranded mRNA workflow (Illumina). Commonalities are in blue and differences in red boxes. The minimal amounts of total or enriched input RNA are the ones stated in the respective user manuals.

b, Dual RNA-seq protocol based on ²⁴. Sequencing depth estimates stem from ⁴² or ^{43,44}, respectively, for conventional bacterial or mammalian RNA-seq and from ²⁹ for dual RNA-seq.

Figure 4: Graphical overview of RNA-seq-based approaches to study inter-species interactions in the mammalian intestine. (1) Metatranscriptomics ^{23,83,85,90,193}; (2) Dual RNA-seq of an enteric pathogen and infected host tissue ^{102,103}; (3) Triple RNA-seq ¹¹² of viral/bacterial co-infections (3a) or of a bacterial pathogen, a competing commensal, and their host (3b); (4) scRNA-seq of individual host cells ^{127,128} (4a) or single bacteria ^{139,140} (4b); (5) NICHE-seq ¹³⁴ to maintain spatial information about the local microenvironment **at single-cell resolution**.

Figure 5: Molecular aspects of host-pathogen interactions revealed by transcriptomics. a,

S. Typhimurium infection of macrophages **represents** a prime example for the strength of combining different, complementing transcriptomic approaches to dissect the molecular aspects of an infection process from various angles. The global picture that emerges from the combined application of conventional bacterial RNA-seq¹⁹¹, dual RNA-seq^{28,92}, and scRNA-seq^{127,128} argues that intra-macrophage *Salmonella* face oxidative, nitrosative, and envelope stress. Moreover, the intracellular replication rates are variable and follow a gradient that correlates with the expression of *Salmonella* pathogenicity island 2 (SPI-2) as well as PhoP-controlled genes and anti-correlates with SPI-1 genes. The extreme populations along this gradient – persistent or fast-replicating bacteria – reside in pro-inflammatory M1 or anti-inflammatory M2-like macrophages, respectively, and host macrophage polarization seems to be actively induced by secreted *Salmonella* effector proteins as exemplified by SteE¹³¹. **b, A similar virulence strategy is applied by *M. tuberculosis*.** Initial bulk dual RNA-seq of *Mycobacterium*-infected THP-1 macrophages revealed the metabolic flexibility of this intracellular pathogen during infection⁹⁷. Macrophage phenotype-specific dual RNA-seq dissected the mycobacterial-host interactions of interstitial (M1-like) and alveolar (M2-like) mouse macrophages¹⁰⁸. **c, Two isolates of the obligate intracellular pathogen *O. tsutsugamushi* – UT176 and Karp – were used to infect human**

endothelial cells and analyzed by dual RNA-seq, revealing a common interferon-based host response in addition to strain-specific immune signaling cascades, likely as a result of differential expression of virulence factors between the two bacterial strains ⁹⁹. **d**, Bacterial RNA-seq of *Y. pseudotuberculosis* in small cecal tissue biopsies revealed an early induction of T3SS genes during the acute phase and their repression when bacteria adopt a persistent state ²⁶. Tissue dual RNA-seq of this pathogen within ileal Peyer's patches recapitulated the early induction of *Yersinia* T3SS – a mechanism dependent on the carbon storage regulator (Csr) system – and identified the host's immune response to infection to be dominated by infiltrating neutrophils ¹⁰³. ROS, reactive oxygen species; RNS, reactive nitrogen species; IFN, interferon; T3SS, type-III secretion system.

Box 1: Orthogonal ‘omics’ technologies and their use for host-microbe interaction studies.

Genomics and epigenomics. Omics techniques aim at the global detection of biomolecules in an organism to infer biological insights (see the Table). Genomics refers to the determination of the genetic content of an organism by sequencing its genomic DNA. Comparative genomics was
5 harnessed to unravel similarities and differences within the (virulence) gene content of related bacterial species ¹⁹⁴, to study human host adaptation of bacterial pathogens ¹⁹⁵, to trace the spread of infectious diseases ¹⁹⁶, or to uncover single nucleotide polymorphisms (SNPs) and insertions/deletions (indels) that render human subjects or whole populations more susceptible to certain infections ¹⁹⁷. However, genomics is also moving away from single-sided studies and
10 striving to set host and pathogen in context, by profiling host genetic susceptibility and pathogen genetic variation together ¹⁹⁸.

Whereas the genetic blueprint defines the molecular process that a given organism is theoretically capable of, it fails to determine the genes which are currently activated under the given condition. Chemical chromatin modifications are marks of the transcriptional activity at a
15 given locus. Technological breakthroughs in deep-sequencing technologies (see Box 2) now enable the detection of modified DNA bases and gave rise to the field of host-pathogen epigenomics ^{199,200}. More commonly, RNA serves as direct readout of transcriptional activity; it should be noted, however, that a cell’s transcriptome always reflects the sum of *de novo* transcription and transcript decay. While mRNA levels are typically used as a proxy for protein
20 expression, there may be cases where this assumption is violated (see main text). Rather, ribosome profiling (Ribo-seq) or mass spectrometry (MS) provide insight into the composition of a cellular proteome at any given point in time.

Translatomics and proteomics. Ribo-seq has long been restricted to cultivatable species and has traditionally been used on pure bacterial cultures ²⁰¹; however, technological improvements
25 – particularly with respect to input material constraints – now enable ‘MetaRibo-seq’ studies of complex bacterial consortia such as the intestinal microbiota ²⁰². Ribo-seq can also be applied to mammalian cells, and even dual Ribo-seq studies, e.g. in diverse virus-infected cells ²⁰³ and of human fibroblasts infected with a eukaryotic parasite ²⁰⁴, were recently reported. However, divergence in the technical details, foremost different compounds being needed to stall the
30 prokaryotic and eukaryotic translation apparatus, has so far prevented cross-kingdom Ribo-seq of bacteria-infected host cells.

While Ribo-seq measures global translation rates, label-free and label-based MS detect translation products directly, providing an accurate reflection of the composition of a cellular proteome at any given point in time ²⁰². However, while RNA-seq-based approaches are capable
35 of discovering novel, previously unannotated transcripts, MS generally relies on reference data,

limiting proteomics to the detection of known proteins, and due to technical limitations, conventional proteomics bias heavily against the detection of membrane proteins ²⁰⁵. Additionally, due to codon degeneracy, the resolving power of MS is inferior to RNA-seq, and, lacking the possibility to amplify peptides prior to MS, current proteomics cannot reach the sensitivity of genomics, transcriptomics or Ribo-seq for which the input nucleic acids may be amplified at will. For example, metaproteomics that profile protein levels in complex bacterial consortia (e.g. the human fecal microbiome) typically detect only a tiny fraction of the total number of proteins present in that sample (only the most abundant 0.1-0.0001%) ²⁰⁶. Consequently, host-microbe proteomics – measuring bacterial and eukaryotic protein levels simultaneously – is uncommon. Nevertheless, one-sided infection studies exist that profiled e.g. *S. Typhimurium*'s proteome changes during host infection ²⁰⁷ or determined host factors required for intra-epithelial *Salmonella* replication ²⁰⁸ or cell death of *Salmonella*-infected macrophages ²⁰⁹. Proteomics methods also lend themselves to map protein-protein interaction networks during infection, e.g. interactions of secreted *Salmonella* effectors with host target proteins ²¹⁰. Thus, proteomic approaches to host-microbe interaction can complement transcriptomics ²¹¹, and both RNA-seq and MS are being increasingly used in the clinics, e.g. to accelerate diagnosis of bacterial infections ^{212,213}.

Metabolomics. Not all proteins lingering in a cell may be functioning. So-called activity-based protein profiling detects active enzymes in biological samples and has been successfully applied to shed light on the enzymatic cross-talk of host and pathogen during enteric infection ²¹⁴. Additionally, metabolomics gained attraction as it offers an even more direct readout of a cell's phenotype. Mirroring infection proteomics, host-pathogen metabolomics is a blooming field ²¹⁵. One-sided metabolomics was performed to measure host ²¹⁶ or – upon physical separation – pathogen metabolic flux changes during an infection ²¹⁷. Analyzing host and microbial metabolism in parallel is challenging, given that many core metabolites are shared across kingdoms and cannot be unequivocally assigned to their source organism. This difficulty may be partially overcome by intelligent experimental design that involves the use of axenic cultures or spike-ins, or by complementing host-pathogen metabolomics with orthogonal information. For example, a recent study combined host-pathogen metabolomics with dual RNA-seq during early infection of human macrophages by *M. tuberculosis* and uncovered the robustness of this obligate intracellular pathogen in the face of metabolic interventions ⁹⁷. Similarly, metabolomics and metagenomics were measured in parallel in lung samples of patients with cystic fibrosis to reconstruct the metabolic flow during disease ²¹⁸.

Multi-omics. More generally, the progress made in the individual omics techniques fosters such integrative multi-omics approaches to dissect host-microbe interactions from multiple angles.

For instance, in a *Salmonella* infection study of mice, metagenomics, proteomics, and metabolomics of fecal samples were combined to link infection with gut microbiota compositional and metabolic changes ²¹⁹. And in an extremely comprehensive survey, microbial metagenomics, -transcriptomics, -proteomics, and -metabolomics were integrated with host expression profiling to unravel the molecular aspects underlying dysbiosis in the human gut microbiome during inflammatory bowel disease ²²⁰. Integrative omics, however, are demanding new bioinformatic tools to take full advantage of the enormous amount of information buried within the massive datasets ²²¹.

Table: Side-by-side comparison of the properties, strengths, and limitations of the different omics methods.

	genomics	transcriptomics	proteomics	metabolomics
target biomolecules	genomic DNA	mRNA, ncRNA	protein	metabolites
key technologies	DNA-seq	RNA-seq	LC-MS/MS (label-based or label-free)	MS, NMR (targeted or untargeted)
sensitivity	high	high	low	low
resolution	high	high	intermediate	low
From genotype to phenotype	genetic blueprint	gene expression (RNA level)	gene expression (protein level)	direct phenotypic readout
cost per sample	low	low	intermediate	high

Box 2: Alternative cDNA library preparation methods and non-Illumina sequencing technologies for cross-species RNA-seq.

RNAtag-seq. The primary aim of most RNA-seq studies is to determine the set of differentially expressed genes in an organism in response to an experimental stimulus. For differential expression profiling, however, full transcript coverage is not needed, resulting in the fact that many RNA-seq studies accumulate more information than is actually intended. In RNAtag-seq²²² a barcoded adapter is ligated to the 3' end of input RNA fragments, thus allowing for multiplexing early during library construction. This represents a time- and cost-efficient alternative to standard cDNA library generation protocols. RNAtag-seq works with both prokaryotic and eukaryotic input RNA and since counting the resulting sequencing reads is sufficient to call differentially expressed genes, it may be incorporated into cross-kingdom RNA-seq approaches.

