Development of novel experimental approaches to decipher host-pathogen interaction at the single-cell level

Entwicklung neuer experimenteller Ansätze zur Entschlüsselung von Wirt-Pathogen-Interaktion auf Einzelzellebene

Dissertation zur Erlangung des naturwissenschaftlichen Doktorgrades

(Dr. rer. nat.)

der Graduate School of Life Sciences,

Bayerische Julius-Maximilians-Universität Würzburg

Infektion und Immunität

Vorgelegt von

Fabian Dominik Imdahl

aus

Aachen

Würzburg, 2022

Eingereicht am: ……

Bürostempel

Mitglieder des Promotionskomitees:

Vorsitzender: Prof. Dr. Georg Gasteiger

1. Betreuer: Dr. Antoine-Emmanuel Saliba

2. Betreuer: Prof. Dr. Roy Gross

3. Betreuer: Jun. Prof. Dr. Alexander Westermann

Tag des Promotionskolloquiums: ……

Doktorurkunden ausgehändigt am: ……..

“Dismantling bacterial RNA on the production line” Illustration by Dr. Sandy R. Pernitzsch

Abstract:

COVID-19 has impressively shown how quickly an emerging pathogen can have a massive impact on our entire lives and show how infectious diseases spread regardless of national borders and economic stability. We find ourselves in a post-antibiotic era and have rested too long on the laurels of past research, so today more and more people are dying from infections with multi-resistant germs.

Infections are highly plastic and heterogeneous processes that are strongly dependent on the individual, whether on the host or pathogen side.

Improving our understanding of the pathogenicity of microorganisms and finding potential targets for a completely new class of drugs is a declared goal of current basic research. To tackle this challenge, single-cell RNA sequencing (scRNA-seq) is our most accurate tool.

In this thesis we implemented different state of the art scRNA-seq technologies to better understand infectious diseases. Furthermore, we developed a new method which is capable to resolve the transcriptome of a single bacterium. Applying a poly(A)-independent scRNA-seq protocol to three different, infection relevant growth conditions we can report the faithful detection of growth-dependent gene expression patterns in individual *Salmonella* Typhimurium and *Pseudomonas* aeruginosa bacteria. The data analysis shows that this method not only allows the differentiation of various culture conditions but can also capture transcripts across different RNA species.

Furthermore, using state of the art imaging and single-cell RNA sequencing technologies, we comprehensively characterized a human intestinal tissue model which in further course of the project was used as a *Salmonella enterica* serovar Typhimurium infection model. While most infection studies are conducted in mice, lacking a human intestinal physiology, the *in vitro* human tissue model allows us to directly infer *in vivo* pathogenesis. Combining immunofluorescent imaging, deep single-cell RNA sequencing and HCR-FISH, applied in time course experiments, allows an unseen resolution for studying heterogeneity and the dynamics of *Salmonella* infection which reveals details of pathogenicity contrary to the general scientific opinion.

Zusammenfassung:

COVID-19 hat eindrucksvoll gezeigt, wie schnell ein neu auftretender Erreger massive Auswirkungen auf unser aller Leben haben kann und wie sich Infektionskrankheiten unabhängig von Landesgrenzen und wirtschaftlicher Stabilität ausbreiten. Wir befinden uns in einer post-antibiotischen Ära und haben uns zu lange auf den Lorbeeren der vergangenen Forschung ausgeruht, so dass heute immer mehr Menschen an Infektionen mit multiresistenten Keimen sterben.

Infektionen sind sehr plastische und variable Prozesse, die stark vom Individuum abhängen, sei es auf Seiten des Wirts oder des Erregers. Die Pathogenität von Mikroorganismen besser zu verstehen und potenzielle Angriffspunkte für eine völlig neue Klasse von Arzneimitteln zu finden ist ein erklärtes Ziel der aktuellen Grundlagenforschung. Um diese Herausforderung zu meistern, ist die Einzelzell-RNA-Sequenzierung (scRNA-seq) unser präzisestes Werkzeug.

In dieser Arbeit haben wir verschiedene hochmoderne scRNA-seq-Technologien eingesetzt, um Infektionskrankheiten besser zu verstehen. Darüber hinaus haben wir eine neue Methode entwickelt, die in der Lage ist, das Transkriptom eines einzelnen Bakteriums aufzulösen. Durch die Anwendung eines poly(A)-unabhängigen scRNA-seq-Protokolls unter drei verschiedenen, infektionsrelevanten Wachstumsbedingungen konnten wir die wachstumsabhängigen Genexpressionsmuster in einzelnen *Salmonella* Typhimurium- und *Pseudomonas aeruginosa*-Bakterien zuverlässig nachweisen. Die Datenanalyse zeigt, dass diese Methode nicht nur die Differenzierung verschiedener Kulturbedingungen ermöglicht, sondern auch Transkripte über verschiedene RNA-Spezies hinweg erfassen kann.

Darüber hinaus haben wir unter Verwendung modernster Bildgebungs- und Einzelzell-RNA-Sequenzierungstechnologien ein menschliches Darmgewebemodell umfassend charakterisiert, das im weiteren Verlauf des Projekts als *Salmonella* Typhimurium-Infektionsmodell verwendet wurde. Während die meisten Infektionsstudien in Mäusen durchgeführt werden, denen die menschliche Darmphysiologie fehlt, ermöglicht uns das *in vitro* Modell des menschlichen Gewebes direkte Rückschlüsse auf die Pathogenese *in vivo*. Die Kombination aus immunfluoreszierender Bildgebung, deep single-cell RNA Sequenzierung und HCR-FISH, angewandt in Zeitverlaufsexperimenten, ermöglicht eine bisher ungesehene Auflösung zur Untersuchung von Heterogenität und Dynamik einer *Salmonella* Infektion, welche Details der Pathogenität entgegen der allgemeinen wissenschaftlichen Meinung offenbaren.

Acknowledgements:

First and foremost, I want to thank my mentor and PhD supervisor, Dr. Antoine-Emmanuel Saliba, for giving me the opportunity to work in his group and in an excellent scientific environment, from the very beginning of his lab. Thank you for teaching and guiding me over several years, and for always being supportive, encouraging and motivating.

Many thanks also to my dissertation committee, Prof. Dr. Roy Gross for the support and also the evaluation of my dissertation. Jun. Prof. Dr. Alexander Westermann for his valuable advice and Prof. Georg Gasteiger for chairing the doctoral committee.

I want to express my greatest gratitude to the SIGA group with all current and former members. Thank you: Ehsan Vafadarnejad, Tobias Krammer, Oliver Dietrich, Christophe Toussaint, Nina DiFabion und Alexander Leipold.

A deep gratitude also goes to my family for their continuous support throughout my career. Thanks to Anne.

Table of contents:

List of figures:

List of tables:

List of Abbreviations:

BA	Bioanalyzer
BC	Barcode
BONCAT	Bioorthogonal non-canonical amino acid tagging
bp	Base pairs
cDNA	Complementary-DNA
CDS	Coding sequence
CEL-seq	Cell expression by linear amplification and sequencing
DASH	Depletion of abundant sequences by hybridization
DMEM	Dulbecco's modified eagle medium
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTP	Deoxyribonucleotide triphosphate
DPBS	Dulbecco's phosphate buffered saline
ds	Double-stranded
DTT	Dithiothreitol
EID	Emerging infectious disease
ERCC	External RNA control consortium
EtOH	Ethanol
FACS	Fluorescence activated cell sorting
FISH	Fluorescence in situ hybridization
FITC	Fluorescein isothiocyanate
g	Relative centrifugal force
gDNA	Genomic DNA
GFP	Green fluorescent protein
GO	Gene ontology
HCR	Hybridization chain reaction
HIRI	Helmholtz institute for RNA-based infection research
hITM	Human Intestinal epithelial Tissue Model
IGR	Intergenic region
IVT	*In vitro* transcription

LB	Lennox broth
LncRNA	Long non-coding RNA
m.o.i.	Multiplicity of infection
MALBAC	Multiple annealing and looping based amplification cycles
MATQ-seq	Multiple annealing and dC-tailing based quantitative single-cell RNA-seq
mRNA	Messenger RNA
NaCl	Sodium chloride
nt	Nucleotide
OD	Optical density
p.i.	Post infection
PBST	PBS with 0.1% Tween
PC	Principle component
PCA	Principal component analysis
PE	Paired end
Petri-seq	Prokaryotic single-cell RNA sequencing
PFA	Paraformaldehyde
RACS	Raman-activated cell-sorting
REM	Raster electron microscopy
RI	RNase Inhibitor
RNA	Ribonucleic acid
RNA-seq	RNA sequencing
RNase	ribonuclease
ROS	Reactive oxygen species
rpm	Rounds per minute
rRNA	Ribosomal RNA
RT	Reverse Transcription
S. Typhimurium	*Salmonella* Typhimurium
Sc	Single cell
SCDE	Single-cell differential expression analysis
SCRB-seq	Single-cell RNA barcoding and sequencing
scRNA-seq	Single-cell RNA sequencing
Seq	Sequencing
SIFT	Single-cell isolation following time-laps imaging

SIGA	Single-cell analysis group
SIS	Small intestinal submucosa
Smart-seq	Switching mechanism at 5’ end or RNA template -seq
SPI	*Salmonella* pathogenicity island
sRNA	Small RNA
SSC	Saline sodium citrate
SSCT	SSC with 0.1% Tween
t-SNE	t-distributed stochastic neighbour embedding
UMAP	Uniform manifold approximation and projection
T3SS	Type III secretion system
TEM	Transmission electron microscopy
tmRNA	Transmitter RNA
tRNA	Transfer RNA
UMI	Unique molecular identifier
VSG	Variant surface glycoprotein
WHO	World health organization
YscSeq	Yeast single-cell RNA sequencing

1. Introduction:

1.1 Infectious disease and antimicrobial crisis

Hygiene measures and the introduction of antibiotics since the 1950s have created the illusion that infectious diseases are a phenomenon of the past. Since 2020, infectious disease research is given a special status as the COVID-19 pandemic is affecting our daily lives around the world and has so far claimed the lives of about 5 million people in two years and brought the advanced healthcare systems of the western world to their knees[1,2]. Yet, COVID-19 adds up to the multitude of other infectious diseases still affecting the world and the recurrent emergence of new pathogens. The World Health Organization (WHO) reported for over 17 million deaths connected with infections of bacteria, fungi or viruses within a year and warns of a global crisis[3]. While most deaths are due to lower respiratory tract infections, observation of disease statistics over the last decade shows a marked increase in mortality rates from diarrhoeal diseases[4]. The steadily growing population and locally undeveloped hygiene infrastructure makes the spread of infectious diseases an enormous health, but also economic problem, especially in developing countries[5,6]. In addition, lacking environmental protection at pharmaceutical industrial plants creates breeding sites for multi-resistant germs[7]. With advancing globalisation, however, this is no longer just a problem of developing countries; nosocomial multi-resistant pathogens have long since found their way into our hospitals.

Most emerging infectious diseases (EIDs) are caused by bacteria[8] which feature mechanisms, such as horizontal gene transfer, to incorporate virulence factors into their genomes that are potentially beneficial for increased fitness or spread frequency[9,10]. Especially drug resistant microbes have led to a massive increase of EIDs in the past 50 years[8]. Supposedly eradicated diseases experience a resurgence through the acquisition of drug resistances via mutation, conjugation, transformation or viral transduction[11]. Simultaneously, the number of new antibiotic drug approvals has decreased constantly. In addition an accelerating resistance can be observed, which means that new antibiotics are only on the market for a short time before the first resistant bacteria appear[12]. The massive overuse of antibiotics in food industries combined with the increasing number of medical prescriptions fuel this trend[13]. While the estimated costs of antibiotic development amounts to US$1.6 billion, the average revenue generated from the approved products sale is roughly $46 million a year[14]. Due to this trend, it

hardly makes sense form an economic point of view for pharmaceutical companies to maintain the costly research and development of new antibiotics[15]. Today, we already treat patients with reserve antibiotics which nevertheless often fail to combat multi-resistant pathogens. In this apparent deadlock of antibiotics research and development, it is more important than ever to find other approaches to give us a leading edge in the everlasting microbial arms-race.

With the advancing and more comprehensive understanding of RNA and simultaneous desperation in the field of antibiotics, the need for RNA therapeutics is emerging. In recent years, major progresses have been made in generation, purification and cellular delivery of RNA, allowing the development of RNA therapeutics for a broad array of applications[16]. This new generation of drugs unites several advantages, such as, simple manufacturing, cost effectiveness or possible individualization, and has the potential to revolutionize the standard of care for many diseases and personalized medicin[16].

1.2 *Salmonella* infection and pathogenicity

Being one of the most frequently isolated foodborne pathogens, accounting for 93.8 million illnesses and 155,000 deaths per year, *Salmonella* is a worldwide public health concern, representing a major economic burden through costs of surveillance, prevention and treatment of disease[5,6,17]. Worldwide the most common manifestation of *Salmonella* infection is gastroenteritis followed by bacteraemia and enteric fever, mainly occurring in developing countries[18]. The rod-shaped, Gram-negative and facultative anaerobe member of the family of Enterobacteriaceae is classified in over 2,500 serovars based on their flagellar and surface antigenic composition[19,20]. While 'typhoidal serovars' like *S. enterica* subsp. *enterica* serovar Typhi or *S. enterica* subsp. *enterica* serovar Paratyphi, cause systemic infection primarily in the human host, other serovars ('non-typhoidal serovars') such as *S. enterica* subsp. *enterica* serovar Typhimurium have a broader host range and most often cause self-limiting gastroenteritis[21–23]. Main source of *Salmonella* infection are eggs and dairy products but also food animals such as swine, poultry and cattle[24]. In the past decades, antibiotic-resistant foodborne pathogens are emerging, becoming a growing public health problem. As these resistant pathogens are more virulent, the mortality rate of infected patients increases[25]. Antimicrobial resistance of *Salmonella* is mainly due to the excessive use of antibiotics in animal breeding. This poses the risk of zoonosis through transmission of multidrug-resistant *Salmonella* from animals to humans, caused by food or water contaminated with animal faeces,

through direct contact or consumption of infected food animals[26]. As multi-drug resistant pathogens become an ever-increasing threat to our society, new methods of basic research are needed to find targets for tailor-made drug development.

After oral ingestion, the bacteria can survive the acidic environment of the gastric tract to access the intestinal epithelium. Here it triggers inflammatory changes leading ultimately to diarrhea[27]. The inflammatory reaction indirectly nourishes the pathogen by releasing compounds like tetrathionate, providing *Salmonella* a growth advantage over the intestinal microbiota[28,29]. Invading the mucosal barrier of the intestinal epithelium the bacterium either gets engulfed by phagocytes or triggers its own uptake by epithelial cells. However, the described standard invasion process occurs over microfold cells (M cells) overlying the lymphoid tissue at the Peyer's patches[30]. Furthermore, *Salmonella* can invade and survive in several cell types, including immune cells like macrophages. As neutrophils can provide a rapid clearing of an infection, the pathogens intracellular behaviour is geared towards avoidance of a neutrophilic immune response[31,32]. To do this, the bacteria must be able to sense and react to certain stress factors such as cationic antimicrobial peptides or the acidic environment of a phagocytic vacuole. The recognition and adaption to the host innate immune system is mediated by transcriptional activation of genes involved in cell surface remodelling and regulatory proteins including two-component systems like PhoP-PhoQ or OmpR-EnvZ [33,34]. Other virulence factors of *S*. Typhimurium are the two distinct type III secretion systems (T3SS). Encoded by *Salmonella* pathogenicity island 1 (SPI-1) and SPI-2, they provide the delivery of bacterial effectors directly into the host cells using their needle like structures. While SPI-1 T3SS is mainly an invasion apparatus which transports effector proteins across the plasma membrane mediating the bacterial uptake, the SPI-2 T3SS releases transport proteins contributing to intracellular survival and vascular movement within the *Salmonella* containing vacuole (SCV)[34,35].

Focussing on the T3SS mediated uptake a subset of effectors, namely SopB, SopE and SopE2 have been identified as major drivers that concert together with SipA and SipC the complex mechanism of membrane and actin rearrangement and signal-pathway activation[36,37]. SopB, an inositol phosphatase is a multifunctional effector involved in membrane ruffling, M-cell development, reduction of SCV acidity and inhibition of SCV-fusion with the lysosome[38–41]. Together with SopE and SopE2, and to indirectly rearrange the host cells actin structure, GTPases of the Rho family are stimulated, while SipA and SipC can directly bind actin[42]. In collaboration SipC and SopE can directly fusion exocytic vesicles with the plasma membrane

at the entry site, leading to pronounced membrane ruffling enabling macropinocytic uptake of the pathogen[43]. The multifactorial entry-mechanism is way more sophisticated than here described and thus is still a major research focus of *Salmonella* researchers. However, several pieces of the puzzle are still missing to unveil the whole complexity of the invasion pathway.

Internalized *Salmonella* survive and proliferate in two major niches, either enclosed by SCV or free-living within the hosts cytosol[44]. As mentioned before, effectors of the T3SS SPI2 are mediating an extensive remodelling of the SCV, leading to a migration of the vacuole from cell periphery towards the nucleus[45]. While SPI-1 induced effectors play a major role in the early stages of SCV biogenesis, it is downregulated once the pathogen is residing in the SCV. In contrast, SPI-2 T3SS effectors which are intravacuolar induced and responsible for progressive biogenesis of the vacuole, the onset of bacterial replication up to the formation of *Salmonella* induced filaments[46]. Timing-wise, the SCV remodelling happens in the first 60 min post infection (p.i.), followed by movement towards the Microtubule organizing center (MTOC) in hour 1-4 p.i.; proximate onset of intravacuolar bacterial replication and extensive SifA mediated tubulation is described at > 4h after infection[47]. However, *Salmonella* residing in the cytosolic compartment of epithelial cells show a robust growth surpassing replication in SCVs[48]. These Cytosolic bacteria therefore require several niche-specific transcriptional adaptions such as upregulation of genes involved in iron uptake and storage, manganese and sugar transport or fostering motility and adhesion, priming for massive proliferation and reinfection upon their release[44].

While most research was focused to elucidate the intricate relation of Salmonella the host and the role of bacterial effectors, the aspect of variability in host-pathogen interaction has been neglected.

1.3 Phenotypic heterogeneity of isogenic bacteria

Bacterial isogenic cell populations have been demonstrated to harbour a considerable phenotypic variation[49]. This heterogeneity can impact a wide variety of biological processes and arises presumably from stochasticity or noise in gene expression[50]. This diversity can be a great advantage for a bacterial population, particularly in a host and pathogen context. In the course of infection, pathogens are often exposed to dramatic environmental changes, so in order to survive in different host niches, bacteria have evolved diverse mechanisms to promote

phenotypic diversity[51]. Besides stochastic events and fluctuations which potentially lead to different transcriptional signatures[52], also the microenvironment of infection sites and related gradients of metabolite concentration, immune response activities or reactive oxygen species might have an influence on phenotypic variation[53]. Some sets of genes have been showed to be more susceptible for stochastic expression than others. Evolutionary conserved and housekeeping genes are less likely to change expression patterns, while genes in metabolic or stress response pathways are more frequently affected[52,54]. The terms *division of labour* and *bet hedging* have been assigned to the two main benefits of heterogeneity in pathogenic gene expression. While both behavioural patterns increase the overall fitness of a population, *division of labour* describes a functional division of pathogenic subpopulations that perform different tasks, such as defence against neutrophils or the activation of metabolic pathways[55]. Bet hedging, on the other hand, describes the formation of a subpopulation that has an indirect influence on a pathogen's fitness being able to survive sudden environmental changes acting as a safety anchor during the process of infection across several host barriers or exposed to immune cells[56,57]. Transcriptomic heterogeneity, however, is also demonstrated to allow pathogenic subpopulations to escape antibiotic therapy[58].

1.4 Persister cells and Infection

It hast long been recognised that numerous bacterial pathogens can exist in both, a replicative and a non-proliferating but metabolically active state[59,60]. While in 1993, Abshire et al. just observed a rapidly replicating and a stagnating population of *Salmonella* within macrophages, in the meantime so called persistent cells have been well described in *Escherichia coli*, *Pseudomonas aeruginosa*, *Mycobacterium tuberculosis*, and *Salmonella* Typhimurium and *Staphylococcus aureus*[61–63]. However, complicating terminology, persistence does not equal persistence, between different research areas the term has varying definitions. Thus, there are major differences between *persistence of an infection*, describing bacteria which remain viable after clearance of infection by the innate and/or adaptive immune response[64] versus *antibiotic persistence* defined as a clonal subpopulation of antibiotic susceptible bacteria that survive antibiotic treatment[65]. Since, as mentioned earlier, antibiotic-resistant bacteria pose a major threat for public health, here I elaborate on the latter ones.

Salmonellae residing in immune cells like macrophages or dendritic cells during a systemic infection are subjected to one of three fates when not treated with antibiotics. In the first case,

the immune cell kills the pathogen, controlling the infection. The second fate is massive pathogenic proliferation laying the foundation for an acute infection, which ultimately leads to the death of infected immune cells. Third, the pathogen enters to a non-growing persister state[66]. Being in a proliferative state Salmonellae are in general susceptible to killing by antibiotics, while non-growing bacteria which are antibiotic tolerant can survive antibiotic treatment and host immune response[67]. Various factors can be involved in the development of bacterial persistence in a host, such as immune evasion of the pathogen, immunosuppression of the host or ineffective clearance of an infection by antibiotics. The latter ones have been shown to lead to the emergence of antibiotic resistance which makes persisters a major public health concern[68–71]. Even without being genetically resistant, persister cells have the inherent ability to survive antibiotic treatment, which might be the key to persistent infections.

Studying persisters has seen an upsurge in new methods over the last few years. While first studied via selective antibiotics which were targeting insensitive cells[72], recently fluorescent reporters can be used to record metabolic activity and growth of persisters[73]. As an example of this, a fluorescence dilution strain which carried a plasmid reporting for bacterial proliferation by fading of a green fluorescent protein (GFP) was engineered to distinguish between actively growing or growth-arrested *Salmonella* within macrophages[60,74]. Those studies are mostly based on imaging and FACS analysis of infected cells, identifying persisters and revealing their growth dynamics[74]. Complemented with early micro-array-based transcriptomics, nowadays next generation sequencing in particular deep RNA sequencing became the method of choice, increasing sensitivity and dynamic range substantially[75,76]. Since persisters mostly represent a very rare subpopulation increased levels of precision and sensitivity are urgently needed.

A variety of environmental influences provide indications to trigger persister cell formation. Mainly cellular stresses such as extreme pH[60], nutrient starvation[77], diauxic shift[78] or DNA damage[79] are thought to be key for persistence. Besides that, environments which are incepted by bacteria like macrophages or biofilms may also be involved in the development of persisters[60,80,81].

Although the importance of persistent infections, especially nosocomial persistent infections, is well known, most scientific studies are limited to laboratory strains grown in nutrient rich culture media[65]. Studying persistent subpopulations in a human tissue would be a powerful approach to get deeper insights in formation and dynamics of persistent infections. Single-cell transcriptomics already revealed distinct subpopulations on the host side fostering non-growing bacterial states in macrophages[73,80]. Recent approaches resolving the transcriptome of individual

bacteria[82–84] could be the key to elucidating the underlying mechanism, and to identify targets to combat persistent populations.

1.5 RNA-sequencing

Ever since Robert Hook and Anton van Leeuvenhoek in the late 17th century presented the first published depiction of a microorganism, made possible by their inventive use and fabrication of simple microscopes, researchers strive for a higher resolution[85]. The ability to observe and characterise individual cells raises more and profound questions, which have been reflected in numerous studies over the past centuries[86]. Beyond the external observation, the functional analysis of cells comes to the fore. Fast forward, the classification of cells has been based on protein secretion under distinct functional conditions[87–89]. Analysing the proteome (entirety of all proteins of an organism) provides insights into the final functional product of gene expression. However, proteomics are limited to a pre-selected repertoire of molecules and show detection biases for low abundance or hydrophobic proteins, impeding a comprehensive phenotypic characterization of a cell[90]. RNA-sequencing (RNA-seq) sheds light to one level below and focusses on the transcriptome (set of all RNA transcripts in a cell). RNA-seq as well as single-cell RNA sequencing (scRNA-seq) has undergone a tremendous development over the last decade. While bulk RNA-seq approaches gained complexity and more refined methodologies, scRNA-seq mostly increased in throughput and sensitivity (Fig. 1, Fig 2)[91,92]. Focussing on the field of infection biology, Westermann and Vogel describe three major phases in the development of RNA-seq protocols for the analysis of host-pathogen interactions (Fig 1)[93]. While in the first phase host and pathogen were analysed separately, in the second phase the full power of the single-nucleotide resolution provided was exploited in order to characterize the interaction of two organisms e.g., host and pathogen, at the same time[94]. The third and present phase is related to the rise of scRNA-seq and its use in the area infection biology[95,96].

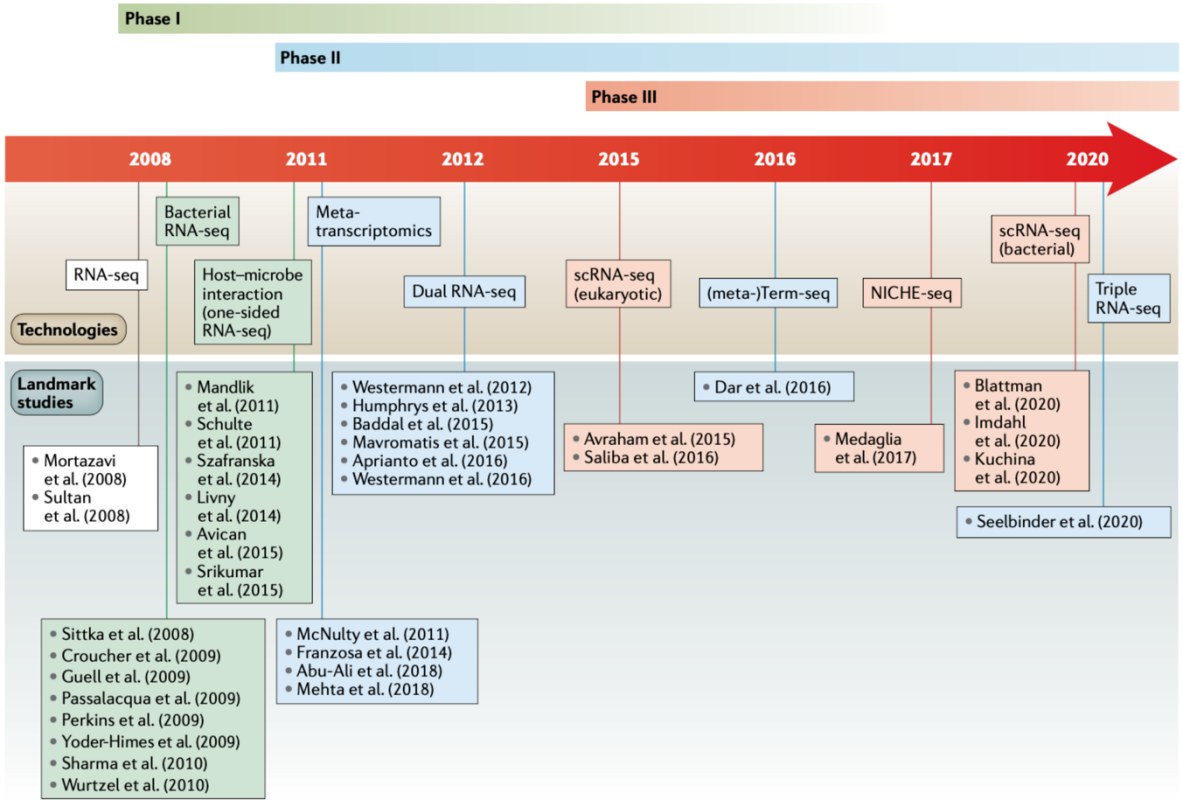

Fig. 1: Timeline of RNA-seq technologies over the past decades. The published methods, arranged in a timeline, can be subdivided in three different phases with increasing complexity. Phase I (green) addresses one-sided RNA-seq of either the host or the pathogen. Phase II (blue) studies address multiple transcriptomes sequenced at the same time. Phase III (red) corresponds to the emergence of scRNA-seq to the field of infection biology. Adapted from Westermann & Vogel 2021[91]

RNA-seq profiles transcript levels, revealing functional elements and molecular constituents of cells and tissues[97]. Measuring the average gene expression across populations is vital for comparative transcriptomics and quantitative expression signatures, for instance in disease studies[98]. However, RNA-seq can only map the broad spectrum of the analysed population-specific common transcriptome; subgroups or even individual outliers are not perceived by this global approach. Transcriptomic heterogeneity and stochastic variances of individual cells get masked by averaging the signal. Cells from the same type or even isogenic cells do not necessarily run the same transcriptional programme and may behave completely different according to their individual transcriptome[95]. In recent years scRNA-seq has become the gold standard to investigate the physiological state of an individual cell in health and disease. The field of infection biology in particular benefits from single-cell approaches, which offer enormous advantages and promise the discovery of unknown regulatory processes or shed light on the microbial dark matter[95,99].

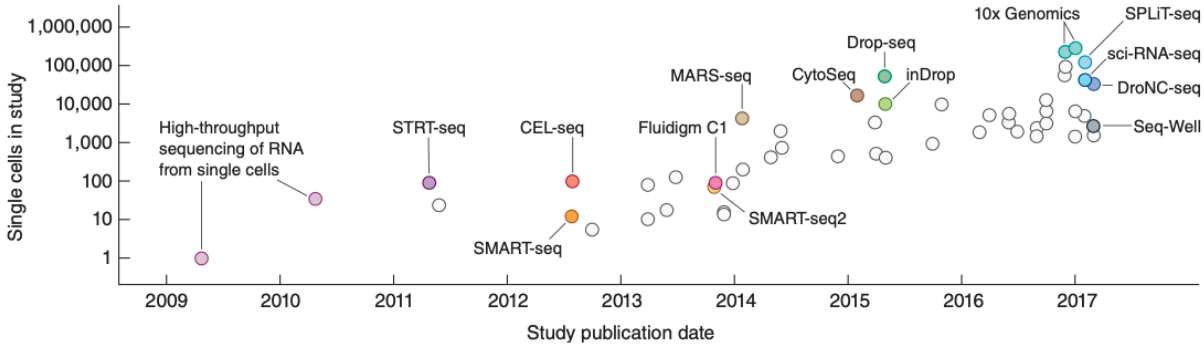

Fig. 2: Evolution of scRNA-seq throughput applied to eukaryotic cells over the past decade. Representative studies and respective cell numbers ordered by publication date. Colored dots represent selected key technologies. Reprint from Svenson et al. 2018[92].

Common to all scRNA-seq methods is the initial step of reverse transcription (RT), in which a cells RNA molecules are converted to more stable cDNA. Subsequently the cDNA gets amplified via polymerase chain reaction (PCR) or *in vitro* transcription (IVT). The ensuing steps of library preparation is followed by sequencing, allowing the analysis and quantification of a cell's transcripts. In pioneering studies the initial methods were applied to only a hand full of individual cells, since the technologies were laborious and quite expensive[100–103]. The introduction of commercial scRNA-seq kits, sophisticated multiplexing strategies and especially droplet based formats lead to an almost exponential increase of cell numbers in studies over the last few years (Fig. 2) [104–106]. Increasing throughput and sequencing cost reduction allow the analysis of complete tissues up to organs, enabling the construction of single-cell Atlases[107,108]. As throughput increases, so does the sensitivity of the technologies to keep pushing the frontiers of what is feasible. However, some aspects of infection research are still underrepresented, as there are some technical hurdles in integrating them into scRNA-seq. One of these hurdles is the deliberate avoidance of ribosomal RNA (rRNA) capture which most methods use to not employ too many sequencing-reads on the comparatively low-information-containing material. As rRNA accounts for the majority of all RNA species, bypassing it, usually by poly(A) capture, is very convenient as it reduces sequencing costs and increases throughput[93]. The downside is that all other non-polyadenylated RNA molecules, such as, lncRNAs, sRNAs or sn(o)RNAs, are excluded. Besides that, this widely used method make all protocols based on it inapplicable for prokaryotic transcripts. Very recent methods, aiming to elucidate regulatory RNA molecules or alternative splicing variants, address this challenge[109–111]. Also, the field of single-bacteria RNA sequencing has recently been kicked off with the publication of three novel methods[83,84,111]. RNA-seq and scRNA-seq have experienced an exponential growth of the past years, starting from a niche application to a ubiquitous gold

standard technology. Also, complexity has increased in terms of host-pathogen interaction. After dual RNA-seq meanwhile represents a standard procedure, more recent triple RNA-seq reveals the synergy in human-virus-fungus infection models[94,112]. Nevertheless, with the higher and further aspiration come new problems. As cell-numbers in studies are facing the 1 million marks, we go beyond the capacity of current sequencers, and the amount of data is not to be overlooked. Single-cell studies of the microbiome will revolutionize the field, however an even higher sensitivity and throughput as well as standardized methods will be needed. Future studies might resolve whole organs and organisms at the single cell level including space and time. Besides all tremendous scientific achievements scRNA-seq has brought already, we find ourselves still at the beginning of the rise of a multidimensional technology.

1.6 Advances and Challenges in single-cell RNA-seq of microbial communities

The intellectual content of the following paragraph is based on the previously published review 'Advances and challenges in single-cell RNA-seq of microbial communities' Imdahl & Saliba Current Opinion in Microbiology, 2020[113].