SEnd-seq. There are opposite cases when ensemble RNA levels are insufficient, but when the detection of full-length transcripts is key to interpret biological processes. Various methods exist that map transcription start sites in bacteria²²³ or eukaryotes²²⁴, and alternative methods that determine 3' ends across bacterial⁶² or eukaryotic transcriptomes²²⁵ are available. Recently, simultaneous 5' and 3' end sequencing (SEnd-seq) was developed that maps both transcription start and end sites in parallel, in the same experiment²²⁶. Developed for *E. coli*, the protocol is generic and should be transferable to metatranscriptomics. However, given that SEnd-seq determines only the extreme 5' and 3' ends of an input cDNA molecule, the protocol is less suited to deduce the full-length mRNA structure in an organism where splicing occurs.

Long-read sequencing. Illumina short-read sequencing technology relies on transcript fragmentation, thereby inevitably losing the complete nucleotide composition of individual transcripts. As alternatives to the above, Illumina-based methods, long-read sequencing platforms – such as Pacific Biosciences (which sequences full-length, unfragmented cDNA molecules) or Oxford Nanopore (with the possibility to directly sequence RNA molecules, omitting the need for RT) – are entering the market^{20,21}. Albeit the current read depth, error rate, and cost of these platforms cannot compete with Illumina technology, they have the ability to read a cDNA/RNA molecule from one end to the other and bear great potential for future bacterial²²⁷ and eukaryotic²²⁸ (epi)transcriptomics.

Box 3: Sequencing-based high-throughput functional screening for genetic factors shaping host-microbe interactions.

Deep-sequencing proved useful for high-throughput genetic perturbation screens to uncover bacterial or host factors contributing to infection that may eventually be exploited as therapeutic targets. For instance, random mutagenesis followed by fitness screening of the resultant mutant pool can be applied to identify factors that are functionally important under the given selection pressure. Transposon insertion sequencing²²⁹ has been widely harnessed to uncover virulence factors of diverse bacterial pathogens, including *S. Typhimurium*, *H. influenza*, *V. cholerae*, *N. meningitidis*, *Streptococcus* spp., *Legionella pneumophila*, *Acinetobacter baumannii*, *Bordetella pertussis*, *Brucella abortus*, *Coxiella burnetii*, and *Enterococcus faecium*²³⁰⁻²⁴³.

Random mutagenesis has also been applied for functional genomics in mice and in human cells^{244,245}. Given their larger genomes, however, genetic screens in mammals are typically performed in a more targeted manner. Initially, RNA interference screens were employed to identify host factors that influence their infection by bacterial pathogens; however, such screens typically yielded a high number of false-positives, likely due to ‘off-targeting’ by the small interfering RNA pool²⁴⁶. This drawback is overcome by recent Clustered Regularly-Interspaced Short Palindromic Repeats (CRISPR)-based screens that have since been intensely used to address infection-related questions²⁴⁷. For human host models, multiplexed pools of single-guide RNAs are available²⁴⁸. This facilitates genome-wide screening approaches, e.g. to identify host factors conferring resistance against intoxication by the Shiga toxin of enterohemorrhagic *E. coli*²⁴⁹, against the type-III secretion system of *Vibrio parahaemolyticus*²⁵⁰, the typhoid toxin of *S. Typhi*²⁵¹ or the *S. aureus* toxin leukocidin²⁵², and to uncover the colon epithelial receptor of *Clostridium difficile* toxin B²⁵³.

In parallel, CRISPR tools have been developed to deliberately perturb target gene sets in bacteria²⁵⁴ and can be exploited for *in vivo* screens of bacterial pathogenesis^{255,256}. We predict CRISPR-based perturbation screens to become a key technology for the discovery of novel bacterial colonization or virulence factors, especially amongst all the short genes^{257,258} that are less likely to be hit by random mutagenesis. Additionally, CRISPR technologies arise – in parallel with antisense oligonucleotides (see main text) – as novel species-specific antibiotic candidates and programmable tools for microbiota editing²⁵⁹⁻²⁶¹.

Glossary definitions

Gnotobiosis: Greek for “known life”. The term generally describes biological systems wherein all present organisms can be accounted for. In the present context, the term refers to ex-germ-free mice that were inoculated and colonized with a defined bacterial species or consortium.

- 5 **Metatranscriptomics:** A method to detect and quantify steady-state transcript levels from multiple bacterial species within a consortium present in a given environmental or host-derived sample.

Axenic culture: Describes a culture comprised of only a single, defined bacterial species or strain.

10

Highlight references (ordered by publication date)

1. Westermann et al., Dual RNA-seq of pathogen and host. Nature Reviews Microbiology 2012.

A study describing a thought experiment that explores the concept of multi-organismal RNA-seq and coining the term “dual RNA-seq”.

2. Avraham et al. Pathogen Cell-to-Cell Variability Drives Heterogeneity in Host Immune Responses. Cell 2015.

Combined eukaryotic single-cell and dual RNA-seq study that associated heterogeneity in *Salmonella* PhoP activity with interferon signaling in infected mouse macrophages.

3. Westermann et al. Dual RNA-seq unveils noncoding RNA functions in host-pathogen interactions. Nature 2016.

Comprehensive evaluation of the dual RNA-seq technology for fourteen different mammalian host cell types infected with *Salmonella* and identification of virulence-related noncoding RNAs in this bacterial pathogen.

4. Saliba et al. Single-cell RNA-seq ties macrophage polarization to growth rate of intracellular *Salmonella*. Nat Microbiol. 2016.

One of the first applications of eukaryotic single-cell RNA-seq to an infection model; revealed *Salmonella*-infected macrophages to differentiate into divergently polarized phenotypes with consequences for the intracellular behavior of this pathogen.

5. Dar et al., Term-seq reveals abundant ribo-regulation of antibiotics resistance in bacteria. Science 2016.

The authors established Term-seq to map prokaryotic RNA 3' ends on genome-wide scale and, among others, applied the method to human oral microbiome samples.

6. Nuss et al. Tissue Dual RNA-seq Allows Fast Discovery of Infection-Specific Functions and Riboregulators Shaping Host-Pathogen Transcriptomes. PNAS 2017.

One of the first applications of Dual RNA-seq to an *in-vivo* setting, analyzing host and bacterial gene expression during the infection of mouse Peyer's patches with *Yersinia pseudotuberculosis*.

7. Zimmermann et al., Integration of Metabolomics and Transcriptomics Reveals a Complex Diet of *Mycobacterium tuberculosis* during Early Macrophage Infection. mSystems 2017.

The study combined dual RNA-seq with host-pathogen metabolomics for a cell-culture model of *Mycobacterium tuberculosis* infection and inferred a system-wide host-pathogen metabolic network.

8. Medaglia et al. Spatial Reconstruction of Immune Niches by Combining Photoactivatable Reporters and scRNA-seq. Science 2017.

This study introduced Niche-seq, which combines eukaryotic gene expression analysis at single-cell resolution with spatial information to dissect cellular and molecular aspects of infection niches.

9. Franzosa et al., Species-level functional profiling of metagenomes and metatranscriptomes. Nature Methods 2018.

Describes the launch of HUMAnN2, i.e. an algorithm to functionally analyze metatranscriptomic data from host-associated and environmental bacterial communities.

10. Mehta et al. Stability of the human faecal microbiome in a cohort of adult men. Nat Microbiol 2018.

Analysis of the human gut microbiome from several hundred subjects revealed the stability of the metatranscriptome over time, highlighting its potential for diagnostics.

11. Abu-Ali et al. Metatranscriptome of human faecal microbial communities in a cohort of adult men. Nat Microbiol 2018.

Comprehensive analysis of the human gut microbiome with focus on ecological and molecular aspects and defining a 'core' and 'variable' metatranscriptome across participants.

12. Stapels et al. Salmonella persists undermine host immune defenses during antibiotic treatment. Science 2018.

Dual RNA-seq of *Salmonella*-infected mouse macrophages linked the polarization toward the M2-like macrophage phenotype with the secreted bacterial effector protein SteE.

13. Montoya et al. Dual RNA-Seq of Human Leprosy Lesions Identifies Bacterial Determinants Linked to Host Immune Response. Cell Reports 2019.

Dual RNA-seq study that measured gene expression of host and pathogen in human leprosy skin lesions, revealing that the human immune response is not primarily shaped by the bacterial dose, but by the virulence programs active in the infecting *Mycobacteria*.

14. Griesenauer et al. Determination of an Interaction Network Between an Extracellular Bacterial Pathogen and the Human Host. mBio 2019.

Combined dual RNA-seq and metabolomics study of punch biopsies from human volunteers inoculated with the bacterial skin pathogen *Haemophilus ducreyi*.

15. Thanert et al., Molecular profiling of tissue biopsies reveals unique signatures associated with streptococcal necrotizing soft tissue infections. Nature Communications 2019.

Comprehensive host-bacterial transcriptomic analysis of necrotizing soft tissue biopsies from human patients with either *Streptococcus pyogenes* mono- or polymicrobial infections.

16. Fremin et al., MetaRibo-Seq measures translation in microbiomes. Nature Communications 2019.

Sensitive protocol for the performance of MetaRibo-seq to globally profile bacterial translation in complex consortia such as the human fecal microbiota.

17. Pisu et al. Dual RNA-Seq of Mtb-Infected Macrophages In Vivo Reveals Ontologically Distinct Host-Pathogen Interactions. Cell Reports 2020.

- 5 ***In-vivo* dual RNA-seq of mouse lungs infected with *Mycobacterium tuberculosis* that revealed the segregation of host-pathogen transcriptomes dependent on macrophage phenotype.**

18. Mika-Gospodorz et al., Dual RNA-seq of *Orientia tsutsugamushi* informs on host-pathogen interactions for this neglected intracellular human pathogen. Nature Communications 2020.

Dual RNA-seq unveiled differences in the pathogenicity of two strains of the obligate intracellular pathogen *Orientia tsutsugamushi* in an infected cell culture model.

19. Minhas et al., In vivo dual RNA-seq reveals that neutrophil recruitment underlies differential tissue tropism of *Streptococcus pneumoniae*. Communications biology 2020.

- 15 **The authors devised a comparative dual RNA-seq approach to identify the impact of a single nucleotide polymorphism in a *Streptococcus* raffinose utilization regulator gene on infection outcome in a mouse model.**

20. Donaldson et al., Spatially distinct physiology of *Bacteroides fragilis* within the proximal colon of gnotobiotic mice. Nature Microbiology 2020.