Over the past decade, RNA sequencing (RNA-seq) has developed to an indispensable tool for molecular biology[114] influencing microbial research in all aspects, from fundamental prokaryote physiology over RNA-based regulation processes to host-pathogen interaction and antibiotic susceptibility diagnostics[94,115,116]. Over the past few years scRNA-seq has developed into the gold standard for mapping the cellular state of eukaryotic cells by generating transcripts at a single nucleotide resolution for each individual cell[117], revolutionizing our view and knowledge of mammalian cellular identities and tissue organization, led by large international consortia such as the Human Cell Atlas or the LifeTime initiative[118–120]. However, in the lack of appropriate techniques, the majority of microbial studies still rely on bulk measurements averaging the transcriptomes of thousands to millions of bacteria.

Microbial cells can be found at almost all parts of the human body colonizing a great variety of niches including mucosal and skin environments[121]. Complex adaptions enable them to bypass our immune system and settle also in hostile environments within their host, making them at least as abundant as somatic cells[122]. In numbers it is estimated that 500-1000 bacterial species exist in the human body at any time[123], sharing their niches with archaeal and fungal species[121,124], influencing host cell physiology and affect the course of many diseases[121] from birth to adulthood[125–127]. Investigating bacterial ecosystems and taxonomic structures was mainly based on metagenomics uncovering the repertoire of available genes. Complementing this approach, metatranscriptomics unveils the transcribed genes at a certain time, enabling the detection and prediction of functional activities of individual species[127,128]. However, studying physiological states and developmental trajectories of microbes remains a fundamental challenge. The role of individuals in an isogenic microbial community was long time

underestimated. Via fluorescence reporters combined with bacterial engineering, time-laps microscopy, and microfluidics, it could be revealed how clonal identical cells residing in a similar environment are expressing differing genetic patterns and developing phenotypical heterogeneity.

Persister cells can shift to transient growth and adopt their metabolic activity to escape antibiotic treatments[50,52,58,129–131]. Deeper investigation of this and other medically relevant phenomena is calling for methods to profile the microbial transcriptome at a single cell level.

Understanding and characterizing the physiology and phenotypic heterogeneity of individual microbes in the environment, as well as in context of colonization and infection, independent from cell culture, was a long-term objective for microbiologists[52,53,132–134]. Single-cell genome sequencing of rare, uncultivable bacteria has changed our view on the tree of live and laid the foundation for high-throughput single-cell workflows and the respective analysis[135–138]. Building upon this, novel technologies capture the transcriptome of individual microbes allowing a deep insight into their physiological states and their phenotypic heterogeneity at the same time[83,84,139].

1.7 Challenges in profiling the RNA content of single bacteria

Performing microbial RNA-seq at the single-cell level requires overcoming some major technical hurdles. First, one needs capture the RNA, which is reasonably accessible in mammals or protozoa but well enclosed behind thick cell walls in bacteria and fungi. While eukaryotic cells can be easily lysed using detergent-based lysis buffers especially bacterial and fungal cells demand for a sophisticated lysis strategy comprising enzymatic digestion or mechanical rupture. The large differences between the species-specific cell walls prevent a universally valid lysis formula, which is reliably lysing the cells without harming the fragile RNA content.

Moreover, the properties of the RNA itself vary a lot between the microbes. Besides differing contents and copy numbers, bacterial mRNA lacks a polyadenylation, complicating the separation from ribosomal RNA (rRNA) which accounts for more than 90% of all RNA species. With an mRNA copy number lower than one in bacteria and yeast, the detection limits of common mammalian single-cell RNA-seq methods, which hardly capture transcripts with fewer than ten copies, are exceeded (Table 1)[140–143]. Combined with the high dropout rate, meaning the method's fails to reliably capture mRNA molecules, which ranges from 26% to

74%, it becomes obvious, why a simple adaption of protocols designed for mammalian cells is often not possible, and in this context why single-microbe RNA-seq is still in its infancy[144].

Table 1: Major physical characteristics of microbial and mammalian cells. *Imdahl & Saliba Current Opinion in Microbiology 2020*[113]

	Bacteria	**Yeast**	**Protozoa**	**Mammal**
Cell wall	Yes	Yes	No	No
Lysis	Enzymatic	Enzymatic	Detergent	Detergent
RNA content	~ 10-100 fg	~ 1 pg	1-10 pg	~ 10 pg
mRNA polyadenylation	No	Yes	Yes	Yes
Size	~ 1 µm	2-5 µm	1-20 µm	10-30 µm
Average mRNA copies/cell/gene	0.4	0.8	1-10	>10

1.8 Manipulation and recording of heterogenous microbial phenotypes

Performing single-cell analysis of microbes classically starts with a potential phenotypical characterization and isolation of individual cells. Fluorescent activated cell sorting (FACS) is probably the most convenient method to isolate microbes, convincing through fast, robust, and high-throughput processing, and its ability to be interfaced with PCR-plates for one-cell per well applications. FACS also enables a preselection of cell subsets based on fluorescence signals. Exemplary for this are genetically engineered *Salmonella* Typhimurium, intracellular bacteria, which express fluorescence reporters based on replication, which enable the tracking of persister cells[145,146], or report on host responses like oxidative stresses[147,148]. Alternatively, biorthogonal non-canonical amino acid tagging (BONCAT) facilitates the labelling, subsequent visualization and quantification of bacterial translationally active subsets[149]. Stable isotope probing and sorting is another way to study the function of microbial taxa in their natural environment. Raman activated cell sorting (RACS) combines isotope labelling with microfluidics, optical tweezing, and Raman microspectroscopy enabling to sort complex intestinal, soil or marine microbial communities based on their ability of deuterium metabolization[150]. Since cells for FACS and RACS are only probed once before analysis, they just allow an end-point analysis of a cellular phenotype. Tracking bacteria over time to select rare or transient phenotypes can be provided by Single-cell isolation following time laps imaging (SIFT). SIFT, an integrative platform which combines long-term bacterial culture, time-lapse imaging, optical trapping and collection for downstream omics analysis[151]. Besides

that, the cultivation of gut bacteria is recently revolutionized by microfluidics. Meanwhile, microfluidic devices can cultivate microbial cells under oxygen exclusion for several days in picolitre sized droplets which subsequently can be sorted, imaged, and enriched for slow growing or rare taxa[152]. End-to-end solutions like this, providing cell-culture, imaging and sorting enable us to analyse rare phenotypes or populations at the single-cell level.

1.9 Single-cell RNA-seq of protozoa and fungi

The highly variable genome of protozoa allows them to quickly adapt to multiple hosts and to escape host immunity. To get deeper insights in theses mechanisms it is crucial to analyse its individuals at different developmental stages. Over the past few years, Smart-seq (switching mechanism at 5'end of RNA template) protocols have been proven to be among the most sensitive and robust methods to generate single-cell libraries on mammalian cells providing a complete coverage across the genome[117,144,153]. With some minor adaptions to minimize the unmapped reads, the Smart-seq2 protocol has been used for single-cell studies of *Plasmodium* and *Trypanosoma*[154,155].

As malaria parasites (*Plasmodium*) adopt a remarkable variety of morphological stages during its life cycle due to their transition through a mosquito vector to the mammalian host, combined with the fact that ~40% of its genome is functionally unassigned, it predestines them to be analysed on a single-cell level. Howick *et al*. provide a single-cell analysis of transcription of ten different life cycle stages of the rodent hosted model organism *Plasmodium berghei* summarised in the comprehensive resource 'Malaria Cell Atlas'[156]. To each stage, the authors assigned a pattern of co-expressed genes, unveiling potential mechanisms of the pathogen to adapt the host via cell-to-cell variability. Another finding is a morphologically indistinguishable subpopulation which can be generated by the pathogen, which can account for 1-30% and are able to persist in red blood cells. A broader approach was taken by Poran *et al*., using droplet based single-cell RNA-seq and screening more than 18.000 cells, they were able to resolve the parasites temporal dynamics regarding asexual replication in a persistent state and the shift back to sexual development[157]. Another representative of protozoan parasites is the unicellular parasite *Trypanosoma brucei* which causes human and animal African trypanosomiases better known as sleeping-sickness, after their transmission to their vertebrate host by the tsetse fly[158]. Once arrived in the mammalian host, versatile sets of variant surface glycoproteins (VSG) coat their surface to prevent host recognition. Key to decode the persistence mechanism of

Trypanosoma, is to understand the VSG switching principle. Using scRNA-seq, Müller *et al*. successfully elucidated the structure of VSG expression and linked it to the global genome architecture, local chromatin conformation, and histone variants[155].

Since protozoans are single-celled eukaryotes, standard eukaryotic scRNA-seq methods can be applied for library preparation and analysis. In contrast, fungi contain only 1-3 pg of total RNA, a fraction (1/10) of an average eukaryotic cell, which is additionally encapsuled by a cell wall and thus difficult to access. Three recent protocols provide different workflows, overcoming challenges like cell wall and low abundant transcripts. While Gasch *et al*. build up on an automated Smart-seq/C1 approach with previous enzymatic lysis[159], Nadal-Ribelles *et al*. and Saint *et al*. developed independent protocols[160,161]. Yeast single-cell RNA sequencing (YscRNA-seq) combined with index sorting could reveal a linear relationship between RNA content and cell size of *Saccharomyces cerevisiae*. The inherently strand-specific protocol captures the 5' End of transcripts and has the ability to record absolute gene expression and transcription start sites (Fig. 3, Table 21). Single-cell RNA barcoding and sequencing (SCRB-seq) was integrated with tetrad dissection microscopy to collect fission yeast (*Schizosaccharomyces pombe*) at precise cell sizes unveiling cell-size-dependent heterogeneity in gene expression programs influenced by environmental changes (Fig. 3). Both protocols provide multiplexing strategies via cell-specific barcodes and introduce unique molecular identifiers (UMIs) allowing the analysis of several hundred cells at a time and supporting normalization of RNA read-counts. Between 1000-3500 genes ca be detected per cell (Table 21). More recently, increasing throughput, the customization of the lysis-step to an enzyme-base permits usage of the 10x genomics platform to sequence thousands of cells at once via droplet-based sequencing[162,163].

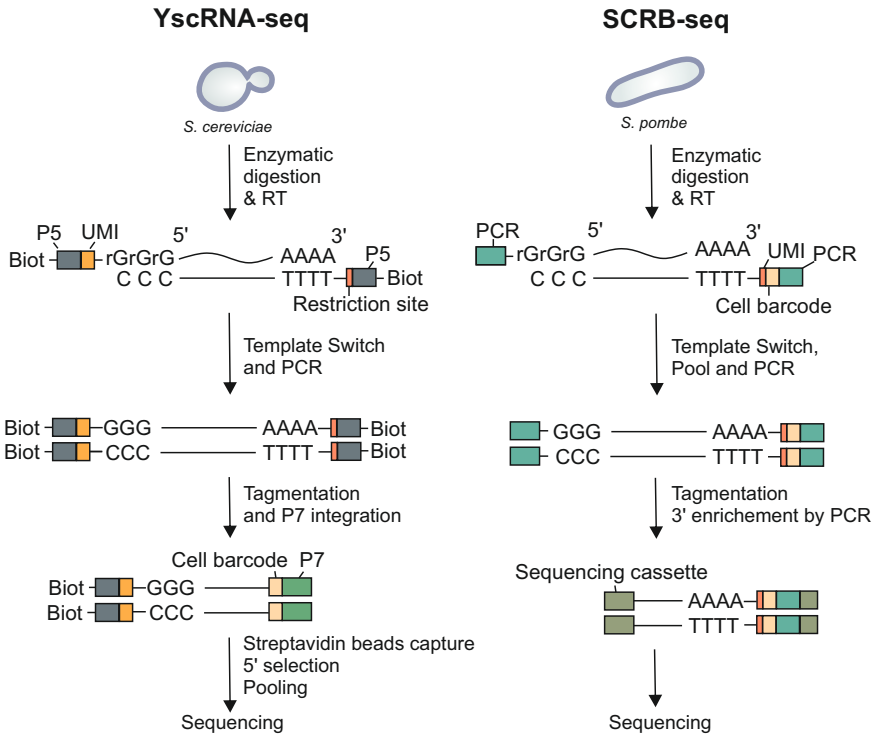

Fig. 3: scRNA-seq methods tailored to capture the transcriptome of yeast. Both protocols target poly(A) transcripts via oligo(dT) initiating reverse transcription. After 5' template switching the cDNA gets amplified in a PCR. Both methods provide UMIs and cell barcodes. (**Left**) YscRNA-seq (Yeast scRNA-seq) maps transcription start-sites using enzymatic 3' cleavage after biotin capture with streptavidin-beads. (**Right**) SCRB-seq (single-cell RNA barcoding and sequencing) specifically enriches for 3'-end transcripts. *Imdahl & Saliba Current Opinion in Microbiology, 2020*[113]

1.10 Single-bacteria RNA-seq

It is a long-held goal to resolve the gene expression heterogeneity of isogenic bacteria. As mentioned before there a several major challenges that must be overcome to achieve this aim. Briefly, the bacterial RNA content is ranging in levels between 1-100 femtograms which is tenfold to a hundredfold less than what can be found in a mammalian cell. Not polyadenylated mRNA transcripts complicate the discrimination of mRNA and rRNA and the high turnover of the bacterial transcriptome require rapid handling and tailored protocols.

Microbial split-pool and ligation transcriptomics (microSPLiT)[84] and prokaryotic expression profiling by tagging RNA in situ and sequencing (PETRI-seq)[83], are recent bacterial scRNA-seq method that are based on combinatorial barcoding to label the cellular origin of RNA[164]. These methods provide high throughput, as they can analyse thousands of bacteria simultaneously (Fig. 4b). In multiple indexing and ligation steps an individual barcode is generated and associated to the cDNA *in situ*. Additionally, the microSPLiT protocol provides a specific mRNA-polyadenylation reaction, with subsequent poly-(A) capture, to avoid the capture of rRNA. These methods are capable to perform scRNA-seq on both Gram-positive and Gram-negative species (*Escherichia coli, Bacillus subtilis, Staphylococcus aureus*). A range of 100-300 transcripts can be identified and assigned to specific growth conditions. The protocols are a great achievement for the investigation of larger bacterial populations or microbiome research, however, when it comes to transient and rare cells e.g., in an infections process, a more tailored and precise method is required. Multiple annealing and dC-tailing based quantitative single-cell RNA sequencing (MATQ-seq) is such a method, which is capable to analyse individual transcriptomes in a one-cell-per-well-approach (1.7.5 MATQ-seq). Imdahl *et al*. could conduct single-bacteria RNA-seq of *Salmonella* Typhimurium and *Pseudomonas aeruginosa* in defined infection relevant growth conditions reporting growth-dependent gene expression patterns (Fig. 4a).

Building on scRNA-seq a further aim is to simultaneously resolve the transcriptome of host and microbe since many pathogens find their niches in the intracellular milieu. Dual-RNA approaches scaled down to the single-cell level will be decisive tool to elucidate pathogenicity mechanisms and the respective host responses[165]. The immense challenge here is that bacterial RNA accounts for only 0.05% of the total RNA of both organisms[166]. Approaching this objective, Avital *et al*. developed scDual-seq to map the gene regulatory program of host and pathogen during *Salmonella* infection utilizing the cell expression by linear amplification and sequencing (CEL-seq2) -protocol, capturing the transcriptome of Macrophages infected with multiple intracellular bacteria (Fig. 4c)[167,168]. Although not reaching a single microbe resolution this technique represents an important advance in understanding the relationships among different states of infection.

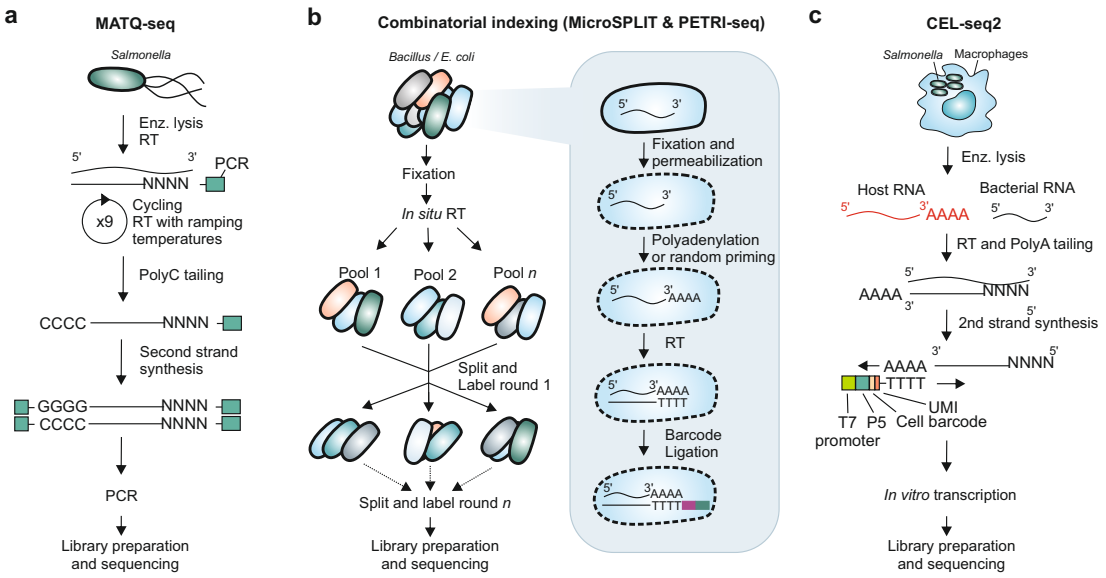

Fig. 4: scRNA-seq methods aiming to capture bacterial transcriptomes. **a**) MATQ-seq captures the transcriptome of single bacteria in a one-cell-per-well approach. RT is carried out with multiple cycles and ramping temperatures before poly(C) tailing of the first strand cDNA and followed by second strand synthesis and PCR amplification. **b**) MicroSPLiT (Microbial split-pool and ligation transcriptomics) and PETRI-seq (prokaryotic expression profiling by tagging RNA in situ and sequencing) are based on combinatorial indexing. After fixation and permeabilization several rounds of barcode ligation create a unique label for each cell. While PETRI-seq captures the transcripts via random primed RT, in MicroSPLiT mRNA is poly-adenylated before reverse transcription. **c**) CEL-seq2 is intended to simultaneously capture the transcriptome of host and invading bacteria. An individual barcode is introduced during RT. After second strand synthesis cells can be pooled for T7-promotor integration and PCR amplification. Imdahl & Saliba Current Opinion in Microbiology, 2020[113]

1.11 MATQ-seq

Gene expression is a fundamentally stochastic process, and inherent transcriptional randomness is leading to considerable cell-to-cell variations[54,169], which should be distinguished from technical noise in library preparation and sequencing. MATQ-seq is a highly sensitive and quantitative method for single cell sequencing which systematically diminishes technical noise, capturing whole transcriptomes of single cells[111]. Its ability to detect biologically unbiased and technically unaffected transcriptional variation among cells of the same population indicates the top-level sensitivity. MATQ-seq provides whole gene body coverage and therefore the detection of total RNA including noncoding and most importantly non-polyadenylated RNA, which enables its adaption to bacterial RNA-seq. Decisive for the sensitivity of single-cell

RNA-seq is the efficiency of reverse transcription and after that the successful generation of PCR amplicons[170]. Improving efficiency, Sheng *et al*. introduce special primers which are based on multiple annealing and looping based amplification cycles (MALBAC), a method enabling quasilinear whole-genome amplification[171]. Originally developed for single-cell DNA sequencing, MALBAC primers create amplicons containing complementary ends that form a loop, preventing exponential amplification of DNA. Hence only the template gets amplified while the copies that form loops are excluded, reducing amplification bias[171]. Hybridizing these primers, containing random nucleotides, to the internal regions of transcripts at low temperatures promotes a successful RT across the transcripts and the detection of non-polyadenylated RNA[111]. Following reverse transcription, a dC-tail is added to first strand cDNA, to enable an efficient second-strand synthesis using G-enriched MALBAC primers. Subsequently the libraries are amplified via PCR.

Aim of the project

Pushing to the limits of single-cell RNA sequencing we try to overcome the technical hurdles regarding accessibility of bacterial RNA in combination with the incompatibilities of most existing scRNA-seq protocols, to capture the transcriptome of individual prokaryotes. The MATQ-seq protocol provides poly-(A) independent priming and an outstanding RT-efficiency and is therefore ideally suited for adaption to bacteria. Using the well characterized microbial model organism *Salmonella* Typhimurium as well as *Pseudomonas* aeruginosa, we want to systematically validate the capabilities of the customized protocol and benchmark it to published bulk-RNA-seq experiments. Applying different infection relevant growth and shock conditions we aim to capture differently expressed genes correlating to the respective conditions and those found in benchmark-studies[172].

Human intestinal tissue model & infection

1.12 Tissue Models – in vitro test systems

The field of tissue engineering is an interdisciplinary research environment which combines cell-biology, material science, electrical and mechanical engineering involving 3D bio-printing and the development of bio-reactors[173]. Over the past few years human tissue models are emerging as an integral part in research of host-pathogen interaction and disease modelling[174,175]. Recent efforts of tissue engineering automation promise great potential and are of particular interest to clinicians and industry[176–178].

Generally, the principle of tissue engineering usually starts with primary cells isolated from a human donor which get expanded and subsequently seeded on a scaffold. The three-dimensional construction generates essential stimuli needed for differentiation and formation of a tissue-like morphology mimicking the original as close as possible. The matured *in vitro* tissue can then be used for research use or clinical implantation[179]. A great variety of human tissues has already been translated to 3D *in vitro* models finding application in numerous research areas[173]. Especially fields like drug-testing and infection research are still highly dependent on animal models which have long time been the gold standard for assessing the safety and efficacy of drug candidates or chemical compounds[180]. Despite the extraordinary scientific research made possible by animal experimentation in recent decades, there are some serious disadvantages in addition to ethical concerns. Clinical research and infection research in particular, suffer from the limited comparability of animals and humans due to serious anatomical and physiological differences[181,182]. Furthermore, several obligate human pathogens such as *Bordetella pertussis* either cannot be studied at all in animal models, or show a different pathogenesis[183]. For cosmetics and household products, the EU implemented a ban of animal testing in 2013, and already in 1959, the basic principle of 3R – Refine, Reduce, Replace animal testing was introduced by Russel an Burch[184]. Consequently, there is great interest in the current attempts aiming to replace animal testing by advanced *in vitro* approaches. However, to develop a valid human *in vitro* test system, one must first understand the physiology of a particular tissue *in vivo* to properly reproduce and validate it.

1.13 Physio-morphological characteristics of the small intestine

In adult humans, the small intestine, with a length of 3-6 m, connects the stomach with the large intestine and fulfils a variety of tasks, such as food digestion or nutrient transportation but also functioning as a protective barrier preventing infiltration of pathogens or harmful substances. The morphologically hollow tube has a multilayer composition. The outer sheath, called *tunica serosa*, forms the demarcation to the peritoneal lumen and is subtended by two muscle layers. This *tunica muscularis* consists of a thin linear and a thick circular muscle layer, responsible for the intestine peristaltic movement. Connecting the muscle layer and the *Mucosa*, the *tela submucosa* is a highly vascularized and collagenous structure. One layer further inside the tripartite *Mucosa* comprising *lamina muscularis*, *lamina propria*, and the luminal epithelium, enables nutrient intake and energy supply. To ensure effective food digestion and absorption of various nutrients such as sugars, amino acids, lipids, fatty acids and carbohydrates, the small intestine secretes mucus containing the necessary enzymes and ions[185,186]. Besides this task, another vital role of the epithelium is its barrier function, delineating the intestinal environment from inner body while enabling trans- and paracellular transportation at the same time[187,188]. At the cellular level, barrier integrity is based on the formation of tight and adherence junction proteins such as E-cadherin, claudin or occludin[189–192]. The small epithelium has a sophisticated microstructure consisting of villi and crypts. The latter ones are tube-shaped invaginations of the epithelium accommodating multipotent intestinal stem cells, which possess the ability to differentiate into all intestinal cell types, and through this regenerate the epithelium every 3-5 days[193–197]. One of the most common markers for adult intestinal stem cells was described as leucine-rich repeat-containing G-protein coupled receptor 5 ($LGR5^{+}$)[198]. Located at the stem cell niche at the bottom of the crypt, they undergo asymmetric cell divisions, where one descendant stays in the niche retaining stem cell characteristics, while a second daughter cell develops to a "transit amplifying cell". With ongoing proliferation, the transit amplifying cells migrate along the villus and differentiate to the different intestinal cell types before they undergo programmed cell death (anoikis) reaching the top of the villus[199–201]. Another cell type resident in the crypt are the long-lived and defensive paneth cells, which secrete antimicrobial peptides, cytokines, and proteases, but also important growth factors such as Wnt3a, Noggin, R-Spondin or the epidermal growth factor (EGF)[202–204]. The protein levels of those factors decrease towards the villus in a gradient manner, while concentrations of morphogenic proteins like hedgehog or ephrin B1 which regulate the differentiation of transit amplifying cells are increasing[191,205–207]. In contrary to the crypts, the villus structure extends finger-shaped to the

lumen of the small intestine which increases the resorption area and optimises the nutrient uptake potential. Digestion of this nutrients is supported by enzymes secreted by the very abundant enterocytes at the villus[208]. Enterocytes are covered by microvilli which protrude into the lumen and foster the uptake of luminal nutrients, vitamins or water[209,210]. The concatenated and polarized layer of absorptive enterocytes is sustained by several differentiated cell types such as goblet cells, enteroendocrine cells, tuft cells and microfold (M) cells[211]. The cell layer is covered with a protective coat of mucus which is produced by the goblet cells and consists of Mucin-2 protein and antimicrobial peptides preventing microbial invasions[211,212]. Enteroendocrine cells account for 1% of the epithelium and combine several cellular subtypes that are responsible for vital functions including blood glucose homeostasis, appetite regulation or gut contractility by secreting specific hormones such as serotonin, glucose-dependent insulinotropic polypeptide or glucan like peptide. Besides that, they are the key sensors of microbial metabolites and able to release cytokines responding to pathogen associated molecules[213]. The small intestine is exposed to an incredible number of microbes, many of which commensals, that are beneficial and aid nutrient uptake. Among these microbes, however, there may also be pathogens whose infection or excessive growth must be prevented. A major sentinel for microbial homeostasis is represented by the goblet cells, which actively respond to microbial invasion with extensive mucus secretion and can thus flush away the invaders[214]. Other players in the defence against invading pathogens are the villus-residing tuft-cells and M-cells, which are part of a concerted immunological system within the small intestine. While the chemo sensitive tuft cells are initiating type 2 immune responses upon intestinal infections, M cells are in close contact to the innate immune system in the underlying basal lamina, where they conduct phagocytosis and transcytosis of antigens, molecules and microbes[215–217].

1.14 Engineering the small intestine

In vitro tissue-engineered small intestine models hold great potential to study and understand an organs physiology in health and disease. Early developed intestinal tissues consisted of a 2D monolayer and were only suitable for short-term culture of primary epithelial cell types. Missing cell-cell or cell-matrix interaction prevented a longer cell culture. Besides primary cells from healthy tissue, adenocarcinoma-derived cell lines such as HT29 or Caco-2 cells have also been used extensively to create *in vitro* models of the human small intestine. Despite their

widespread use, these models have some specific characteristics that make them less than ideal as a general model[218–220]. While HT29 tissue models show a disproportionate differentiation of Goblet cells, Caco-2 based models exhibit a more enterocyte-heavy phenotype compared to the in-vivo tissue[221,222]. Although, the use of cancer-based monolayer cultures led to important findings in function and physiology of the small intestine, they still have limitations in terms of complexity and completeness of their cellular composition. Cell line-specific artificial gene and protein expression do not allow reliable conclusions to be drawn about the *in vivo* tissue.

Representing the functional unit for tissue regeneration *in vivo*, intestinal stem cells harbor the potential to give rise to all intestinal epithelial cell types. Fortunately, they keep their abilities to proliferate, differentiate and self-organise also *in vitro,* where they form three-dimensional structures, mimicking *in vivo*-like tissue composition[223–225]. These so-called intestinal organoids or "mini-guts" are of high interest for basic pharmacological or translational research and are meanwhile very well-established[204]. Organoid cultures may even reflect the small intestines villus morphology including cyst-like structures surrounding the central lumen with highly polarized epithelial cells[226]. Organoids that are grown from isolated intestinal crypts from a surgical or endoscopic biopsy (Enteroids) retain their genetic background and thus a donor-specificity. Although this background complicates reproduction and standardization, it has a great utility for establishing of biobanks of health and disease conditions, and further enables high throughput drug screening that may facilitate personalized therapy[227]. Besides their use as disease models, translational research or drug-screening, organoids can also be used to investigate infection processes and host-pathogen interaction[228,229]. However, organoid culture remains difficult, and infection via microinjection lacks throughput. Technical limitations also arise from the use of 3D Matrigel for the growth of organoids. Recently established Transwell systems are based on synthetic or biological membranes, that function as a scaffold for the intestinal cells and provides two separated compartments. Thus, the apical, luminal side and the basolateral side of the small intestine can be recreated, making these models suitable for absorption or transport studies. The membranes used can either be based on polyester, polycarbonate, or polyethylene terephthalate which come in different thicknesses and pore sizes, or be made from biological materials such as the decellularized small intestinal submucosa (SIS)[230–232]. Synthetic materials, while offering a high degree of standardization, they require coating with extra cellular matrix proteins to allow cell growth and may affect distribution and diffusion of cellular compounds. Although biological matrices lack standardisability, they do form a biocompatible 3D microenvironment[233]. Next-generation models try to guide morphogenesis of stem cells extrinsically via specialized hybrid scaffolds

out of type-I collagen and Matrigel containing key constituents of the native basement membrane. Aiming to create a functional organ-on-a-chip those tissues are mounted on a perfusable platform containing microchannel with microcavities mimicking the structure of human crypts[234]. Approaching the goal of a standardized *in vitro* model representing the entire diversity and functionality of the human gut, several studies try to increase the complexity in terms of vascular systems, microbiota exposition, the inclusion of immune cells or the application of growth-factor, nutrient, or oxygen gradients[235–239]. Even though there is still a long way to go to engineer a universal and standardized small intestinal model, the research in this field is very active and constantly publishes versatile promising approaches.

Aim of the Project:

In this project we want to characterize a human intestinal epithelia tissue model (hITM) using top-notch imaging and scRNA-seq technologies to identify physiological and compositional similarities and differences to the *in vivo* tissue. Starting with histochemistry, immunofluorescence microscopy up to electron microscopy, we want to unveil integrity and polarization of the model. ScRNA-seq enables us to reveal the cellular composition and associated proportion of different cell types within the tissue model, so that we can validate the degree to which it reflects the real tissue. Furthermore, we aim to assess hITM as an infection model, exposing it to *Salmonella* Typhimurium. Besides immunofluorescence microscopic analysis of the infection dynamics and heterogeneity, we apply deep single-cell RNA sequencing to different time points of the infection with the objective to unveil favourable entry points and launch of *Salmonella* replication. Further validation of the single-cell data can be conducted using HCR-FISH to demonstrate transcripts directly in the tissue.

2. Material and Methods:

2.1 MATQ-seq

2.1.1 Cell culture – single bacteria RNA-seq

Salmonella Typhimurium

For our experiments we cultured *Salmonella enterica ssp. enterica* serovar Typhimurium. *Salmonella* Typhimurium SL1344 strain (internal reference: JVS 1574) was grown in 5 ml Lennox broth (LB) in a 37°C at 220 rpm in a shaking incubator overnight and diluted 1:1000 in a new 250ml flask containing 10 ml of LB. Again, the cells were grown at 37°C and 220 rpm to an OD_{600} of 0.3 indicating that the mid exponential phase (MEP) had been reached. All incubation steps were performed in a New Brunswick Innova 44 shaking incubator. Overnight cultures for SPI-1 condition were grown to stationary phase, indicated by an OD_{600} of 2.0. Immediately after reaching the wanted phases, cells of each condition were pelleted by centrifugation at 14.000 g for 4 minutes and resuspended in 1x Dulbecco's phosphate buffered saline (DPBS, Gibco). This washing step was repeated before resuspending the pellet in 1 ml of a 1:1 RNA-later DPBS solution to prevent RNA form degradation.
Using FACS for the separation of the cells, a reference SL1344 GFP-strain (internal reference: JVS 3858; Papenfort *et al*. 2009) was needed, to calibrate the gating of the flow cytometer. This strain was cultured in the same way as the SPI-1 condition strain (JVS 1574).

NaCl shock- and anaerobic shock-condition

Shock conditions were generated as described in Kroger, Colgan *et al*. 2013[172]. For osmotic shock generation, NaCl was added to MEP-bacteria cultures to a final concentration of 0.3 M, followed by 10 minutes continued incubation at 37°C shaking. For anaerobic shock generation, 5 ml of MEP-bacteria cultures were filled in a test-tube, covered with 2 ml of mineral oil and subsequently incubated at 37°C for 30 min without agitation. After final incubation step, 2 ml of SPI1 untreated o/n culture or shock cultures, cells were pelleted and washed twice with 1 ml of DPBS (Gibco) before resuspension in 1ml of a 1:1 RNAlater (Thermo Fischer) - DPBS solution to prevent RNA degradation.