- 20 **The study employed hybrid-selection RNA-seq to enrich *Bacteroides fragilis* transcripts from host-bacterial RNA mixtures, to compare *in-vivo* transcriptome signatures of this gut commensal between different host niches.**

21. Blattman et al., Prokaryotic single-cell RNA sequencing by in situ combinatorial indexing. Nature Microbiology 2020.

- 25 **First demonstration of near-genome-wide bacterial single-cell RNA-seq that applied a SPLiT-seq-related protocol to individual *E. coli* and *S. aureus* cells.**

22. Imdahl et al., Single-cell RNA-sequencing reports growth-condition-specific global transcriptomes of individual bacteria. Nature Microbiology 2020.

- 30 **Proof-of-concept study that successfully applied MATQ-seq-based single-bacterium RNA-seq to *Salmonella* and *Pseudomonas*.**

23. Seelbinder et al., Triple RNA-seq reveals synergy in a human virus-fungus co-infection model. Cell Reports 2020.

First (and so far only) Triple RNA-seq study; revealed molecular basis of synergistic virulence strategies of two frequently co-occurring pulmonary pathogens.

REFERENCES

- 1 Bumann, D. Heterogeneous host-pathogen encounters: act locally, think globally. *Cell Host Microbe* **17**, 13-19, doi:10.1016/j.chom.2014.12.006 (2015).
- 2 Sarkar, S. & Heise, M. T. Mouse Models as Resources for Studying Infectious Diseases. *Clinical therapeutics* **41**, 1912-1922, doi:10.1016/j.clinthera.2019.08.010 (2019).
- 5 3 Barrila, J. *et al.* Modeling Host-Pathogen Interactions in the Context of the Microenvironment: Three-Dimensional Cell Culture Comes of Age. *Infect Immun* **86**, doi:10.1128/IAI.00282-18 (2018).
- 4 Colgan, A. M., Cameron, A. D. & Kroger, C. If it transcribes, we can sequence it: mining the complexities of host-pathogen-environment interactions using RNA-seq. *Curr Opin Microbiol* **36**, 37-46, doi:10.1016/j.mib.2017.01.010 (2017).
- 10 5 Westermann, A. J., Gorski, S. A. & Vogel, J. Dual RNA-seq of pathogen and host. *Nat Rev Microbiol* **10**, 618-630, doi:nrmicro2852 10.1038/nrmicro2852 (2012).
- 6 Eulalio, A., Schulte, L. & Vogel, J. The mammalian microRNA response to bacterial infections. *RNA Biol* **9**, 742-750, doi:10.4161/rna.20018 (2012).
- 15 7 Agliano, F., Rathinam, V. A., Medvedev, A. E., Vanaja, S. K. & Vella, A. T. Long Noncoding RNAs in Host-Pathogen Interactions. *Trends in immunology* **40**, 492-510, doi:10.1016/j.it.2019.04.001 (2019).
- 8 Westermann, A. J. Regulatory RNAs in Virulence and Host-Microbe Interactions. *Microbiology spectrum* **6**, doi:10.1128/microbiolspec.RWR-0002-2017 (2018).
- 20 9 Saliba, A. E., Westermann, A. J., Gorski, S. A. & Vogel, J. Single-cell RNA-seq: advances and future challenges. *Nucleic Acids Res* **42**, 8845-8860, doi:10.1093/nar/gku555 (2014).
- 10 Penaranda, C. & Hung, D. T. Single-Cell RNA Sequencing to Understand Host-Pathogen Interactions. *ACS infectious diseases* **5**, 336-344, doi:10.1021/acsinfecdis.8b00369 (2019).
- 25 11 Dreyfus, M. & Regnier, P. The poly(A) tail of mRNAs: Bodyguard in eukaryotes, scavenger in bacteria. *Cell* **111**, 611-613, doi:Doi 10.1016/S0092-8674(02)01137-6 (2002).
- 12 Wolin, S. L. & Steitz, J. A. Genes for two small cytoplasmic Ro RNAs are adjacent and appear to be single-copy in the human genome. *Cell* **32**, 735-744, doi:10.1016/0092-8674(83)90059-4 (1983).
- 30 13 Sim, S. & Wolin, S. L. Bacterial Y RNAs: Gates, Tethers, and tRNA Mimics. *Microbiology spectrum* **6**, doi:Unsp Rwr-0023-2018 10.1128/Microbiolspec.Rwr-0023-2018 (2018).
- 14 Lundblad, E. W. & Altman, S. Inhibition of gene expression by RNase P. *New Biotechnol* **27**, 212-221, doi:10.1016/j.nbt.2010.03.003 (2010).
- 15 Kroger, C. *et al.* An infection-relevant transcriptomic compendium for *Salmonella enterica* Serovar Typhimurium. *Cell Host Microbe* **14**, 683-695, doi:S1931-3128(13)00411-3 10.1016/j.chom.2013.11.010 (2013).
- 35 16 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**, doi:10.1128/ecosalplus.ESP-0030-2019 (2020).
- 40 17 Rion, N. & Ruegg, M. A. LncRNA-encoded peptides: More than translational noise? *Cell Res* **27**, 604-605, doi:10.1038/cr.2017.35 (2017).
- 18 Yao, R. W., Wang, Y. & Chen, L. L. Cellular functions of long noncoding RNAs. *Nat Cell Biol* **21**, 542-551, doi:10.1038/s41556-019-0311-8 (2019).
- 19 Milo, R. & Phillips, R. Cell Biology by the Numbers. *Garland Science* (2015).
- 45 20 Stark, R., Grzelak, M. & Hadfield, J. RNA sequencing: the teenage years. *Nat Rev Genet* **20**, 631-656, doi:10.1038/s41576-019-0150-2 (2019).
- 21 van Dijk, E. L., Jaszczyszyn, Y., Naquin, D. & Thermes, C. The Third Revolution in Sequencing Technology. *Trends Genet* **34**, 666-681, doi:10.1016/j.tig.2018.05.008 (2018).

22 Cox, M. L. *et al.* Investigating fixative-induced changes in RNA quality and utility by
microarray analysis. *Exp Mol Pathol* **84**, 156-172, doi:10.1016/j.yexmp.2007.11.002 (2008).

23 Franzosa, E. A. *et al.* Relating the metatranscriptome and metagenome of the human gut.
Proc Natl Acad Sci U S A **111**, E2329-2338, doi:10.1073/pnas.1319284111 (2014).

5 24 Westermann, A. J. & Vogel, J. Host-Pathogen Transcriptomics by Dual RNA-Seq. *Methods Mol
Biol* **1737**, 59-75, doi:10.1007/978-1-4939-7634-8_4 (2018).

25 Fronicke, L. *et al.* Toward Cell Type-Specific In Vivo Dual RNA-Seq. *Methods Enzymol* **612**,
505-522, doi:10.1016/bs.mie.2018.08.013 (2018).

26 Avican, K. *et al.* Reprogramming of Yersinia from virulent to persistent mode revealed by
10 complex in vivo RNA-seq analysis. *PLoS Pathog* **11**, e1004600,
doi:10.1371/journal.ppat.1004600 (2015).

27 Connolly, J. P. R. *et al.* Host-associated niche metabolism controls enteric infection through
fine-tuning the regulation of type 3 secretion. *Nature Communications* **9**, doi:Artn 4187
10.1038/S41467-018-06701-4 (2018).

15 28 Westermann, A. J. *et al.* Dual RNA-seq unveils noncoding RNA functions in host-pathogen
interactions. *Nature* **529**, 496-501, doi:10.1038/nature16547 (2016).

29 Westermann, A. J., Barquist, L. & Vogel, J. Resolving host-pathogen interactions by dual RNA-
seq. *PLoS Pathog* **13**, e1006033, doi:10.1371/journal.ppat.1006033 (2017).

30 Jin, D. J., Cagliero, C. & Zhou, Y. N. Growth rate regulation in Escherichia coli. *Fems
20 Microbiology Reviews* **36**, 269-287, doi:10.1111/j.1574-6976.2011.00279.x (2012).

31 Goodfellow, S. J. & Zomerdiijk, J. C. Basic mechanisms in RNA polymerase I transcription of
the ribosomal RNA genes. *Subcell Biochem* **61**, 211-236, doi:10.1007/978-94-007-4525-4_10
(2013).

32 Montoya, D. J. *et al.* Dual RNA-Seq of Human Leprosy Lesions Identifies Bacterial
25 Determinants Linked to Host Immune Response. *Cell Reports* **26**, 3574-+,
doi:10.1016/j.celrep.2019.02.109 (2019).

33 Sintsova, A. *et al.* Genetically diverse uropathogenic Escherichia coli adopt a common
transcriptional program in patients with UTIs. *Elife* **8**, doi:ARTN e49748 10.7554/eLife.49748
(2019).

30 34 Culviner, P. H., Guegler, C. K. & Laub, M. T. A Simple, Cost-Effective, and Robust Method for
rRNA Depletion in RNA-Sequencing Studies. *mBio* **11**, doi:10.1128/mBio.00010-20 (2020).

35 Giannoukos, G. *et al.* Efficient and robust RNA-seq process for cultured bacteria and complex
community transcriptomes. *Genome Biol* **13**, R23, doi:10.1186/gb-2012-13-3-r23 (2012).

36 Petrova, O. E., Garcia-Alcalde, F., Zampaloni, C. & Sauer, K. Comparative evaluation of rRNA
35 depletion procedures for the improved analysis of bacterial biofilm and mixed pathogen
culture transcriptomes. *Sci Rep-Uk* **7**, doi:Artn 41114 10.1038/Srep41114 (2017).

37 Prezza, G. *et al.* Improved bacterial RNA-seq by Cas9-based depletion of ribosomal RNA
reads. *RNA*, doi:10.1261/rna.075945.120 (2020).

38 Gu, W. *et al.* Depletion of Abundant Sequences by Hybridization (DASH): using Cas9 to
40 remove unwanted high-abundance species in sequencing libraries and molecular counting
applications. *Genome Biol* **17**, 41, doi:10.1186/s13059-016-0904-5 (2016).

39 Zheng, G. *et al.* Efficient and quantitative high-throughput tRNA sequencing. *Nat Methods* **12**,
835-837, doi:10.1038/nmeth.3478 (2015).

40 Cozen, A. E. *et al.* ARM-seq: AlkB-facilitated RNA methylation sequencing reveals a complex
45 landscape of modified tRNA fragments. *Nat Methods* **12**, 879-884, doi:10.1038/nmeth.3508
(2015).