2.1.2 Flow cytometry and fluorescence-activated cell sorting (FACS)

Gating and Selection of Bacteria

Gating and selection of bacteria was performed by comparison of JVS 3858 (GFP) and wild type Salmonella (JVS 1574). Among granularity and cell diameter, the FITC-Channel dependent on cell count, which is defined by the GFP expression intensity, was used to discriminate single cells form doublets, debris, and dust. To make sure that only single cells were sorted, the gate for the GFP expressing strain was adapted for the wild-type bacteria.

The flow-cytometer based analysis and separation of the cells by sorting was performed using a FACS aria III (BD Biosciences) using a 70 μm nozzle and medium flow rate. Before sorting, RNA*later* treated cells were diluted 1:100 in PBS. Single cells were sorted in unskirted 48-well plates (Brand) containing 2.6 μl lysis-buffer per well (see below). Sorted plates were stored at -20°C.

2.1.3 Cell-lysis

Single bacteria were directly sorted in wells containing 2.6 μl lysis-buffer with the following composition:

Table 2: Lysis buffer

Reagent	**Volume in μl**
10 x Lysisbuffer (Takara)	0.26
RNase Inhibitor (40 U/μl; Takara)	0.03
1 x DPBS (Gibco)	0.26
Lysozym (50 U/μl; Epicentre)	0.1
Nuclease free H_2O (Ambion)	1.95
Total Volume:	2.6

Before sorting bacteria, the lysis buffer was freshly prepared and filled into unskirted 48 well plates (Brand) with 2.6 μl per well, sealed with micro seal ‚B' (BioRad) and kept on ice until sorting.

2.1.4 MATQ-seq protocol

The MATQ-seq protocol was performed according to the initial report [111] with small adjustments. All primers were ordered at IDT (Integrated DNA Technologies). The primer mix was prepared using following primers:

Table 3: Primer mix MATQ-seq

Primer/Reagent	**Sequence 5' – 3'**	**Volume in µl**
GAT27 dt	GTG AGT GAT GGT TGA GGA TGT GTG GAG NNN NN TTTTTTTTTTTTTTTTTTT	0.12
GAT27 5N3G	GTG AGT GAT GGT TGA GGA TGT GTG GAG NNN NNG GG	0.4
GAT27 5N3T	GTG AGT GAT GGT TGA GGA TGT GTG GAG NNN NNT TT	0.4
Nuclease free H_2O (Ambion)		7.08
Total Volume:		8.0

Pre-RT:

Every well, containing lysed cells, was supplemented with a pre-RT mix, containing the following ingredients:

Table 4: pre-RT mix

Reagent	**Volume in µl**
DTT	0.05
Primer mix	0.4
dNTPs (10mM each nucleotide, NEB)	0.12
Total Volume:	0.57

The plate was sealed and incubated for 3 minutes at 72°C in a thermocycler and subsequently kept on ice for at least 1 minute.

Reverse transcription:

During incubation the reverse transcription mix (RT-mix) was prepared as follows:

Table 5: RT-mix MATQ-seq

Reagent	**Volume in μl**
5 x RT-Buffer (Life Technologies)	0.8
DTT	0.2
RNase Inhibitor (40 U/μl, Takara)	0.1
Superscript III (LifeTechnologies)	0.15
Nuclease free H_2O (Ambion)	1.15
Total Volume:	2.4

After addition of 2.4 μl of RT-mix to every reaction-well, plates were sealed again, and placed in the thermocycler for the RT-program:

Table 6: Thermocycler program for reverse transcription MATQ-seq

Time	Temperature	
12 sec	8°C	10x
45 sec	15°C	
45 sec	20°C	
30 sec	30°C	
2 min	42°C	
3 min	50°C	
15 min	50°C	
forever	4°C	

Primer-digestion:

Primers for reverse transcription were then digested using 0.2 μl T4 DNA Polymerase (New England Biolabs). The polymerase was added after 1 minute of incubation at 50°C at 37°C. The actual digestion was performed with the following steps in the thermocycler.

Table 7: Thermocycler program for primer digestion

Time	Temperature
40’	37°C
20’	75°C
40’	37°C
20’	80°C
forever	4°C

RNA digestion:

Reverse transcribed RNA was digested with 0.1 μl RNase If (New England Biolabs) and 0.1 μl RNase H (New England Biolabs). After mixing the enzymes to every well, the plate was sealed and incubated for 15 min at 37°C and 15 min at 72°C before keeping temperature at 4°C in the thermocycler.

Tailing:

The following mix was prepared to perform the tailing reaction:

Table 8: dC tailing mix

Reagent	**Volume in μl**
10x Tdt-Buffer (New England Biolabs)	0.4
dCTP (100 μM, Life Technologies)	0.4
Tdt Terminal Transferase (New England Biolabs)	0.1
Nuclease free H_2O (Ambion)	3.13
Total Volume:	4.03

For the tailing, 4.03 μl of tailing-mix were added to each well, before sealing and incubating in the same way as for RNA-digestion step.

Second strand synthesis:

The second strand synthesis required the following mix:

Table 9: Mix for second strand synthesis

Reagent	**Volume in µl**
10x Thermopol-Buffer (New England Biolabs)	1.5
dNTP (10 µM each nucleotide)	1.25
GAT 21 6N3T (100 µM)	0.125
Nuclease free H_2O (Ambion)	12.925
Total Volume:	15.8

Table 10: Primer sequence GAT21 6N3G

Primer	**Sequence 5' – 3'**
GAT21 6N3G	GAT GGT TGA GGA TGT GTG GAG NNN NNN GGG

15.8 µl of the reaction-mix were added to each well and mixed thoroughly. Plates were then sealed and heated at 98°C for 1 min in the thermocycler. After cooling down to 48°C, the plates were removed from the cycler, and 0.4 µl Deepvent exo- DNA Polymerase (New England Biolabs) were added to the wells. After resealing, the cycler program was proceeded.

Table 11: Thermocycler program for second strand synthesis

Time	Temperature	
1 min	98°C	
hold	48°C	add 0,4 µl Deepvent exo-
20 sec	48°C	11x
1 min	72°C	
2 min	72°C	
forever	4°C	

Amplification:

During second strand synthesis, the master mix for the final amplification was prepared as follows:

Table 12: PCR mix for MATQ-seq

Amplification	**Volume in µl**
10x Thermopol-Buffer (New England Biolabs)	13
dNTP (10 µM each nucleotide)	3
GAT27 PCR	0.8
Nuclease free H_2O (Ambion)	114
Deepvent exo- DNA polymerase	3
Total Volume:	133.8

Table 13: Primer sequence GAT27 PCR

Primer	**Sequence 5' – 3'**
GAT27 PCR	GTG AGT GAT GGT TGA GGA TGT GTG GAG

For amplification 133.8 µl were added to each reaction well and mixed thoroughly. Each well containing 160 µl, was split into four wells with 40 µl each. The plates were placed in a thermocycler to run the following PCR program.

Table 14: Thermocycler program for amplification MATQ-seq

Time	Temperature	
30 sec	95°C	
15 sec	95°C	
20 sec	62°C	24x
2 min	72°C	
5 min	72°C	
forever	4°C	

Amplified samples were directly processed further or stored at -20°C.
All thermo cycling steps were performed in a BioRad T100 thermal cycler.

2.1.5 cDNA purification

Amplified cDNA was purified using Ampure XP-beads (Beckmann Coulter). At first beads were equilibrated to room temperature for at least 30 minutes. Then beads were added to each reaction well in a 1:1 ratio (40 µl) and mixed until a homogenous phase was reached. The

mixture was then incubated for 8 minutes at room temperature. Subsequently samples were placed on a magnetic stand until the liquid became clear, but at least for 5 min. With the samples kept on the magnetic stand, the supernatant was removed and 200 μl 80% EtOH were added without mixing, and reaspirated after 30 seconds. This washing step was repeated, before beads were dried for about 8-10 minutes. Dried beads were resuspended in 17 μl Nuclease free H_2O or Elution buffer and incubated off the magnet for 2 min. After following 2 min of incubation again on the magnet the purified cDNA containing supernatant was aspirated and placed in a new PCR tube.

2.1.6 cDNA quantification & quality control

Qubit and Bioanalyzer

The cDNA is quantified using a Qubit™ 4 Fluorometer (ThermoFischer) and a Qubit HS-DNA kit (ThermoFischer). Briefly, 198 μl of Qubit dsDNA HS Buffer and 1 μl of the Qubit dsDNA HS dye were mixed and 1 μl purified cDNA was added to a supplied Qubit HS tube. After mixing thoroughly, the tube was incubated for at least 1 min at room temperature. For calibration of the Qubit Fluorometer, the two supplied standards were used. Ten μl of Standard-1 and Standard-2 were mixed with 190 μl Qubit ds DNA HS Buffer in two separate tubes, also incubated for 1 min at room temperature, before they were transferred to the fluorometer for calibration.

To observe the fragment size distribution a chip-based gelelectrophoresis was performed using an Agilent 2100 Bioanalyzer and an Agilent HS DNA Chip. Following the Qubit results cDNA samples were diluted to 1 ng/μl. Of this dilution 1 μl was used for input. The Bioanalyzer was used according to the manufacturer's instructions (Aglient Technologies 2017).

2.1.7 cDNA library preparation and sequencing

Nextera XT and Nextseq

Libraries were generated using Nextera XT (Illumina) with small modifications, previously described in [241,242]. Summarized 1 ng of cDNA was used as input. Tagmentation was performed with one-quarter of the recommended volumes and an elongated fragmentation time (10 min). 15 μl Elution Buffer (EB, Quiagen) were used for resuspension.

Pooled libraries were sequenced using an Illumina Nextseq 500 platform and an Illumina Nextseq high output 2 x 75 bp PE (paired end) sequencing kit.

2.1.8 Bioinformatic analysis

The bioinformatic analysis was processed using Spliced Transcripts Alignment to a Reference (START) which is a R-based toolkit for mapping the sequencing reads to a reference genome [243]. Count-tables were generated using HTSeq, which is a Python framework for high-throughput sequencing data [244]. Principal Component Analysis (PCA) have been generated via the R package factoextra (*Package "factoextra" Extract and Visualize the Results of Multivariate Data Analyses*, 2017). All computational analysis was performed form Ehsan Vafadarnejad.

2.2 Human intestinal epithelium tissue model infection with *Salmonella* Typhimurium

All experiments of this project have been performed in close collaboration with Thomas Däullary from the chair of Tissue Engineering and Regenerative Medicine (Würzburg). The HCR-FISH was performed by Tobias Krammer from the SIGA group (HIRI). The Bioinformatic data procession was conducted by Ehsan Vafadarnejad and Oliver Dietrich (SIGA group HIRI). Imaging was mainly contributed by Thomas Däullary.

2.2.1 2D human intestinal tissue model (hITM)

To generate the human intestinal epithelial tissue model (hITM), human-derived primary small intestinal enteroids were applied on a decellularized porcine matrix deriving from the small intestinal submucosa (SIS). The initial proliferation takes 5-7 days and is subsequently followed by a 4-day differentiation phase. The hITM generation was performed by Thomas Däullary from the chair of Tissue Engineering and Regenerative Medicine (Würzburg). The overall procedure was performed following the published protocol [230].

2.2.2 hITM infection

Cell culture – Inoculum

The constitutively GFP expressing Wild-type derivative of *Salmonella* Typhimurium strain SL1344 (JVS-3858, [240]) was used for imaging and SMART-seq experiments. It was cultured in 5ml Lennox broth (LB) at 37°C under constant agitation at 220 rpm (New Brunswick, Innova 44) over night. A 1:100 dilution of the latter culture was then grown to an OD of 2.0.

Where required, the fluorescence dilution mutant pFCcGi (JVS-11424)[73] (Stapels et al. 2018) was used. It was cultured as described above. When setting up the day-culture 0.1% arabinose was added to the LB-media to induce GFP.

Tissue infection with *Salmonella* Typhimurium

1 ml of *Salmonella* cultures (JVS-3858 or JVS-11424) at OD 2 were pelleted by centrifugation, washed with PBS and resuspended in 1 ml of DMEM. For each MOI the correct amount of inoculum was calculated and transferred to the Crypt-media. The media covering the model was removed before 300μl of the infection-media were added to the upper compartment, and 900μl DMEM were added to the lower compartment of the transwell inserts.

For adhesion the plates were centrifuged for 10 min at RT using a rcf of 250 and incubated at 37°C and 5% CO_2, 95% O_2 (Heracell 240i Thermo Scientific) for 1h. After the adhesion the supernatant was collected and exchanged to high-gentamycin-media containing 50μg Gentamycin per ml. After an incubation of 30 min at 37°C and 5% CO_2, 95% O_2, the high-gentamicin media gets exchanged to a low-gentamycin-media (10μg Gentamycin/ml) for (x)h at 37°C and 5% CO_2, 95% O_2.

Tissue model dissociation

After the desired infection time (0.5 – 16h) the supernatant of upper and lower compartment of the transwell was collected. Models then were then washed with 1x PBS + EDTA (conc?). Prewarmed Accutase (Gibco) + 1:1000 Rock-inhibitor (company?) was then added, 400μl to the upper compartment and 1000μl to the lower compartment, and incubated for 10 min at 37°C. To detach cells from the scaffold the 400μl in the upper compartment are pipetted up and down a few times and then transferred to a 2 ml reaction tube. Using tweezers, the scaffold was carefully transferred to the corresponding 2 ml reaction tube. Cells and scaffold were then incubated for 10 minutes at 37°C shaking (ThermoMixer, Eppendorf). After careful resuspension with a 1 ml Pipet, the scaffold was removed cautiously. Next the cells were centrifuged at 500 g for 3 min at room temperature. The supernatant was discarded, and the pellet washed with 1x PBS. This step was repeated twice before resuspending cells in 700 μl

1x PBS. From here on cells were stored on ice and filtered through a 40 μm Strainer (Miltanyi) before FACS-sorting or Drop-seq (10x Genomics).

2.2.3 Flow cytometry and fluorescence-activated cell sorting (FACS) hITM

Fluorescence-activated cell sorting (FACS):

FACS was used to discriminate between infected and uninfected cells. More in depth, infected cells were separated in three different clusters (H: High; M: Medium; L: Low) according to their bacterial load. Therefore, gates were set based on the GFP intensity. Infected cells were sorted to a 48 well plate into a lysis buffer under permanent cooling to 4°C by a BD Aria III.

FACS and flow-cytometry for fluorescent dilution strain:

Intestinal cells infected with the fluorescence dilution strain (JVS-11424) of the different time-points were gated as described based on their mCherry intensity. H, M, and L gates were sorted separately into reaction tubes containing 500 μl 1x PBS + 0,1% Triton x-100 (Sigma) to lyse the host-cells but not the bacteria. Cell-bacteria suspension was then centrifuged for 2 min at 14,000g. Following the supernatant was discarded. The pellet was then resuspended in 500 μl 1xPBS, before the samples were supplied to the FACS ARIA III again, to quantify the intracellular bacteria. Gates were set to detect all GFP-positive cells within the mCherry positive population.

2.2.4 Library preparation and Sequencing via SMART-seq

Salmonella infected cells were FACS-sorted into 2.6 μl of Lysisbuffer (0.26 μl 10x Lysisbuffer (Takara), 0.03 μl RNase inhibitor (40 U/μl; Takara) and 2.31 μl nuclease free H_2O (Ambion)) presented in a 48 well plate (Brand). Cell sorting was conducted using a BD FACS Aria III. After sorting cells were spun down quickly and immediately frozen at -80°C until library preparation. Libraries were produced using the SMART-seq2 V4 kit (Takara) using ¼ of the recommended volumes except for the volumes we followed the manufacturers manual. [117]. ERCC RNA Spike-Ins Mix1 (Invitrogen) were added to each library diluted 1:2,000,000. In brief, 0.3 μl CDS primer and 0.2 μl ERCC RNA Spike-In Mix 1 were added to each library which was subsequently incubated at 72°C for 3 minutes in a thermal cycler. Then the RT-mix containing 1 μl 5x Ultra Low First-Strand Buffer; 0.25 μl SMART-Seq v4 Oligonucleotide (48 μM), 0.125 μl RNase Inhibitor (40U/μl) and 0.5 μl SMARTScribe Reverse Transcriptase was

prepared. 1.9 μl RT-mix were added to each library before transferring it to the thermal cycler for reverse transcription. In the reverse transcription program, a 90-minute incubation at 42°C is followed by 10 minutes on 70°C before the reaction is kept on 4°C. For the cDNA amplification a mix containing 6.25 μl 2X SeqAmp PCR Buffer; 0.25 μl PCR Primer II A; 0.25 μl SeqAmp DNA Polymerase; and 0.75 μl Nuclease-Free water, was prepared and a total of 7.5 μl was added to each library. Subsequently the plate was transferred to the thermal cycler and the following amplification program was executed.