41 Shigematsu, M. *et al.* YAMAT-seq: an efficient method for high-throughput sequencing of
mature transfer RNAs. *Nucleic Acids Res* **45**, e70, doi:10.1093/nar/gkx005 (2017).

42 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, doi:10.1186/1471-2164-13-734 (2012).

43 Liu, Y., Zhou, J. & White, K. P. RNA-seq differential expression studies: more sequence or
5 more replication? *Bioinformatics* **30**, 301-304, doi:10.1093/bioinformatics/btt688 (2014).

44 Ching, T., Huang, S. & Garmire, L. X. Power analysis and sample size estimation for RNA-Seq differential expression. *RNA* **20**, 1684-1696, doi:10.1261/rna.046011.114 (2014).

45 Engstrom, P. G. *et al.* Systematic evaluation of spliced alignment programs for RNA-seq data. *Nat Methods* **10**, 1185-1191, doi:10.1038/nmeth.2722 (2013).

10 46 Dillies, M. A. *et al.* A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. *Briefings in bioinformatics* **14**, 671-683, doi:10.1093/bib/bbs046 (2013).

47 Berghoff, B. A., Karlsson, T., Kallman, T., Wagner, E. G. H. & Grabherr, M. G. RNA-sequence data normalization through in silico prediction of reference genes: the bacterial response to
15 DNA damage as case study. *BioData mining* **10**, 30, doi:10.1186/s13040-017-0150-8 (2017).

48 Munro, S. A. *et al.* Assessing technical performance in differential gene expression experiments with external spike-in RNA control ratio mixtures. *Nat Commun* **5**, 5125, doi:10.1038/ncomms6125 (2014).

49 Consortium, S. M.-I. A comprehensive assessment of RNA-seq accuracy, reproducibility and
20 information content by the Sequencing Quality Control Consortium. *Nat Biotechnol* **32**, 903-914, doi:10.1038/nbt.2957 (2014).

50 Loven, J. *et al.* Revisiting global gene expression analysis. *Cell* **151**, 476-482, doi:10.1016/j.cell.2012.10.012 (2012).

51 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for
25 RNA-seq data with DESeq2. *Genome Biol* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).

52 Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140, doi:10.1093/bioinformatics/btp616 (2010).

53 Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. voom: Precision weights unlock linear model
30 analysis tools for RNA-seq read counts. *Genome Biol* **15**, R29, doi:10.1186/gb-2014-15-2-r29 (2014).

54 Garber, M., Grabherr, M. G., Guttman, M. & Trapnell, C. Computational methods for transcriptome annotation and quantification using RNA-seq. *Nat Methods* **8**, 469-477, doi:10.1038/nmeth.1613 (2011).

35 55 Gene Ontology, C. Gene Ontology Consortium: going forward. *Nucleic Acids Res* **43**, D1049-1056, doi:10.1093/nar/gku1179 (2015).

56 Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res* **44**, D457-462, doi:10.1093/nar/gkv1070 (2016).

40 57 Caspi, R. *et al.* The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res* **44**, D471-480, doi:10.1093/nar/gkv1164 (2016).

58 Breuer, K. *et al.* InnateDB: systems biology of innate immunity and beyond--recent updates and continuing curation. *Nucleic Acids Res* **41**, D1228-1233, doi:10.1093/nar/gks1147 (2013).

45 59 Sharma, C. M. *et al.* The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature* **464**, 250-255, doi:nature08756 10.1038/nature08756 (2010).

60 Kroger, C. *et al.* The transcriptional landscape and small RNAs of *Salmonella enterica* serovar Typhimurium. *Proc Natl Acad Sci U S A* **109**, E1277-1286, doi:1201061109 10.1073/pnas.1201061109 (2012).

61 Slager, J., Aprianto, R. & Veening, J. W. Deep genome annotation of the opportunistic human
pathogen *Streptococcus pneumoniae* D39. *Nucleic Acids Res* **46**, 9971-9989,
doi:10.1093/nar/gky725 (2018).

62 Dar, D. *et al.* Term-seq reveals abundant ribo-regulation of antibiotics resistance in bacteria.
5 *Science* **352**, aad9822, doi:10.1126/science.aad9822 (2016).

63 Lallane, J. B. *et al.* Evolutionary Convergence of Pathway-Specific Enzyme Expression
Stoichiometry. *Cell* **173**, 749-761 e738, doi:10.1016/j.cell.2018.03.007 (2018).

64 Kroger, C. *et al.* The primary transcriptome, small RNAs and regulation of antimicrobial
resistance in *Acinetobacter baumannii* ATCC 17978. *Nucleic Acids Res* **46**, 9684-9698,
10 doi:10.1093/nar/gky603 (2018).

65 Dugar, G. *et al.* High-resolution transcriptome maps reveal strain-specific regulatory features
of multiple *Campylobacter jejuni* isolates. *PLoS Genet* **9**, e1003495,
doi:10.1371/journal.pgen.1003495 (2013).

66 Albrecht, M. *et al.* The transcriptional landscape of *Chlamydia pneumoniae*. *Genome Biol* **12**,
15 R98, doi:gb-2011-12-10-r98 10.1186/gb-2011-12-10-r98 (2011).

67 Thomason, M. K. *et al.* A rhII 5' UTR-Derived sRNA Regulates RhIR-Dependent Quorum
Sensing in *Pseudomonas aeruginosa*. *Mbio* **10**, doi:ARTN e02253-19 10.1128/mBio.02253-19
(2019).

68 Heidrich, N. *et al.* The primary transcriptome of *Neisseria meningitidis* and its interaction
20 with the RNA chaperone Hfq. *Nucleic Acids Res* **45**, 6147-6167, doi:10.1093/nar/gkx168
(2017).

69 Thomason, M. K. *et al.* Global Transcriptional Start Site Mapping Using Differential RNA
Sequencing Reveals Novel Antisense RNAs in *Escherichia coli*. *Journal of Bacteriology* **197**, 18-
28, doi:10.1128/Jb.02096-14 (2015).

25 70 Ryan, D., Jenniches, L., Reichardt, S., Barquist, L. & Westermann, A. J. A high-resolution
transcriptome map identifies small RNA regulation of metabolism in the gut microbe
Bacteroides thetaiotaomicron. *Nat Commun* **11**, 3557, doi:10.1038/s41467-020-17348-5
(2020).

71 Kemal Avican, J. A., Matteo Togninalli, Jing Tang, Karsten M. Borgwardt, Mikael Rhen, Maria
30 Fällman. RNA Atlas of Human Bacterial Pathogens Uncovers Stress Dynamics Linked to
Infection. *bioRxiv*, doi:<https://doi.org/10.1101/2020.06.29.177147> (2020).

72 Cossart, P., Boquet, P., Normark, S. & Rappuoli, R. Cellular microbiology emerging. *Science*
271, 315-316, doi:10.1126/science.271.5247.315 (1996).

73 Raynaud, S., Le Pabic, H. & Felden, B. Selective recovery of RNAs from bacterial pathogens
35 after their internalization by human host cells. *Methods*, doi:10.1016/j.ymeth.2018.04.035
(2018).

74 Hinton, J. C., Hautefort, I., Eriksson, S., Thompson, A. & Rhen, M. Benefits and pitfalls of using
microarrays to monitor bacterial gene expression during infection. *Curr Opin Microbiol* **7**,
277-282, doi:10.1016/j.mib.2004.04.009 S1369527404000475 [pii] (2004).

40 75 Szafranska, A. K. *et al.* High-resolution transcriptomic analysis of the adaptive response of
Staphylococcus aureus during acute and chronic phases of osteomyelitis. *MBio* **5**,
doi:10.1128/mBio.01775-14 (2014).

76 Rossi, E., Falcone, M., Molin, S. & Johansen, H. K. High-resolution in situ transcriptomics of
Pseudomonas aeruginosa unveils genotype independent patho-phenotypes in cystic fibrosis
45 lungs. *Nat Commun* **9**, 3459, doi:10.1038/s41467-018-05944-5 (2018).

77 Mandlik, A. *et al.* RNA-Seq-based monitoring of infection-linked changes in *Vibrio cholerae*
gene expression. *Cell Host Microbe* **10**, 165-174, doi:S1931-3128(11)00224-1
10.1016/j.chom.2011.07.007.

78 Donaldson, G. P. *et al.* Spatially distinct physiology of *Bacteroides fragilis* within the proximal colon of gnotobiotic mice. *Nature microbiology* **5**, 746-756, doi:10.1038/s41564-020-0683-3 (2020).

79 Betin, V. *et al.* Hybridization-based capture of pathogen mRNA enables paired host-pathogen transcriptional analysis. *Sci Rep* **9**, 19244, doi:10.1038/s41598-019-55633-6 (2019).

5 80 Bashiardes, S., Zilberman-Schapira, G. & Elinav, E. Use of Metatranscriptomics in Microbiome Research. *Bioinformatics and biology insights* **10**, 19-25, doi:10.4137/BBI.S34610 (2016).

81 Reck, M. *et al.* Stool metatranscriptomics: A technical guideline for mRNA stabilisation and isolation. *BMC Genomics* **16**, 494, doi:10.1186/s12864-015-1694-y (2015).

10 82 Engel, P. & Moran, N. A. The gut microbiota of insects - diversity in structure and function. *FEMS Microbiol Rev* **37**, 699-735, doi:10.1111/1574-6976.12025 (2013).

83 Abu-Ali, G. S. *et al.* Metatranscriptome of human faecal microbial communities in a cohort of adult men. *Nature microbiology* **3**, 356-366, doi:10.1038/s41564-017-0084-4 (2018).

84 Yan, Y. *et al.* Structure of the Mucosal and Stool Microbiome in Lynch Syndrome. *Cell Host & Microbe* **27**, 585+, doi:10.1016/j.chom.2020.03.005 (2020).

15 85 Mehta, R. S. *et al.* Stability of the human faecal microbiome in a cohort of adult men. *Nature microbiology* **3**, 347-355, doi:10.1038/s41564-017-0096-0 (2018).

86 Franzosa, E. A. *et al.* Species-level functional profiling of metagenomes and metatranscriptomes. *Nat Methods* **15**, 962-968, doi:10.1038/s41592-018-0176-y (2018).

20 87 Menzel, P., Ng, K. L. & Krogh, A. Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat Commun* **7**, 11257, doi:10.1038/ncomms11257 (2016).

88 Wood, D. E. & Salzberg, S. L. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol* **15**, R46, doi:10.1186/gb-2014-15-3-r46 (2014).

89 Forouzan, E., Shariati, P., Mousavi Maleki, M. S., Karkhane, A. A. & Yakhchali, B. Practical evaluation of 11 de novo assemblers in metagenome assembly. *J Microbiol Methods* **151**, 99-105, doi:10.1016/j.mimet.2018.06.007 (2018).

25 90 Zhang, Y. *et al.* Compositional and Functional Differences in the Human Gut Microbiome Correlate with Clinical Outcome following Infection with Wild-Type *Salmonella enterica* Serovar Typhi. *mBio* **9**, doi:10.1128/mBio.00686-18 (2018).

30 91 Westermann, A. J. *et al.* The Major RNA-Binding Protein ProQ Impacts Virulence Gene Expression in *Salmonella enterica* Serovar Typhimurium. *MBio* **10**, doi:10.1128/mBio.02504-18 (2019).

92 Stapels, D. A. C. *et al.* *Salmonella* persists undermine host immune defenses during antibiotic treatment. *Science* **362**, 1156-1160, doi:10.1126/science.aat7148 (2018).

35 93 Aprianto, R., Slager, J., Holsappel, S. & Veening, J. W. Time-resolved dual RNA-seq reveals extensive rewiring of lung epithelial and pneumococcal transcriptomes during early infection. *Genome Biol* **17**, 198, doi:10.1186/s13059-016-1054-5 (2016).

94 Mavromatis, C. H. *et al.* The co-transcriptome of uropathogenic *Escherichia coli*-infected mouse macrophages reveals new insights into host-pathogen interactions. *Cell Microbiol* **17**, 730-746, doi:10.1111/cmi.12397 (2015).

40 95 Baddal, B. *et al.* Dual RNA-seq of Nontypeable *Haemophilus influenzae* and Host Cell Transcriptomes Reveals Novel Insights into Host-Pathogen Cross Talk. *MBio* **6**, e01765-01715, doi:10.1128/mBio.01765-15 (2015).

96 Humphrys, M. S. *et al.* Simultaneous transcriptional profiling of bacteria and their host cells. *PLoS One* **8**, e80597, doi:10.1371/journal.pone.0080597 (2013).

45 97 Zimmermann, M. *et al.* Integration of Metabolomics and Transcriptomics Reveals a Complex Diet of *Mycobacterium tuberculosis* during Early Macrophage Infection. *Msystems* **2**, doi:UNSP e00057-17 10.1128/mSystems.00057-17 (2017).

98 Rienksma, R. A. *et al.* Comprehensive insights into transcriptional adaptation of intracellular mycobacteria by microbe-enriched dual RNA sequencing. *Bmc Genomics* **16**, doi:ARTN 34 10.1186/s12864-014-1197-2 (2015).

50

99 Mika-Gospodorz, B. *et al.* Dual RNA-seq of *Orientia tsutsugamushi* informs on host-pathogen interactions for this neglected intracellular human pathogen. *Nat Commun* **11**, 3363, doi:10.1038/s41467-020-17094-8 (2020).

100 Schulte, L. N. *et al.* An Advanced Human Intestinal Coculture Model Reveals
5 Compartmentalized Host and Pathogen Strategies during *Salmonella* Infection. *mBio* **11**, doi:10.1128/mBio.03348-19 (2020).

101 Hannemann, S., Gao, B. & Galan, J. E. *Salmonella* modulation of host cell gene expression promotes its intracellular growth. *PLoS Pathog* **9**, e1003668, doi:10.1371/journal.ppat.1003668 (2013).

102 Thanert, R., Goldmann, O., Beineke, A. & Medina, E. Host-inherent variability influences the transcriptional response of *Staphylococcus aureus* during in vivo infection. *Nat Commun* **8**, 14268, doi:10.1038/ncomms14268 (2017).

103 Nuss, A. M. *et al.* Tissue dual RNA-seq allows fast discovery of infection-specific functions and riboregulators shaping host-pathogen transcriptomes. *Proc Natl Acad Sci U S A* **114**, E791-E800, doi:10.1073/pnas.1613405114 (2017).

104 Damron, F. H., Oglesby-Sherrouse, A. G., Wilks, A. & Barbier, M. Dual-seq transcriptomics reveals the battle for iron during *Pseudomonas aeruginosa* acute murine pneumonia. *Sci Rep* **6**, 39172, doi:10.1038/srep39172 (2016).

105 Minhas, V. *et al.* In vivo dual RNA-seq reveals that neutrophil recruitment underlies differential tissue tropism of *Streptococcus pneumoniae*. *Communications biology* **3**, 293, doi:10.1038/s42003-020-1018-x (2020).

106 Griesenauer, B. *et al.* Determination of an Interaction Network between an Extracellular Bacterial Pathogen and the Human Host. *mBio* **10**, doi:10.1128/mBio.01193-19 (2019).

107 Thanert, R. *et al.* Molecular profiling of tissue biopsies reveals unique signatures associated with streptococcal necrotizing soft tissue infections. *Nat Commun* **10**, 3846, doi:10.1038/s41467-019-11722-8 (2019).

108 Pisu, D., Huang, L., Grenier, J. K. & Russell, D. G. Dual RNA-Seq of *Mtb*-Infected Macrophages In Vivo Reveals Ontologically Distinct Host-Pathogen Interactions. *Cell Rep* **30**, 335-350 e334, doi:10.1016/j.celrep.2019.12.033 (2020).

109 Jew, B. *et al.* Accurate estimation of cell composition in bulk expression through robust integration of single-cell information. *Nat Commun* **11**, 1971, doi:10.1038/s41467-020-15816-6 (2020).

110 Zaitsev, K., Bambouskova, M., Swain, A. & Artyomov, M. N. Complete deconvolution of cellular mixtures based on linearity of transcriptional signatures. *Nat Commun* **10**, 2209, doi:10.1038/s41467-019-09990-5 (2019).

111 Qiu, J. *et al.* Mixed-species RNA-seq for elucidation of non-cell-autonomous control of gene transcription. *Nat Protoc* **13**, 2176-2199, doi:10.1038/s41596-018-0029-2 (2018).

112 Seelbinder, B. *et al.* Triple RNA-seq reveals synergy in a human virus-fungus co-infection model. *Cell Rep* (2020).

113 Le-Bury, G. & Niedergang, F. Defective Phagocytic Properties of HIV-Infected Macrophages: How Might They Be Implicated in the Development of Invasive *Salmonella* Typhimurium? *Front Immunol* **9**, doi:Artn 531 10.3389/Fimmu.2018.00531 (2018).

114 Gordon, M. A. *Salmonella* infections in immunocompromised adults. *The Journal of infection* **56**, 413-422, doi:10.1016/j.jinf.2008.03.012 (2008).

115 Chertow, D. S. & Memoli, M. J. Bacterial Coinfection in Influenza A Grand Rounds Review. *Jama-J Am Med Assoc* **309**, 275-282, doi:DOI 10.1001/jama.2012.194139 (2013).

116 Steben, M. & Duarte-Franco, E. Human papillomavirus infection: epidemiology and pathophysiology. *Gynecologic oncology* **107**, S2-5, doi:10.1016/j.ygyno.2007.07.067 (2007).

117 Perez-Losada, M., Castro-Nallar, E., Bendall, M. L., Freishtat, R. J. & Crandall, K. A. Dual Transcriptomic Profiling of Host and Microbiota during Health and Disease in Pediatric Asthma. *PLoS One* **10**, e0131819, doi:10.1371/journal.pone.0131819 (2015).

118 Ren, L. *et al.* Transcriptionally Active Lung Microbiome and Its Association with Bacterial
Biomass and Host Inflammatory Status. *Msystems* **3**, doi:10.1128/mSystems.00199-18
(2018).

119 Bordenstein, S. R. & Theis, K. R. Host Biology in Light of the Microbiome: Ten Principles of
5 Holobionts and Hologenomes. *PLoS Biol* **13**, e1002226, doi:10.1371/journal.pbio.1002226
(2015).

120 Birnbaum, K. D. Power in Numbers: Single-Cell RNA-Seq Strategies to Dissect Complex
Tissues. *Annu Rev Genet* **52**, 203-221, doi:10.1146/annurev-genet-120417-031247 (2018).

121 Hautefort, I., Proenca, M. J. & Hinton, J. C. Single-copy green fluorescent protein gene fusions
10 allow accurate measurement of Salmonella gene expression in vitro and during infection of
mammalian cells. *Appl Environ Microbiol* **69**, 7480-7491, doi:10.1128/aem.69.12.7480-
7491.2003 (2003).

122 Tsai, C. N. & Coombes, B. K. The Role of the Host in Driving Phenotypic Heterogeneity in
Salmonella. *Trends in microbiology* **27**, 508-523, doi:10.1016/j.tim.2019.01.004 (2019).

15 123 Ackermann, M. *et al.* Self-destructive cooperation mediated by phenotypic noise. *Nature*
454, 987-990, doi:10.1038/nature07067 (2008).

124 Rivera-Chavez, F. *et al.* Salmonella uses energy taxis to benefit from intestinal inflammation.
PLoS Pathog **9**, e1003267, doi:10.1371/journal.ppat.1003267 (2013).

125 Vieth, B., Parekh, S., Ziegenhain, C., Enard, W. & Hellmann, I. A systematic evaluation of
20 single cell RNA-seq analysis pipelines. *Nat Commun* **10**, 4667, doi:10.1038/s41467-019-
12266-7 (2019).

126 Attar, M. *et al.* A practical solution for preserving single cells for RNA sequencing. *Sci Rep* **8**,
2151, doi:10.1038/s41598-018-20372-7 (2018).

127 Avraham, R. *et al.* Pathogen Cell-to-Cell Variability Drives Heterogeneity in Host Immune
25 Responses. *Cell* **162**, 1309-1321, doi:10.1016/j.cell.2015.08.027 (2015).

128 Saliba, A. E. *et al.* Single-cell RNA-seq ties macrophage polarization to growth rate of
intracellular Salmonella. *Nature microbiology* **2**, 16206, doi:10.1038/nmicrobiol.2016.206
(2016).

129 Helaine, S. *et al.* Internalization of Salmonella by macrophages induces formation of
30 nonreplicating persisters. *Science* **343**, 204-208, doi:10.1126/science.1244705 (2014).

130 Claudi, B. *et al.* Phenotypic variation of salmonella in host tissues delays eradication by
antimicrobial chemotherapy. *Cell* **158**, 722-733, doi:10.1016/j.cell.2014.06.045 (2014).