Table 15: Thermocycler program for amplification Smart-seq2 v4 amplification

Time	Temperature	
1 min	95°C	
10 sec	98°C	22x
30 sec	65°C	
3 min	68°C	
10 min	72°C	
forever	4°C	

All library preparation steps were carried out using the Mantis Microfluidic Liquid Handler (Formulatrix). All thermal cycling steps were performed in a T100 Thermal Cycler (Bio-Rad). Separate cyclers were used for pre-RT steps and Amplification. Ready amplified libraries were cleaned up using 12.8 μl AMPure XP beads (Beckman Coulter) and 0.3 μl 10X Lysisbuffer (Takara) (compare 2.1.5 cDNA purification). The cDNA was eluted in 15.5 μl Nuclease Free water. Ready libraries underwent a quality control via Qubit™ (Thermo Fischer) and 2100 Bioanalyzer (Agilent). Libraries containing less then 0.8 ng/μl of cDNA or showing an inappropriate bioanalyzer trace were excluded from further processing. The remaining libraries were diluted with nuclease free water to a final concentration of 0.4 ng/μl as input for the Nextera XT (Illumina) library preparation protocol (compare 2.1.7.1). The tagmented libraries underwent another quality control as mentioned before and were pooled according to their individual molarity. Finally, the ready pool was sequenced using the NextSeq 500 (Illumina) platform with a 75bp paired-end high output kit.

10x Genomics Chromium single-cell RNA-seq

Single cells were dissociated form the tissue model as described above. The cells of two different tissues were hashtagged with TotalSeq-A antibodies (Biolegend) following the manufacturers protocol for TotalSeq™-A antibodies and cell hashing with 10X Single 3' Reagent kit v3.1. Approximately 400.000 cells per sample were resuspended in 100 μl Cell Staining Buffer (Biolegend) and 5 μl Human TruStain FcXTM FcBlocking (Biolegend) reagent were added. For the blocking reaction the cells were kept on 4°C for 10 minutes. 1 μg of a unique TotalSeq™-Antibody was then added to each sample followed by a 30-minute incubation at 4°C. Afterwards cells were washed three times with 1 ml Cell Staining Buffer and spun down for 5 min at 350g and 4°C. Finally, the cells were resuspended in an appropriate volume of DPBS (Gibco) and passed through a 40 μm cell strainer (Flowmi™ Cell Strainer, Merck). To adjust the concentration to 1000 cells per μl, cells were counted in a Neubauer Hemacytometer (Marienfeld) and the according amount of DPBS (Gibco) was added. The hashtagged cells were pooled equally and ~20.000 cells were loaded in the Chromium™ Controller. The machine creates Gel Bead-In-Emulsions (GEMs) to separate single cells into a nanoliter compartment together with an individual barcode. Reverse transcription, cDNA amplification and the construction of gene expression libraries was performed using the Single Cell 3' reaction kit v3.1 (10x Genomics) and the associated protocol. The hashtag libraries were prepared following the cell hashing protocol for 10x Single Cell 3' Reagent Kit v3.1 from Biolegend. Incubation and amplification steps were carried out using a SimpliAmp Thermal Cycler (ThermoFisher). Library quantification and quality control was observed using a Qubit™ 4.0 Fluorometer (ThermoFischer) and a 2100 Bioanalyzer with High Sensitivity DNA kit (Agilent). Sequencing was performed on a NextSeq 500 sequencer (Illumina).

2.2.5 Histological and immunofluorescent characterization

Fixation of tissue models

For histological analysis human intestinal tissue models along with the SIS scaffold was washed with PBS^- and resuspended in 4% paraformaldehyde (PFA) for 4-20h at 4°C. After fixation, tissue models have been stored in PBS^- until they were stained.

Paraffin embedding, sectioning and rehydration

For section analysis, fixed tissues have been embedded in paraffine. For this purpose, the SIS scaffold was transferred to an embedding cassette in a filter paper. Loaded embedding cassettes were then paraffinized automatically in an embedding automate.

Table 16: Paraffin embedding procedure

Step	Solution	Time (h)
Removal of PFA	dH_2O	2
Dehydration	50% EtOh	1
	70% EtOh	1
	90% EtOh	1
	96% EtOh	1
	Isopropanol I	1
	Isopropanol II	1
	Isopropanol/Xylene (1:2)	1
Removal of alcohol	Xylene I	1
	Xylene II	1
Paraffinization	Paraffin I	3
	Paraffin II	3

Paraffin blocking was performed by transferring the models to a casting mold filled with liquid paraffin. Solid blocks were then cut into 5 μm thick sections using a microtome. Via a 60° C water bath, sections were transferred onto glass slides and subsequently dried over night at 37°C.

Before staining, fixed and embedded samples must be deparaffinized and rehydrated. Therefore, the paraffin was melted for at least 15 min in a 60°C water bath. Remaining paraffin was removed, before the section was rehydrated, via the following protocol:

Table 17: Deparaffinization and rehydration of tissue sections

Step	Solution	Time (min)
Rehydration	Xylene I	10
	Xylene II	10
	96% EtOh I	Dip 3x

96% EtOh II	Dip 3x
70% EtOh	Dip 3x
50% EtOh	Dip 3x
dH_2O	

Fixation embedding rehydration and staining was performed by Thomas Däullary.

Alcian blue staining

Visualization of acidic glycoproteins was performed via Alcian blue staining. After rehydration of the tissue sections, they were incubated for 30 min in 1% Alcian blue solution in order to stain the negatively charged sulphated proteoglycans. Subsequently, slides were washed in dH_2O and exposed to nuclear fast red solution for 5 min, counterstaining the nuclei. Samples were than washed again with dH_2O and dehydrated (Table 16, followed from bottom to top). For final conservation, samples were mounted with Entellan (Sigma-Aldrich) and covered with a coverslip.

Immunofluorescence staining of hITM

For wholemount staining of the hITM in a first step the SIS matrix was removed from the cell crown, and fixed tissue models were transferred to a 48-well plate. Matrices were then treated for 20 min with 200 μl 0,2% Triton X-100 shaking. Between two washing steps with PBS-T, samples were blocked in 5% donkey serum for 20 min shaking. Primary antibodies (Table 17) were then applied over-night at 4°C shaking followed by another three washes before secondary antibodies (Table 18) were applied for 1-2h at RT shaking. After further three washes fluorescently labeled Phalloidin (AF555 1:1000, AF488 1:500) and DAPI (1mg/ml, 1:10.000) were added for 20 min shaking for respective actin and nuclei staining. Finally, samples were washed again three times before the matrices were transferred to glas slides and mounted with Fluoromount G™ (Thermo Fisher) and subsequently sealed and conserved with a coverslip.

Image acquisition

Imaging of immunofluorescence labeled staining was performed using a confocal microscope TCS SP8 (Leica). For brightfield images of Alcian blue stained tissue sections an inverse fluorescence microscope BZ-9000 (Keyence) was used. Imaging was primarily performed by Thomas Däullary.

Antibodies

Table 18: List of primary antibodies used in this study

Antigen	Host	Clonality	Manufacturer/ Cat. #
Mucin 2	rabbit	polyclonal	Abcam/ 76774
Mucin 1	rabbit	monoclonal	Abcam/109185
Lysozyme	goat	polyclonal	SantaCruz/ sc27958
Villin	goat	polyclonal	SantaCruz/ sc7672
Occludin	mouse	monoclonal	Thermo Fisher / 33-1500
ZO-1	rabbit	polyclonal	Ptglab/ 21773-1-AP

Table 19: List of secondary antibodies used in this study

Antigen	Host	Conjugated Fluorochrome	Manufacturer
Rabbit	Donkey	Alexa Fluor 647	Invitrogen
Rabbit	Donkey	Alexa Fluor 555	Invitrogen
Mouse	Donkey	Alexa Fluor 555	Invitrogen
Goat	Donkey	Alexa Fluor 647	Invitrogen

2.2.6 HCR-Fluorescence In-situ Hybridization

Tissue samples were fixed at room temperature for 2h with paraformaldehyde followed by a permeabilization with 70% EtOH for at least 1h. Specimens prepared in this way can be stored up to four weeks at 4°C. For further preparation, models were taken out of the transwells and edges were trimmed using scalpel and tweezers. Then samples were washed for 5 min in 1ml PBST, followed by a wash with 500 μl 50% PBST and 50% 5x SSCT buffer and a further wash with 500 μl SSCT for 5 min each. All washing steps were performed on ice.

Detection stage

300 μl of the probe hybridization buffer was pre-hybridized to the sample for 30 min at 37°C. Meanwhile the probe solution was prepared by adding 5 pmol (10 μl of a 1 μM Stock) of the probe set pool to 500 μl probe hybridization buffer at 37°C. Then the pre-hybridization buffer was replaced by the probe solution followed by an overnight incubation at 37°C. To remove

the probes the sample was washed three times with 300 µl probe wash buffer, each time for 10 min at 37°C. In a final washing step, samples were treated twice with 5x SSCT (500 µl) for 5 min at room temperature.

Amplification stage

Pre-amplification was carried out by adding 400 µl amplification buffer for 30 min at room temperature. 30 pmol hairpin h1 and 30 pmol hairpin h2 were separately prepared by snap cooling 3 µl of each stock. This was performed by heating the hairpins to 95°C for 90 seconds and a subsequent cool down phase for 30 min in the dark. The hairpin solution was then prepared by adding hairpin h1 and h2 to 250 µl amplification buffer at room temperature. The pre-amplification solution was furthermore exchanged by the hairpin solution before the sample was incubated for 16h at room temperature in the dark. Exceeding hairpins get rinsed off with 1 ml 5x SSCT followed by three washes for 10 min each with 5x SSCT at room temperature.

Antibody staining

First, samples were blocked for mouse primary antibodies by incubating in blocking solution (5% donkey serum and 1:400 diluted unlabeled affinity purified Fab fragment donkey anti-mouse IgG (H+L) in TBS (Thermo Scientific)) for 1h. After three washes with TBS for 10min each, samples were incubated with primary mouse antibody (diluted 1:1000 in TBS) for 2h. Again, the samples were washed three times in TBS for 10min each before they were incubated with the secondary antibody AF647 (diluted in 1:1000 in TBS) for 1h. For nuclei staining, DAPI dye was diluted to 2 µg/ml in 5x TBS before 200 µl of the dilution were added to each sample for 45min. After another washing procedure as described above, samples were transferred to a glass slide and 50µl mounting medium was added and dried for 24h at RT. Storage temperature was 4°C.

Buffers, hairpins and oligos

Buffers (whole mount reagents) and hairpins (B3, Alexa 594, and B4, Alexa 647) were ordered from Molecular Instruments (USA, Los Angeles). Oligo nucleotides were ordered from IDT (Integrated DNA Technologies), standard desalted; diluted to 1 µM and pooled.

Probe design:

For probe design, a NCBI mRNA sequence (NM_006418.5) was used. Non-overlapping probes of 25nt sequences, separated by a 2nt Spacer (S) were selected based on GC content (35-65%).

Highly homologous mRNAs to target mRNAs were excluded. Additionally, a BLAST search against the human transcriptome ensured the specificity of hybridization probes.

Table 20: Probe design HCR-FISH for *OLFM4*

	1st Half of Initiator	S	Probe sequence (25nt)		Probe sequence (25nt)	S	2nd Half of Initiator
1	gTCCCTgCCTCT ATATCT	TT	tctataatactccaacagtc tccca	2	tagcaaatcatccagtgtgt tgtac	TT	CCACTCAACTTTAAC CCg
3	gTCCCTgCCTCT ATATCT	TT	gtgttgtacatgttgacgtac atgt	4	gtcaggttaactctggcaat attcc	TT	CCACTCAACTTTAAC CCg
5	gTCCCTgCCTCT ATATCT	TT	acatggatgaggactagtc attggg	6	tccaggcatggaagaatt agtggtg	TT	CCACTCAACTTTAAC CCg
9	gTCCCTgCCTCT ATATCT	TT	gtggtagaactatgcacct aaacat	8	tcaagacaaatgtcctaga tctcta	TT	CCACTCAACTTTAAC CCg
11	gTCCCTgCCTCT ATATCT	TT	ggtttccaactactgcactg attaa	10	acagaagcacatcacata caccagc	TT	CCACTCAACTTTAAC CCg
13	gTCCCTgCCTCT ATATCT	TT	ttatactgcttggtataccaa gtgt	12	accatgaaggcgttagaa gcagatg	TT	CCACTCAACTTTAAC CCg
15	gTCCCTgCCTCT ATATCT	TT	tgtgctaacgacactgagtt tgaga	14	gattacgacggatattattg gcaaa	TT	CCACTCAACTTTAAC CCg
17	gTCCCTgCCTCT ATATCT	TT	atctctagatcctgtaaaca gaact	16	acctctcaagagaaccctt agtaga	TT	CCACTCAACTTTAAC CCg

2.2.7 Bioinformatic analysis

Raw sequencing data were converted to FASTQ format, de-multiplexed and quality controlled using the Cell Ranger (https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/what-is-cell-ranger) version 3.0.2 pipeline mkfastq. Reads were aligned to the GRCH38 human genome assembly, filtered, UMI & barcode counted using the cellranger count pipeline.

The count matrix, barcodes and features were loaded into R (version 4.0.3) using the packages Matrix (version 1.3.4)and readr (version 2.0.1). Hashtag counts were separated from transcript counts and stored in different assays in SeuratObject (version 4.0.2) Hashtag count thresholds were set manually for Hashtag 1 (20) and Hashtag 2 (30). Transcript counts were normalized, scaled and highly variable genes identified using the Seurat (version 4.0.3) functions "NormalizeData", "ScaleData" and "FindVariableFeatures" with default settings. Dimensional reduction and clustering were performed using "RunPCA" and "FindNeighbors" (dims = 1:25) from Seurat, umap (X = [, 1:25]) from the uwot (version 0.1.10) package and leiden (resolution_parameter = 0.5, seed = 1992) from the leiden (version

0.3.9) package. Leiden clusters indicative of low quality (2,3,4,7) were removed. The functions "NormalizeData", "ScaleData", "FindVariableFeatures" (nfeatures = 5000), "RunPCA" (npcs = 40), "FindNeighbors" (dims = 1:40), umap(X = SeuratObject@reductionspcacell.embeddings[,1:40]) and leiden(resolution_parameter = 0.9) were re-run based on the reduced matrix. Cluster 5 was subclustered using the Leiden algorithm (resolution_parameter = 0.5, seed = 1992) to reveal a cluster of stem cells (cluster 3) and M-like cells (cluster 4).

Differential gene expression tests were performed using the "FindAllMarkers" function from the Seurat package.

Data wrangling and visualization was performed in R using the packages dplyr (version 1.0.7), tidyr (version 1.1.3), ggplot2 (version 3.3.5) and pheatmap (version 1.0.12).

3. Results:

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

These results were published in *Imdahl et al. 2020 Nat. Microbiology*[82].

Aiming to resolve the transcriptomes of individual bacteria, we had to overcome some major challenges. First, a single bacterium only contains femtograms of RNA[246] which is more than 100-fold less than in eukaryotic cells and demands an extremely sensitive DNA synthesis protocol. Additionally, functional bacterial transcripts lack a poly(A) tail, making it inaccessible for most straight forward reverse transcription (RT) strategies which are designed to exclude ribosomal RNA (rRNA) from eukaryotic transcripts. Another aspect is that bacterial RNAs are labile by nature and have half-lives of only several minutes which requires a rapid and efficient permeabilization and lysis of the robust cell envelope and a subsequent RNA stabilization. While to most common eukaryotic scRNA-seq Protocols have the lower detection limit adjusted to five to ten copies per transcript in a cell[143], the average mRNA copy number in a single bacterium ranges in a region of 0.4 copies per cell[140].

To overcome these challenges we developed a generic workflow (Fig. 5a), using FACS to isolate single bacteria from culture with a subsequent cell lysis followed by a highly sensitive random-hexamer priming based scRNA-seq protocol[111] for cDNA generation and amplification.

The poly(A)-independent multiple annealing and dC-tailing based quantitative scRNA-seq (MATQ-seq)[111] protocol ranges on the top level of sensitivity using low temperatures for hybridization of RT primers to internal transcript regions, enabling the detection of also low abundant transcripts.

To benchmark our workflow, we used the model organism *Salmonella enterica* serovar Typhimurium (henceforth *Salmonella*), which is extensively studied and whose transcriptome is strongly annotated[172,247,248,249].

In our initial experiment we analysed *Salmonella* in three different growth conditions which are mimicking relevant stages during the infection process: (1) ‘late stationary phase’

constituting mostly resting cells; (2) 'anaerobic shock' induced by interruption of the oxygen supply via mineral oil overlay, representing the intestinal environment; and (3) 'NaCl shock' creating an osmotic stress by addition of NaCl to the Medium (Fig. 5b). Approaching the detection limit we isolated either pools of 10 bacteria (10-pooled) or single bacteria in individual wells via FACS. To validate the reliability and precision of our isolation method, we sorted single bacteria on a fresh agarose dish next to controls of 2-, 10- and 100-bacteria and let them grow over night at 37°C. The formation of single colonies proved an accuracy of >97% (Fig. 6).

After cultivation or shock-generation, cells were immediately treated with an RNA stabilization solution (RNAlater) to avoid transcriptional changes in the rapidly turning over bacterial RNA. Individual bacteria were then sorted in PCR plates containing freshly prepared lysis buffer (Fig. 5a).

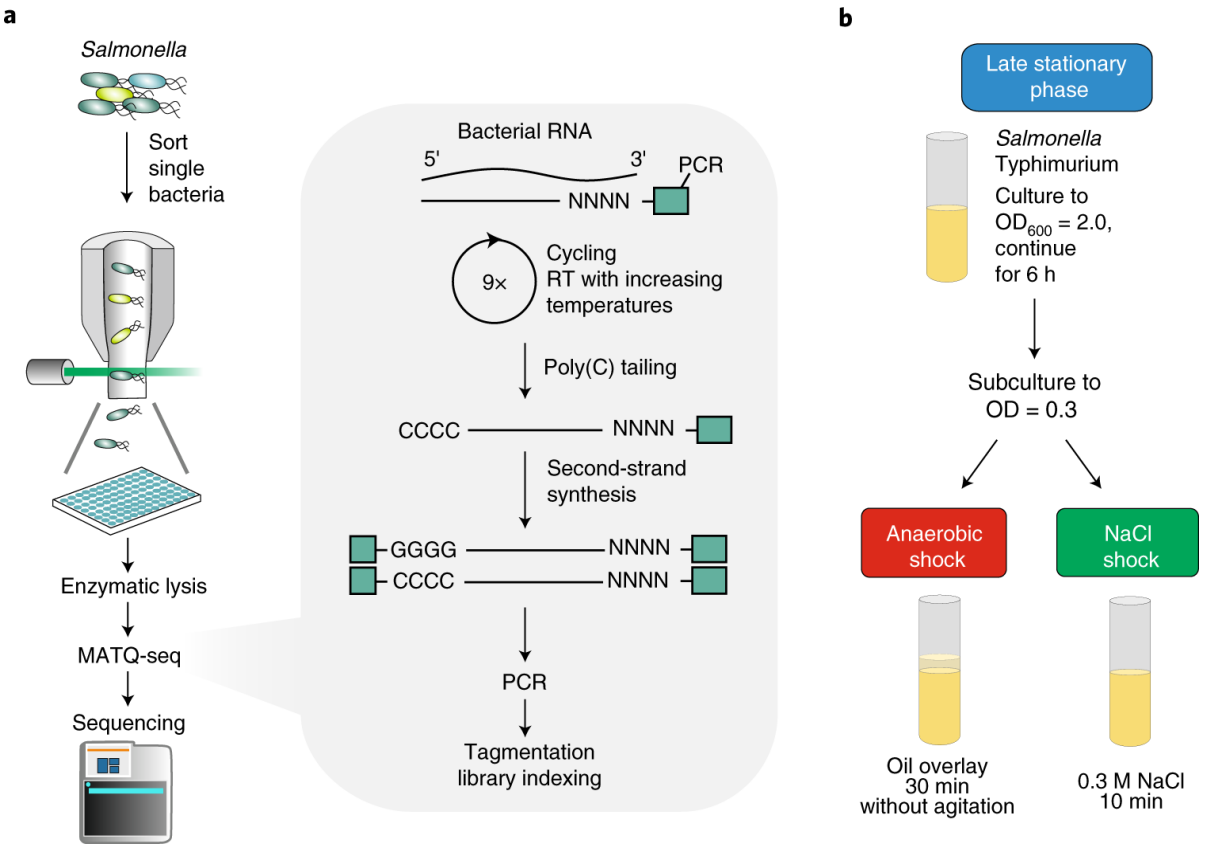

Fig. 5: Generic workflow for single-bacteria RNA sequencing **a**)Cultured *Salmonella* are FACS-sorted in individual wells containing lysis buffer before the MATQ-seq protocol was applied. Multiple cycles of ramping temperatures and the use of MALBAC-primers allow a sensitive and linear reverse transcription. Poly(C)-tailing and second-strand synthesis precede the PCR amplification. Amplified cDNA libraries were tagmented and indexed for illumina-sequencing. **b**) *Salmonella* cell culture and shock generation. Bacteria were cultured to an OD_{600} of two with continued growth for 6h for the late stationary phase. Subcultures were grown to an OD_{600} of 0.3 before shock generation. Anaerobic-shock bacteria were deprived from oxygen by covering the culture with

oil for 30 min. For the NaCl-shock, 0.3M NaCl were added for 10 min. Late stationary and shocked cells were mildly fixed with RNAlater immediately after treatment, to avoid further transcriptomic changes. Imdahl et al. 2020 Nature Microbiology[139].

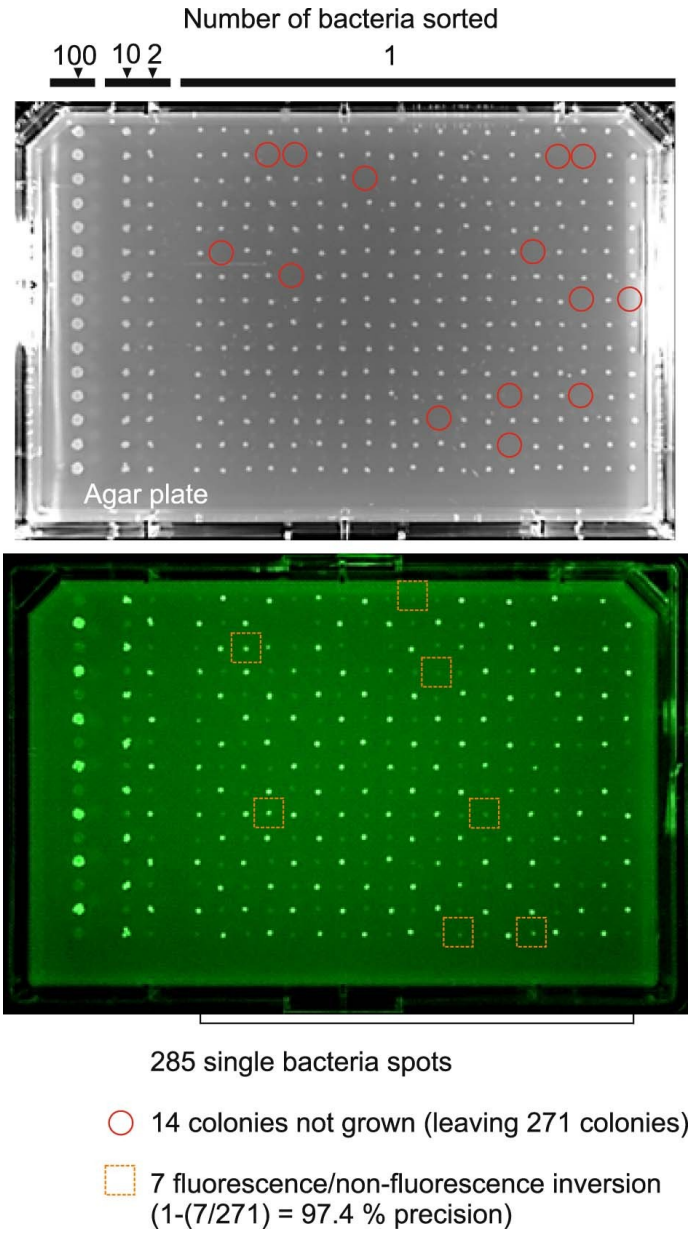

Fig. 6: Assessment of precision and efficiency of single bacteria sorting. LB agar was cast in a 96 well format dish and fluorescent (GFP-expressing) and non-fluorescent *Salmonella* were systematically sorted on top in alternate positions and differing population sizes (from left to right: 100, 10, 2 and single bacteria). After overnight growth, colonies were observed to determine the efficiency. Red circles indicate open positions where no colonies could be detected. Orange squares indicate sorting mismatches between fluorescent and non-fluorescent bacteria. An overall sorting precision of 97.4% could be demonstrated.

The thick peptidoglycan cell wall was digested using 5 U of Lysozyme, an amount that is capable of a sufficient lysis without harming the reverse transcriptase in the subsequent RT-step. We found out that amounts of >20 U are obstructing cDNA synthesis. Following lysis, reverse transcription was performed with several rounds of annealing (Fig. 5a). cDNA libraries were amplified using 23 PCR cycles, before they were indexed, pooled, and finally sequenced to a depth of 62.4±20.9 million reads.

3.1.1 Characterization of transcriptomes down to a single-bacterium level under different infection relevant growth conditions

In total 121 samples were analysed, containing 60 samples of 10-pooled bacteria libraries and 71 single bacteria libraries (Table 22(appendix)). We were able to consistently capture and identify all classes of RNA present in bacteria (Fig. 6a). Not using any rRNA depletion and therefore priming total RNA, the main share of reads was mapped to rRNA and transfer RNA transcripts, which accounted on average for 93% of all mapped reads while mRNA reads constituted on average for just 5% and small non-coding RNAs for 1.2%, with no major difference between pooled or single-cell libraries (Fig. 7, Table 23 and 24 (appendix)). Comparable relative proportions of RNA classes can be found in bulk RNA-seq of *Salmonella* without rRNA depletion[250] (Fig. 7a). For further analysis all reads mapped to rRNA and tRNA were removed, resulting in ~1,500,000 uniquely mapped reads for 10-pooled bacteria and 800,000 for single bacteria respectively (Fig. 8a) which is describe to reflect an adequate sequencing depth to assign physiological states of eukaryotic cells[117] and should therefore also be reasonable for bacteria.

Finally, we were able to detect 413± 237 and 170± 81 (average ± standard, with at least 5 reads per gene) genes in the 10-pooled and the respective single-cell libraries (Fig 7b). Strikingly, we could detect ~2.5-fold more genes in the bacteria treated with oxygen deprivation or NaCl addition compared to the late stationary phase (Fig. 7b and Fig. 8b). Determining a sufficient sequencing depth was applied, we performed a saturation analysis attesting that after 500,00 uniquely aligned reads per cell, the number of reads is marginally increasing (Fig. 7c).

Verifying that actual cDNA deriving from RNA was sequenced instead of genomic DNA we investigated the read distribution along the detected genes exemplary for ssrA (tRNA), fliC (flagellin) and different ribosomal protein genes, proving that only the respective genes

transcribed regions are covered by reads (Fig. 7d and Fig 11). As also shown in bulk RNA-seq of *Salmonella* we obtained an uneven read-coverage throughout the transcript[172].

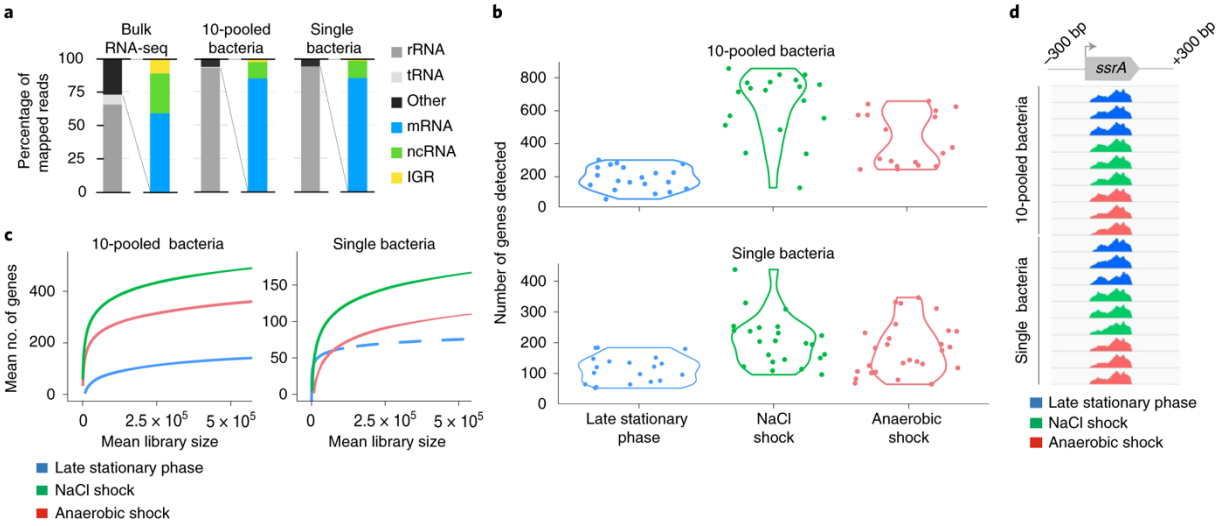

Fig. 7: Characterization of transcriptomes down to the single-cell level under different growth conditions **a)** Percentage of mapped reads to the transcripts of different RNA species in *Salmonella* bulk RNA-seq generated with random hexamers (data from ref.[250]) IGR, intergenic region; ncRNA, non-coding RNA; other, all other RNA classes (Tables 23 and 24 (appendix)). **b)** Violin plots depict the number of detected genes across the libraries. Single bacteria, late stationary phase, n = 19; NaCl shock, n = 23; anaerobic shock, n = 27; 10-pooled bacteria, late stationary phase, n = 20; NaCl shock, n = 19; anaerobic shock, n = 18. **c)** Number of genes detected per cell (counts >5) downsampling total read counts to the indicated depths. Dashed line for single bacteria (late stationary phase) data represents an extrapolated asymptotic fit. **d)** Gene body coverage to the reference sequence of the tmRNA-encoding gene *ssrA* across all conditions of 10-pooled and single bacteria. Imdahl et al. 2020 Nature Microbiology[139].

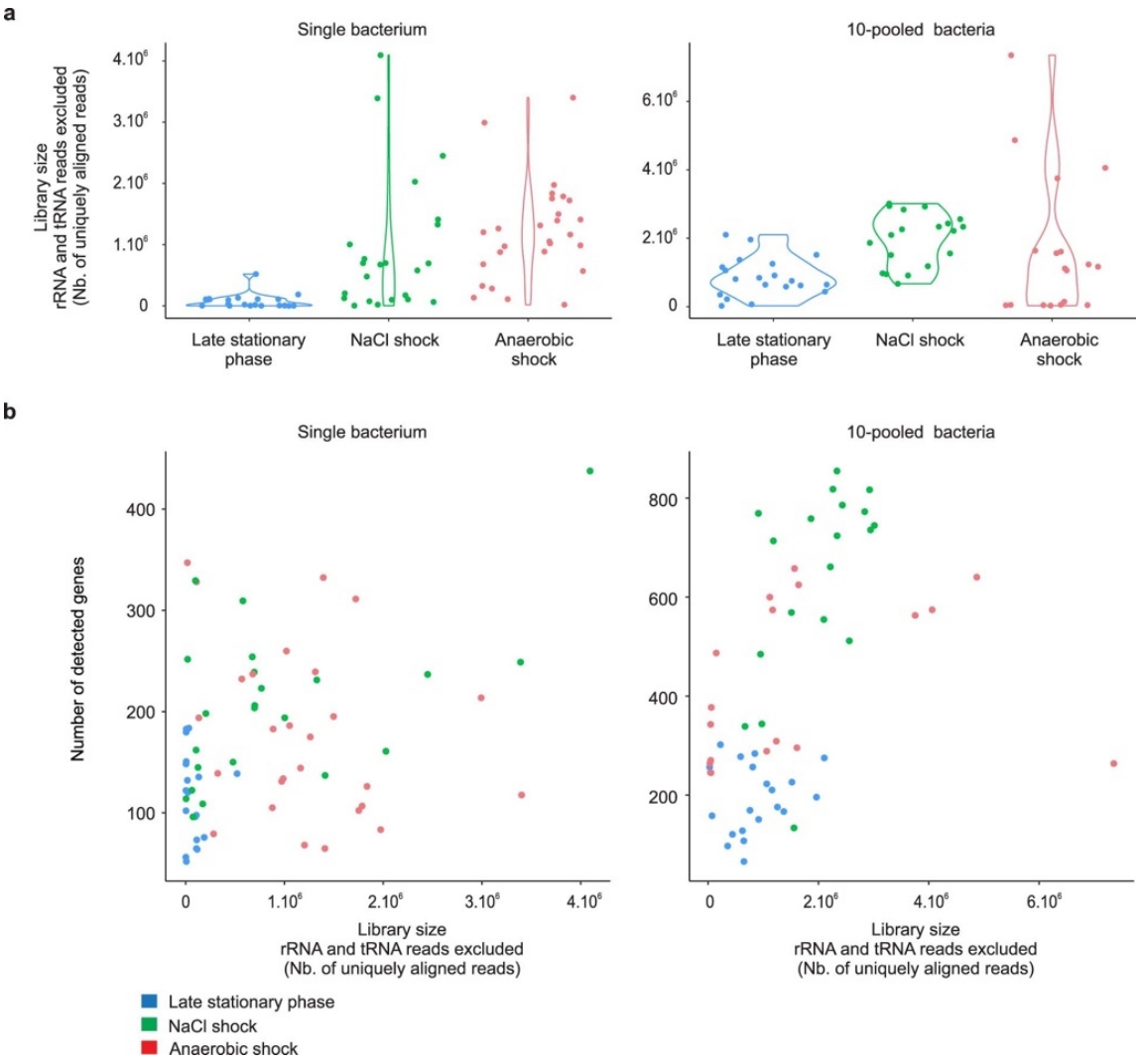

Fig. 8: Comparison of library size and number of detected genes. **a)** Violin plots depict the number of uniquely aligned genes (rRNA and tRNA reads excluded) in the different growth conditions for single bacteria (left) and 10-pooled bacteria (right). Scatter plots represent the relation of detected genes versus the number of uniquely aligned reads in 10-pooled or single bacteria and the respective growth conditions. Outliers were removed from analysis. Imdahl et. al 2020 Nature Microbiology[139]

Assessing the technical noise in our data we plotted the coefficient of variation for average read counts, demonstrating comparable profiles to those achieved using conventional eukaryotic single-cell RNA-seq[251] (Fig. 12). Spearman's correlations display the correlation between the matching conditions of ten-pooled and single bacteria with values of $\rho = 0.5$ between NaCl and anaerobic shock, $\rho = 0.42$ for late stationary phase. In comparison to bulk RNA-seq si andmilar values could be obtained (Fig. 12 a,b). As described before, on average, bacterial mRNA feature a copy number less than one per cell[140]. This, together with some dropouts could explain the substantial proportion of values close to zero in the scatter plots.