131 Panagi, I. *et al.* Salmonella Effector SteE Converts the Mammalian Serine/Threonine Kinase
GSK3 into a Tyrosine Kinase to Direct Macrophage Polarization. *Cell Host & Microbe* **27**, 41-+,
35 doi:10.1016/j.chom.2019.11.002 (2020).

132 Ben-Moshe, N. B. *et al.* Predicting bacterial infection outcomes using single cell RNA-
sequencing analysis of human immune cells. *Nature Communications* **10**, doi:Artn 3266
10.1038/S41467-019-11257-Y (2019).

133 Haber, A. L. *et al.* A single-cell survey of the small intestinal epithelium. *Nature* **551**, 333-+,
40 doi:10.1038/nature24489 (2017).

134 Medaglia, C. *et al.* Spatial reconstruction of immune niches by combining photoactivatable
reporters and scRNA-seq. *Science* **358**, 1622-1626, doi:10.1126/science.aao4277 (2017).

135 Hagemann-Jensen, M., Abdullayev, I., Sandberg, R. & Faridani, O. R. Small-seq for single-cell
small-RNA sequencing. *Nat Protoc* **13**, 2407-2424, doi:10.1038/s41596-018-0049-y (2018).

45 136 Faridani, O. R. *et al.* Single-cell sequencing of the small-RNA transcriptome. *Nat Biotechnol*
34, 1264-1266, doi:10.1038/nbt.3701 (2016).

137 Xiao, Z. *et al.* Holo-Seq: single-cell sequencing of holo-transcriptome. *Genome Biol* **19**, 163,
doi:10.1186/s13059-018-1553-7 (2018).

138 Taniguchi, Y. *et al.* Quantifying E. coli proteome and transcriptome with single-molecule
50 sensitivity in single cells. *Science* **329**, 533-538, doi:10.1126/science.1188308 (2010).

139 Blattman, S. B., Jiang, W., Oikonomou, P. & Tavazoie, S. Prokaryotic single-cell RNA
sequencing by in situ combinatorial indexing. *Nature microbiology*, doi:10.1038/s41564-020-
0729-6 (2020).

140 Imdahl, F., Vafadarnejad, E., Homberger, C., Saliba, A.-E. & Vogel, J. Single-cell RNA-
5 sequencing reports growth-condition-specific global transcriptomes of individual bacteria
Nature Microbiology (2020).

141 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, doi:10.1038/nmeth.4145 (2017).

142 Rosenberg, A. B. *et al.* Single-cell profiling of the developing mouse brain and spinal cord with
10 split-pool barcoding. *Science* **360**, 176-182, doi:10.1126/science.aam8999 (2018).

143 Bartholomaeus, A. *et al.* Bacteria differently regulate mRNA abundance to specifically respond
to various stresses. *Philos T R Soc A* **374**, doi:Artn 20150069 10.1098/Rsta.2015.0069 (2016).

144 Haque, A., Engel, J., Teichmann, S. A. & Lonnberg, T. A practical guide to single-cell RNA-
sequencing for biomedical research and clinical applications. *Genome medicine* **9**, 75,
15 doi:10.1186/s13073-017-0467-4 (2017).

145 Luecken, M. D. & Theis, F. J. Current best practices in single-cell RNA-seq analysis: a tutorial.
Mol Syst Biol **15**, e8746, doi:10.15252/msb.20188746 (2019).

146 Fisher, R. A., Gollan, B. & Helaine, S. Persistent bacterial infections and persister cells. *Nat*
Rev Microbiol **15**, 453-464, doi:10.1038/nrmicro.2017.42 (2017).

20 147 Gomez, J. A. *et al.* The NeST long ncRNA controls microbial susceptibility and epigenetic
activation of the interferon-gamma locus. *Cell* **152**, 743-754, doi:10.1016/j.cell.2013.01.015
(2013).

148 Imamura, K. *et al.* Diminished nuclear RNA decay upon Salmonella infection upregulates
antibacterial noncoding RNAs. *EMBO J* **37**, doi:10.15252/emboj.201797723 (2018).

25 149 Schulte, L. N., Eulalio, A., Mollenkopf, H. J., Reinhardt, R. & Vogel, J. Analysis of the host
microRNA response to Salmonella uncovers the control of major cytokines by the let-7
family. *EMBO J* **30**, 1977-1989, doi:emboj201194 10.1038/emboj.2011.94 (2011).

150 Schulte, L. N., Westermann, A. J. & Vogel, J. Differential activation and functional
specialization of miR-146 and miR-155 in innate immune sensing. *Nucleic Acids Research* **41**,
30 542-553, doi:Doi 10.1093/Nar/Gks1030 (2013).

151 Maudet *et al.* Functional high-throughput screening identifies the miR-15 microRNA family as
cellular restriction factors for Salmonella infection. *Nat Commun* (2014).

152 Rodriguez, A. *et al.* Requirement of bic/microRNA-155 for normal immune function. *Science*
316, 608-611, doi:10.1126/science.1139253 (2007).

35 153 Aguilar, C. *et al.* Functional screenings reveal different requirements for host microRNAs in
Salmonella and Shigella infection. *Nature microbiology* **5**, 192-205, doi:10.1038/s41564-019-
0614-3 (2020).

154 Gong, H. *et al.* A Salmonella small non-coding RNA facilitates bacterial invasion and
intracellular replication by modulating the expression of virulence factors. *PLoS Pathog* **7**,
40 e1002120, doi:10.1371/journal.ppat.1002120 (2011).

155 Lee, E. J. & Groisman, E. A. An antisense RNA that governs the expression kinetics of a
multifunctional virulence gene. *Molecular Microbiology* **76**, 1020-1033, doi:DOI
10.1111/j.1365-2958.2010.07161.x (2010).

156 Padalon-Brauch, G. *et al.* Small RNAs encoded within genetic islands of Salmonella
45 typhimurium show host-induced expression and role in virulence. *Nucleic Acids Res* **36**, 1913-
1927, doi:gkn050 10.1093/nar/gkn050 (2008).

157 Ellis, M. J. *et al.* Silent but deadly: IS200 promotes pathogenicity in Salmonella Typhimurium.
RNA Biol **15**, 176-181, doi:10.1080/15476286.2017.1403001 (2018).

158 Hofer, K. & Jaschke, A. Epitranscriptomics: RNA Modifications in Bacteria and Archaea.
50 *Microbiology spectrum* **6**, doi:10.1128/microbiolspec.RWR-0015-2017 (2018).

- 159 Frye, M., Jaffrey, S. R., Pan, T., Rechavi, G. & Suzuki, T. RNA modifications: what have we
learned and where are we headed? *Nat Rev Genet* **17**, 365-372, doi:10.1038/nrg.2016.47
(2016).
- 160 Rimbach, K., Kaiser, S., Helm, M., Dalpke, A. H. & Eigenbrod, T. 2'-O-Methylation within
5 Bacterial RNA Acts as Suppressor of TLR7/TLR8 Activation in Human Innate Immune Cells.
Journal of innate immunity **7**, 482-493, doi:10.1159/000375460 (2015).
- 161 Shippy, D. C. & Fadl, A. A. tRNA modification enzymes GidA and MnmE: potential role in
virulence of bacterial pathogens. *International journal of molecular sciences* **15**, 18267-
18280, doi:10.3390/ijms151018267 (2014).
- 10 162 Helm, M. & Motorin, Y. Detecting RNA modifications in the epitranscriptome: predict and
validate. *Nat Rev Genet* **18**, 275-291, doi:10.1038/nrg.2016.169 (2017).
- 163 Marbaniang, C. N. & Vogel, J. Emerging roles of RNA modifications in bacteria. *Curr Opin
Microbiol* **30**, 50-57, doi:10.1016/j.mib.2016.01.001 (2016).
- 164 Sonesson, C. *et al.* A comprehensive examination of Nanopore native RNA sequencing for
15 characterization of complex transcriptomes. *Nat Commun* **10**, 3359, doi:10.1038/s41467-
019-11272-z (2019).
- 165 Garalde, D. R. *et al.* Highly parallel direct RNA sequencing on an array of nanopores. *Nat
Methods* **15**, 201-206, doi:10.1038/nmeth.4577 (2018).
- 166 Workman, R. E. *et al.* Nanopore native RNA sequencing of a human poly(A) transcriptome.
20 *Nat Methods* **16**, 1297-1305, doi:10.1038/s41592-019-0617-2 (2019).
- 167 Tsatsaronis, J. A., Franch-Arroyo, S., Resch, U. & Charpentier, E. Extracellular Vesicle RNA: A
Universal Mediator of Microbial Communication? *Trends in microbiology* **26**, 401-410,
doi:10.1016/j.tim.2018.02.009 (2018).
- 168 Brody, H. Extracellular RNA. *Nature* (2020).
- 25 169 Das, S. *et al.* The Extracellular RNA Communication Consortium: Establishing Foundational
Knowledge and Technologies for Extracellular RNA Research. *Cell* **177**, 231-242,
doi:10.1016/j.cell.2019.03.023 (2019).
- 170 Lecrivain, A. L. & Beckmann, B. M. Bacterial RNA in extracellular vesicles: A new regulator of
host-pathogen interactions? *Biochimica et biophysica acta. Gene regulatory mechanisms*
30 **1863**, 194519, doi:10.1016/j.bbagr.2020.194519 (2020).
- 171 Baptista, M. A. P. & Dolken, L. RNA dynamics revealed by metabolic RNA labeling and
biochemical nucleoside conversions. *Nat Methods* **15**, 171-172, doi:10.1038/nmeth.4608
(2018).
- 172 Erhard, F. *et al.* scSLAM-seq reveals core features of transcription dynamics in single cells.
35 *Nature* **571**, 419-423, doi:10.1038/s41586-019-1369-y (2019).
- 173 Crosetto, N., Bienko, M. & van Oudenaarden, A. Spatially resolved transcriptomics and
beyond. *Nat Rev Genet* **16**, 57-66, doi:10.1038/nrg3832 (2015).
- 174 Xia, C., Fan, J., Emanuel, G., Hao, J. & Zhuang, X. Spatial transcriptome profiling by MERFISH
reveals subcellular RNA compartmentalization and cell cycle-dependent gene expression.
40 *Proc Natl Acad Sci U S A* **116**, 19490-19499, doi:10.1073/pnas.1912459116 (2019).
- 175 Ke, R. *et al.* In situ sequencing for RNA analysis in preserved tissue and cells. *Nat Methods* **10**,
857-860, doi:10.1038/nmeth.2563 (2013).
- 176 Wang, X. *et al.* Three-dimensional intact-tissue sequencing of single-cell transcriptional
states. *Science* **361**, doi:10.1126/science.aat5691 (2018).
- 45 177 Liao, J., Lu, X., Shao, X., Zhu, L. & Fan, X. Uncovering an Organ's Molecular Architecture at
Single-Cell Resolution by Spatially Resolved Transcriptomics. *Trends in biotechnology*,
doi:10.1016/j.tibtech.2020.05.006 (2020).
- 178 Wilbrey-Clark, A., Roberts, K. & Teichmann, S. A. Cell Atlas technologies and insights into
tissue architecture. *The Biochemical journal* **477**, 1427-1442, doi:10.1042/BCJ20190341
50 (2020).

179 Avital, G. *et al.* scDual-Seq: mapping the gene regulatory program of Salmonella infection by
host and pathogen single-cell RNA-sequencing. *Genome Biology* **18**, doi:Artn 200
10.1186/S13059-017-1340-X (2017).

180 Sano, T., Smith, C. L. & Cantor, C. R. Immuno-PCR: very sensitive antigen detection by means
5 of specific antibody-DNA conjugates. *Science* **258**, 120-122, doi:10.1126/science.1439758
(1992).

181 Baddal, B. Next-generation technologies for studying host-pathogen interactions: a focus on
dual transcriptomics, CRISPR/Cas9 screening and organs-on-chips. *Pathog Dis* **77**, doi:ARTN
ftz060 10.1093/femspd/ftz060 (2019).

10 182 Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying
mammalian transcriptomes by RNA-Seq. *Nat Methods* **5**, 621-628, doi:nmeth.1226
10.1038/nmeth.1226 (2008).

183 Sultan, M. *et al.* A global view of gene activity and alternative splicing by deep sequencing of
the human transcriptome. *Science* **321**, 956-960, doi:1160342 10.1126/science.1160342
15 (2008).

184 Croucher, N. J. *et al.* A simple method for directional transcriptome sequencing using Illumina
technology. *Nucleic Acids Res* **37**, e148, doi:gkp811 10.1093/nar/gkp811 (2009).

185 Guell, M. *et al.* Transcriptome complexity in a genome-reduced bacterium. *Science* **326**,
1268-1271, doi:10.1126/science.1176951 (2009).

20 186 Passalacqua, K. D. *et al.* Structure and complexity of a bacterial transcriptome. *J Bacteriol*
191, 3203-3211, doi:JB.00122-09 10.1128/JB.00122-09 (2009).

187 Perkins, T. T. *et al.* A strand-specific RNA-Seq analysis of the transcriptome of the typhoid
bacillus *Salmonella typhi*. *PLoS Genet* **5**, e1000569, doi:10.1371/journal.pgen.1000569
(2009).

25 188 Yoder-Himes, D. R. *et al.* Mapping the *Burkholderia cenocepacia* niche response via high-
throughput sequencing. *Proc Natl Acad Sci U S A* **106**, 3976-3981, doi:0813403106
10.1073/pnas.0813403106 (2009).

189 Wurtzel, O. *et al.* A single-base resolution map of an archaeal transcriptome. *Genome Res* **20**,
133-141, doi:gr.100396.109 10.1101/gr.100396.109 (2010).

30 190 Mandlik, A. *et al.* RNA-Seq-based monitoring of infection-linked changes in *Vibrio cholerae*
gene expression. *Cell Host Microbe* **10**, 165-174, doi:S1931-3128(11)00224-1
10.1016/j.chom.2011.07.007 (2011).

191 Srikumar, S. *et al.* RNA-seq Brings New Insights to the Intra-Macrophage Transcriptome of
Salmonella Typhimurium. *PLoS Pathog* **11**, e1005262, doi:10.1371/journal.ppat.1005262
35 (2015).

192 Sittka, A. *et al.* Deep sequencing analysis of small noncoding RNA and mRNA targets of the
global post-transcriptional regulator, Hfq. *PLoS Genet* **4**, e1000163, doi:10.1371/journal.pgen.1000163 (2008).

193 Galvez, E. J. C. *et al.* Distinct Polysaccharide Utilization Determines Interspecies Competition
40 between Intestinal *Prevotella* spp. *Cell Host Microbe*, doi:10.1016/j.chom.2020.09.012
(2020).

194 Wilson, D. J. Insights from genomics into bacterial pathogen populations. *PLoS Pathog* **8**,
e1002874, doi:10.1371/journal.ppat.1002874 (2012).

195 Sheppard, S. K., Guttman, D. S. & Fitzgerald, R. Population genomics of bacterial host
45 adaptation. *Nature Reviews Genetics* **19**, 549-565, doi:10.1038/s41576-018-0032-z (2018).

- 196 Moustafa, A. M., Lal, A. & Planet, P. J. Comparative genomics in infectious disease. *Curr Opin Microbiol* **53**, 61-70, doi:10.1016/j.mib.2020.02.009 (2020).
- 197 Kellam, P. & Weiss, R. A. Infectogenomics: Insights from the host genome into infectious diseases. *Cell* **124**, 695-697, doi:10.1016/j.cell.2006.02.003 (2006).
- 5 198 Khor, C. C. & Hibberd, M. L. Host-pathogen interactions revealed by human genome-wide surveys. *Trends Genet* **28**, 233-243, doi:10.1016/j.tig.2012.02.001 (2012).
- 199 de Monerri, N. C. S. & Kim, K. Pathogens Hijack the Epigenome A New Twist on Host-Pathogen Interactions. *American Journal of Pathology* **184**, 897-911, doi:10.1016/j.ajpath.2013.12.022 (2014).
- 10 200 Gomez-Diaz, E., Jorda, M., Peinado, M. A. & Rivero, A. Epigenetics of Host-Pathogen Interactions: The Road Ahead and the Road Behind. *Plos Pathogens* **8**, doi:ARTN e1003007 10.1371/journal.ppat.1003007 (2012).
- 201 Ingolia, N. T., Ghaemmaghami, S., Newman, J. R. & Weissman, J. S. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* **324**, 218-223, doi:10.1126/science.1168978 (2009).
- 15 202 Fremin, B. J., Sberro, H. & Bhatt, A. S. MetaRibo-Seq measures translation in microbiomes. *Nat Commun* **11**, 3268, doi:10.1038/s41467-020-17081-z (2020).
- 203 Stern-Ginossar, N. Decoding viral infection by ribosome profiling. *J Virol* **89**, 6164-6166, doi:10.1128/JVI.02528-14 (2015).
- 20 204 Holmes, M. J., Shah, P., Wek, R. C. & Sullivan, W. J., Jr. Simultaneous Ribosome Profiling of Human Host Cells Infected with *Toxoplasma gondii*. *mSphere* **4**, doi:10.1128/mSphere.00292-19 (2019).
- 205 Leary, D. H., Hervey, W. J. t., Deschamps, J. R., Kusterbeck, A. W. & Vora, G. J. Which metaproteome? The impact of protein extraction bias on metaproteomic analyses. *Molecular and cellular probes* **27**, 193-199, doi:10.1016/j.mcp.2013.06.003 (2013).
- 25 206 Wilmes, P., Heintz-Buschart, A. & Bond, P. L. A decade of metaproteomics: where we stand and what the future holds. *Proteomics* **15**, 3409-3417, doi:10.1002/pmic.201500183 (2015).
- 207 Wang, Z., Han, Q. Q., Zhou, M. T., Chen, X. & Guo, L. Protein turnover analysis in *Salmonella* Typhimurium during infection by dynamic SILAC, Topograph, and quantitative proteomics. *Journal of basic microbiology* **56**, 801-811, doi:10.1002/jobm.201500315 (2016).
- 30 208 Kaloyanova, D., Vogels, M., van Balkom, B. W. & Helms, J. B. Quantitative proteomic identification of host factors involved in the *Salmonella* typhimurium infection cycle. *Methods Mol Biol* **1225**, 29-45, doi:10.1007/978-1-4939-1625-2_2 (2015).
- 209 Selkig, J. *et al.* Spatiotemporal proteomics uncovers cathepsin-dependent macrophage cell death during *Salmonella* infection. *Nature microbiology*, doi:10.1038/s41564-020-0736-7 (2020).
- 35 210 Auweter, S. D. *et al.* Quantitative mass spectrometry catalogues *Salmonella* pathogenicity island-2 effectors and identifies their cognate host binding partners. *J Biol Chem* **286**, 24023-24035, doi:10.1074/jbc.M111.224600 (2011).
- 40 211 Jean Beltran, P. M., Federspiel, J. D., Sheng, X. & Cristea, I. M. Proteomics and integrative omic approaches for understanding host-pathogen interactions and infectious diseases. *Mol Syst Biol* **13**, 922, doi:10.15252/msb.20167062 (2017).
- 212 Byron, S. A., Van Keuren-Jensen, K. R., Engelthaler, D. M., Carpten, J. D. & Craig, D. W. Translating RNA sequencing into clinical diagnostics: opportunities and challenges. *Nat Rev Genet* **17**, 257-271, doi:10.1038/nrg.2016.10 (2016).
- 45 213 Seger, C. & Salzmann, L. After another decade: LC-MS/MS became routine in clinical diagnostics. *Clinical biochemistry*, doi:10.1016/j.clinbiochem.2020.03.004 (2020).
- 214 Hatzios, S. K. *et al.* Chemoproteomic profiling of host and pathogen enzymes active in cholera. *Nat Chem Biol* **12**, 268-274, doi:10.1038/nchembio.2025 (2016).

215 Newsom, S. N. & McCall, L. I. Metabolomics: Eavesdropping on silent conversations between
 hosts and their unwelcome guests. *Plos Pathogens* **14**, doi:ARTN e1006926
 10.1371/journal.ppat.1006926 (2018).

216 Kentner, D. *et al.* Shigella reroutes host cell central metabolism to obtain high-flux nutrient
 5 supply for vigorous intracellular growth. *Proc Natl Acad Sci U S A* **111**, 9929-9934,
 doi:10.1073/pnas.1406694111 (2014).

217 Beste, D. J. *et al.* 13C-flux spectral analysis of host-pathogen metabolism reveals a mixed diet
 for intracellular Mycobacterium tuberculosis. *Chemistry & biology* **20**, 1012-1021,
 doi:10.1016/j.chembiol.2013.06.012 (2013).

10 218 Garg, N. *et al.* Three-Dimensional Microbiome and Metabolome Cartography of a Diseased
 Human Lung. *Cell Host Microbe* **22**, 705-716 e704, doi:10.1016/j.chom.2017.10.001 (2017).

219 Deatherage Kaiser, B. L. *et al.* A Multi-Omic View of Host-Pathogen-Commensal Interplay in
 Salmonella-Mediated Intestinal Infection. *PLoS One* **8**, e67155,
 doi:10.1371/journal.pone.0067155 (2013).

15 220 Lloyd-Price, J. *et al.* Multi-omics of the gut microbial ecosystem in inflammatory bowel
 diseases. *Nature* **569**, 655-662, doi:10.1038/s41586-019-1237-9 (2019).

221 Bolyen, E. *et al.* Reproducible, interactive, scalable and extensible microbiome data science
 using QIIME 2. *Nat Biotechnol* **37**, 852-857, doi:10.1038/s41587-019-0209-9 (2019).

222 Shishkin, A. A. *et al.* Simultaneous generation of many RNA-seq libraries in a single reaction.
 20 *Nat Methods* **12**, 323-325, doi:10.1038/nmeth.3313 (2015).

223 Wade, J. T. Where to begin? Mapping transcription start sites genome-wide in Escherichia
 coli. *J Bacteriol* **197**, 4-6, doi:10.1128/JB.02410-14 (2015).

224 Adiconis, X. *et al.* Comprehensive comparative analysis of 5'-end RNA-sequencing methods.
Nat Methods **15**, 505-511, doi:10.1038/s41592-018-0014-2 (2018).

25 225 Tian, B. & Manley, J. L. Alternative cleavage and polyadenylation: the long and short of it.
Trends Biochem Sci **38**, 312-320, doi:10.1016/j.tibs.2013.03.005 (2013).

226 Ju, X., Li, D. & Liu, S. Full-length RNA profiling reveals pervasive bidirectional transcription
 terminators in bacteria. *Nature microbiology* **4**, 1907-1918, doi:10.1038/s41564-019-0500-z
 (2019).

30 227 Felix Grünberger, R. K., Michael Jüttner, Martin Fenk, Andreas Borst, Robert Reichelt,
 Winfried Hausner, Jörg Soppa, Sébastien Ferreira-Cerca, Dina Grohmann. Exploring
 prokaryotic transcription, operon structures, rRNA maturation and modifications using
 Nanopore-based native RNA sequencing. *bioRxiv*,
 doi:<https://doi.org/10.1101/2019.12.18.880849> (2019).

35 228 Byrne, A., Cole, C., Volden, R. & Vollmers, C. Realizing the potential of full-length
 transcriptome sequencing. *Philos Trans R Soc Lond B Biol Sci* **374**, 20190097,
 doi:10.1098/rstb.2019.0097 (2019).

229 Cain, A. K. *et al.* A decade of advances in transposon-insertion sequencing. *Nat Rev Genet*,
 doi:10.1038/s41576-020-0244-x (2020).

40 230 Gawronski, J. D., Wong, S. M., Giannoukos, G., Ward, D. V. & Akerley, B. J. Tracking insertion
 mutants within libraries by deep sequencing and a genome-wide screen for Haemophilus
 genes required in the lung. *Proc Natl Acad Sci U S A* **106**, 16422-16427,
 doi:10.1073/pnas.0906627106 (2009).

231 Gonyar, L. A. *et al.* In Vivo Gene Essentiality and Metabolism in Bordetella pertussis. *mSphere*
 45 **4**, doi:10.1128/mSphere.00694-18 (2019).

232 Sternon, J. F. *et al.* Transposon Sequencing of Brucella abortus Uncovers Essential Genes for
 Growth In Vitro and Inside Macrophages. *Infect Immun* **86**, doi:10.1128/IAI.00312-18 (2018).

233 Zhang, X. *et al.* RNA-seq and Tn-seq reveal fitness determinants of vancomycin-resistant
 Enterococcus faecium during growth in human serum. *BMC Genomics* **18**, 893,
 doi:10.1186/s12864-017-4299-9 (2017).

50

234 Fu, Y., Waldor, M. K. & Mekalanos, J. J. Tn-Seq analysis of *Vibrio cholerae* intestinal
colonization reveals a role for T6SS-mediated antibacterial activity in the host. *Cell Host*
Microbe **14**, 652-663, doi:10.1016/j.chom.2013.11.001 (2013).

235 Chaudhuri, R. R. *et al.* Comprehensive assignment of roles for *Salmonella typhimurium* genes
5 in intestinal colonization of food-producing animals. *PLoS Genet* **9**, e1003456,
doi:10.1371/journal.pgen.1003456 (2013).

236 Vohra, P. *et al.* Retrospective application of transposon-directed insertion-site sequencing to
investigate niche-specific virulence of *Salmonella Typhimurium* in cattle. *BMC Genomics* **20**,
20, doi:10.1186/s12864-018-5319-0 (2019).

10 237 Shames, S. R. *et al.* Multiple *Legionella pneumophila* effector virulence phenotypes revealed
through high-throughput analysis of targeted mutant libraries. *Proc Natl Acad Sci U S A* **114**,
E10446-E10454, doi:10.1073/pnas.1708553114 (2017).

238 Zhu, L. *et al.* Novel Genes Required for the Fitness of *Streptococcus pyogenes* in Human
Saliva. *mSphere* **2**, doi:10.1128/mSphereDirect.00460-17 (2017).

15 239 Subashchandrabose, S. *et al.* *Acinetobacter baumannii* Genes Required for Bacterial Survival
during Bloodstream Infection. *mSphere* **1**, doi:10.1128/mSphere.00013-15 (2016).

240 Wang, H. *et al.* Hypermutation-induced in vivo oxidative stress resistance enhances *Vibrio*
cholerae host adaptation. *PLoS Pathog* **14**, e1007413, doi:10.1371/journal.ppat.1007413
(2018).

20 241 Crabill, E., Schofield, W. B., Newton, H. J., Goodman, A. L. & Roy, C. R. Dot/Icm-Translocated
Proteins Important for Biogenesis of the *Coxiella burnetii*-Containing Vacuole Identified by
Screening of an Effector Mutant Sublibrary. *Infect Immun* **86**, doi:10.1128/IAI.00758-17
(2018).

242 Mann, B. *et al.* Control of virulence by small RNAs in *Streptococcus pneumoniae*. *PLoS Pathog*
25 **8**, e1002788, doi:10.1371/journal.ppat.1002788 (2012).

243 Capel, E. *et al.* Comprehensive Identification of Meningococcal Genes and Small Noncoding
RNAs Required for Host Cell Colonization. *mBio* **7**, doi:10.1128/mBio.01173-16 (2016).

244 Carrette, J. E. *et al.* Global gene disruption in human cells to assign genes to phenotypes by
deep sequencing. *Nat Biotechnol* **29**, 542-546, doi:10.1038/nbt.1857 (2011).

30 245 Friedrich, M. J. *et al.* Genome-wide transposon screening and quantitative insertion site
sequencing for cancer gene discovery in mice. *Nat Protoc* **12**, 289-309,
doi:10.1038/nprot.2016.164 (2017).

246 Franceschini, A. *et al.* Specific inhibition of diverse pathogens in human cells by synthetic
microRNA-like oligonucleotides inferred from RNAi screens. *P Natl Acad Sci USA* **111**, 4548-
4553, doi:10.1073/pnas.1402353111 (2014).

35 247 Strich, J. R. & Chertow, D. S. CRISPR-Cas Biology and Its Application to Infectious Diseases.
Journal of Clinical Microbiology **57**, doi:UNSP e01307-18 10.1128/JCM.01307-18 (2019).

248 Shalem, O. *et al.* Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells. *Science*
343, 84-87, doi:10.1126/science.1247005 (2014).

40 249 Pacheco, A. R. *et al.* CRISPR Screen Reveals that EHEC's T3SS and Shiga Toxin Rely on Shared
Host Factors for Infection. *mBio* **9**, doi:10.1128/mBio.01003-18 (2018).

250 Blondel, C. J. *et al.* CRISPR/Cas9 Screens Reveal Requirements for Host Cell Sulfation and
Fucosylation in Bacterial Type III Secretion System-Mediated Cytotoxicity. *Cell Host &*
Microbe **20**, 226-237, doi:10.1016/j.chom.2016.06.010 (2016).

45 251 Chang, S. J., Jin, S. C., Jiao, X. & Galan, J. E. Unique features in the intracellular transport of
typhoid toxin revealed by a genome-wide screen. *PLoS Pathog* **15**, e1007704,
doi:10.1371/journal.ppat.1007704 (2019).

252 Tromp, A. T. *et al.* Human CD45 is an F-component-specific receptor for the staphylococcal
toxin Panton-Valentine leukocidin. *Nature microbiology* **3**, 708-717, doi:10.1038/s41564-018-
0159-x (2018).

50

253 Tao, L. *et al.* Frizzled proteins are colonic epithelial receptors for C. difficile toxin B. *Nature*
538, 350-355, doi:10.1038/nature19799 (2016).

254 Vigouroux, A. & Bikard, D. CRISPR Tools To Control Gene Expression in Bacteria. *Microbiology
and molecular biology reviews : MMBR* **84**, doi:10.1128/MMBR.00077-19 (2020).

5 255 Qu, J. *et al.* Modulating Pathogenesis with Mobile-CRISPRi. *J Bacteriol* **201**,
doi:10.1128/JB.00304-19 (2019).

256 Liu, X. *et al.* Exploration of Bacterial Bottlenecks and Streptococcus pneumoniae
Pathogenesis by CRISPRi-Seq. *Cell Host Microbe*, doi:10.1016/j.chom.2020.10.001 (2020).

257 Storz, G., Wolf, Y. I. & Ramamurthi, K. S. Small Proteins Can No Longer Be Ignored. *Annual
10 Review of Biochemistry, Vol 83* **83**, 753-+, doi:10.1146/annurev-biochem-070611-102400
(2014).

258 Duval, M. & Cossart, P. Small bacterial and phagic proteins: an updated view on a rapidly
moving field. *Current Opinion in Microbiology* **39**, 81-88, doi:10.1016/j.mib.2017.09.010
(2017).

15 259 Bikard, D. *et al.* Exploiting CRISPR-Cas nucleases to produce sequence-specific antimicrobials.
Nat Biotechnol **32**, 1146-1150, doi:10.1038/nbt.3043 (2014).

260 Citorik, R. J., Mimee, M. & Lu, T. K. Sequence-specific antimicrobials using efficiently
delivered RNA-guided nucleases. *Nature Biotechnology* **32**, 1141-1145, doi:10.1038/nbt.3011
(2014).

20 261 Beisel, C. L., Gomia, A. A. & Barrangou, R. A CRISPR design for next-generation
antimicrobials. *Genome Biol* **15**, 516, doi:10.1186/s13059-014-0516-x (2014).