3.1.2 Single-bacterium RNA-seq reveals specific growth condition associated transcriptional signatures

To investigate whether there are differences in the growth condition specific transcriptomes, an unsupervised clustering via principal component analysis (PCA) of the 10-pooled or single-cell libraries was performed (Fig.9a and Fig. 10). Libraries (10-pooled or single bacteria) and their respective transcriptomes are represented as Individual dots in the PCA plots while the colours indicate the growth condition (Fig. 9a). A considerable aggregation of libraries from the same growth condition went alongside with a clear segregation form the other libraries indicating various genes expressed in the different clusters. To further investigate the differentially expressed genes across the libraries a DESeq2 analysis was performed, revealing 101 anaerobic and 274 NaCl shock-specific genes for the 10-pooled bacteria. The DESeq2 for the single bacteria libraries resulted in 63 anaerobic and 131 NaCl-shock genes (Fig. 9b and Tables 25 and 26 (appendix)).

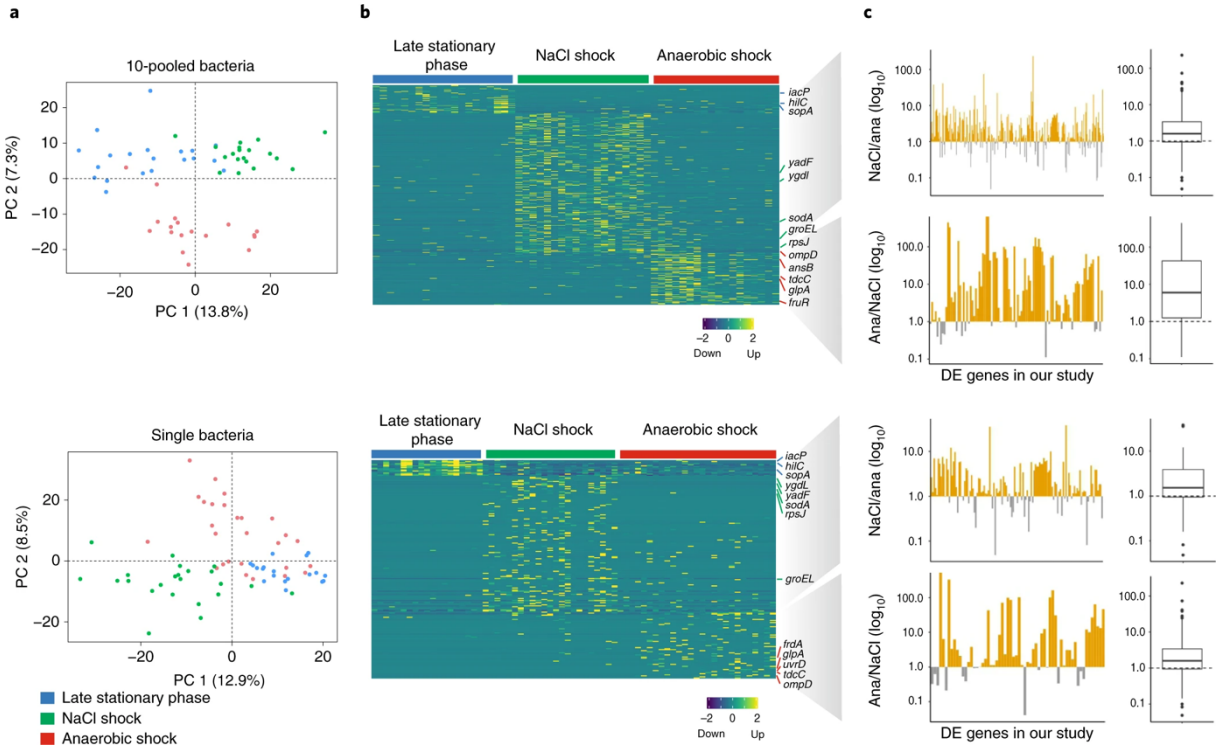

Fig. 9: Single-bacterium RNA-seq reveals specific transcriptional signatures associated with growth conditions **a)** Principal component analysis of the three growth conditions of 10-pooled or single bacteria libraries. Each dot represents either 10-pooled bacteria (upper panel) or single bacteria (lower panel). Colours mark the respective growth condition. **b)** Heatmaps show the differentially expressed genes for the specific growth conditions. DESeq2

could identify a total of 431 genes which were differentially expressed in 10-pooled libraries (upper panel) and 209 genes differentially expressed in the single-bacteria (lower panel) libraries (Tables 25 and 26). **c)** For each differentially expressed (DE) genes between anaerobic and NaCl shock the $\log_{10}$ transformed ratio of expression values independently measured under both shock conditions in a bulk RNA-seq benchmark study[172] (in transcripts per million) was calculated and displayed in histograms (yellow colour indicates a log transformed ratio >1 while grey bars are <1) and boxplots (median, first and third quartiles, lower and upper whiskers and outliers are shown). Imdahl et al. 2020 Nature Microbiology[139].

Among these genes, several were characteristic for metabolic shift under anaerobic shock condition like *glpA* and *tdcC*, as well as indicative genes for NaCl-shock like *yadF*, *ygdl*, or *sodA* (Fig. 9b). A gene ontology (GO) analysis of enriched biological processes of the differential expressed genes provides a more global representation of the metabolic state of individual cells in their respective condition. Typical signatures associated with mainly resting cells at late stationary phase could be observed while cells under anaerobic shock condition revealed a shift to anabolism and catabolism.

To benchmark our work we compared our single-cell results to previously published bulk RNA-seq data of the SalCom database[172]. For this purpose, we correlated our NaCl and anaerobic shock gene expression patterns to those of the database (Fig. 9c, Fig. 12, Table 25 and 26 (appendix)). Essentially, upregulated genes in NaCl or anaerobic shock condition in our experiment matched by ~75% to the respective conditions in SalCom (Fig. 9c) confirming the robustness of gene expression signatures achieved by single-bacteria RNA-seq. However, even if a comprehensive compendium such as SalCom is available, we recommend performing bulk RNA-seq in parallel to the single-cell libraries to validate the scRNA-seq data.

Establishing our bacterial scRNA-seq protocol for broader applications we sought to apply MATQ-seq to other species. Distinguished from *Salmonella* by its significantly higher GC content (>67 versus ~50), we decided to investigate the clinically relevant pathogen *Pseudomonas aeruginosa*. Following the approved workflow, we again generated 10-pooled and single *P. aeruginosa* libraries (Fig. 13). We were able to capture on average 102 genes on the single-bacteria level, which is in good agreement with what we obtained from single *Salmonella* as shown above.

Adopting the MATQ-seq protocol, we were able to capture low abundant transcripts of minimal microbial samples down to a single bacterium level and furthermore infer different growth conditions. By choosing very well characterized growth conditions and independent datasets, robust transcriptomic signatures could be assigned to individual bacteria.

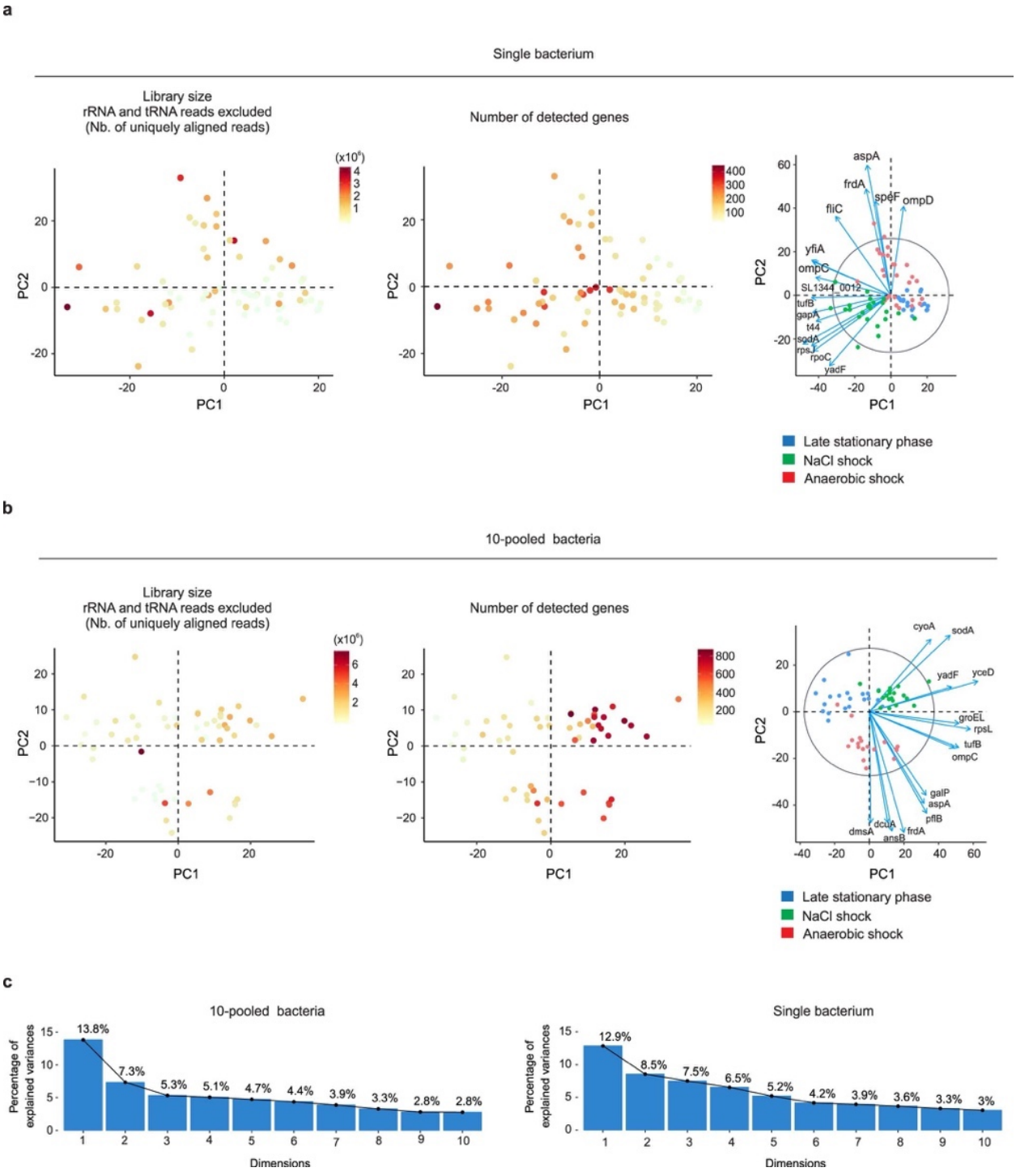

Fig. 10: Technical parameters associated to the Principal Component Analysis (PCA) of 10-pooled and single bacteria libraries. **a)** library size (left), number of detected genes (middle) are plotted on a PCA plot along with the top 15 genes (right) contributing to the principal components in single-cell libraries, and **b)** 10-pooled libraries. **c)** Scree plots display the variance associated with the PC dimension. Imdahl et al. 2020 Nature Microbiology[139].

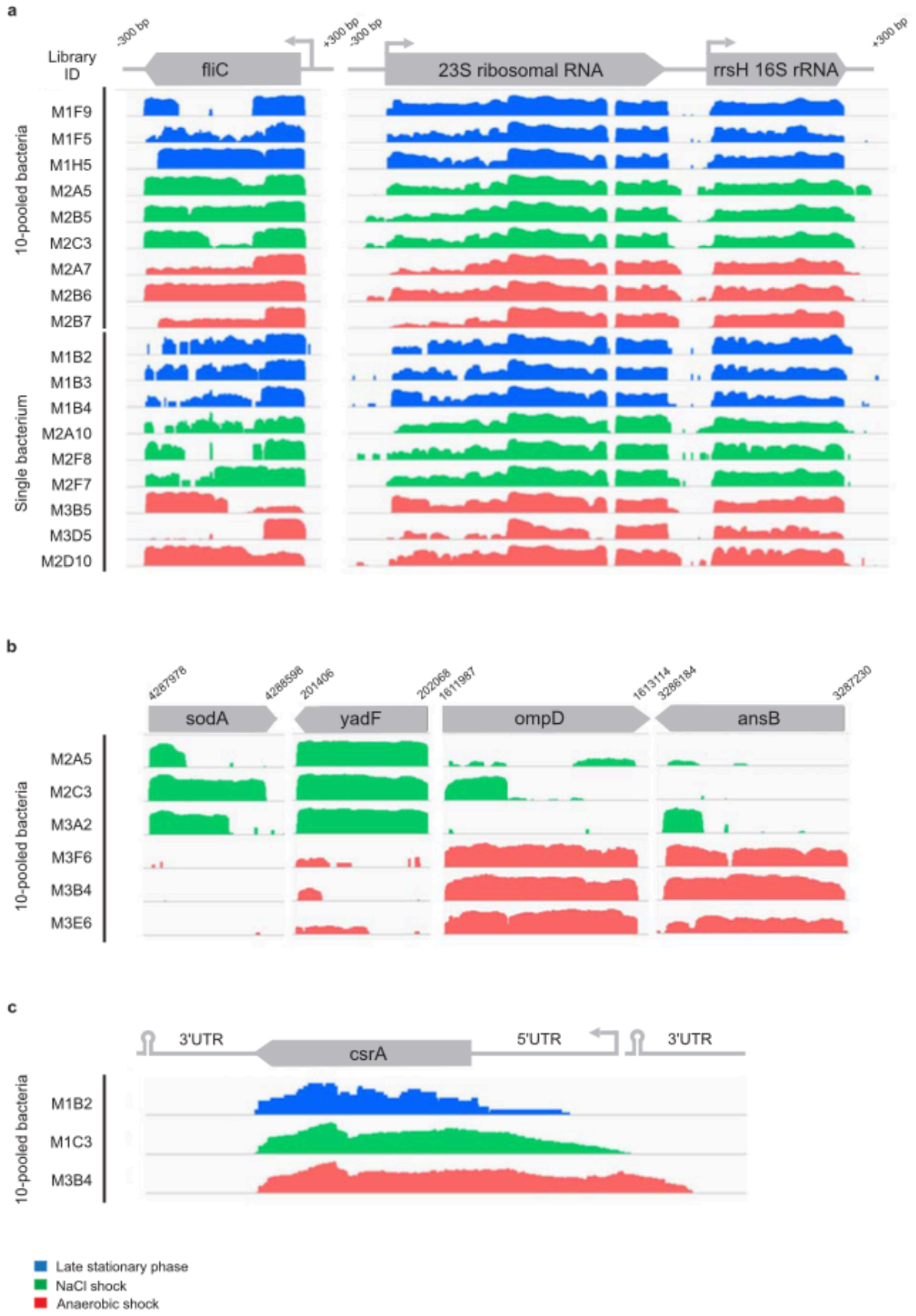

Fig. 11: Coverage plots displaying the density of reads for highly expressed or differentially expressed genes. **a)** Highly expressed structural genes in 10- and single-cell libraries (Colour indicates the growth condition). **b)** Read densities of selected differentially expressed genes in NaCl-shock vs. Anaerobic-shock condition observed in 10-

cell libraries. **c)** Gene body coverage of *csrA* in 10-pooled libraries and the different shock conditions overlapping to the 5' UTR. Libraries have been automatically log scaled by Integrative Genome Viewer. Imdahl et al 2020 Nature Microbiology[139].

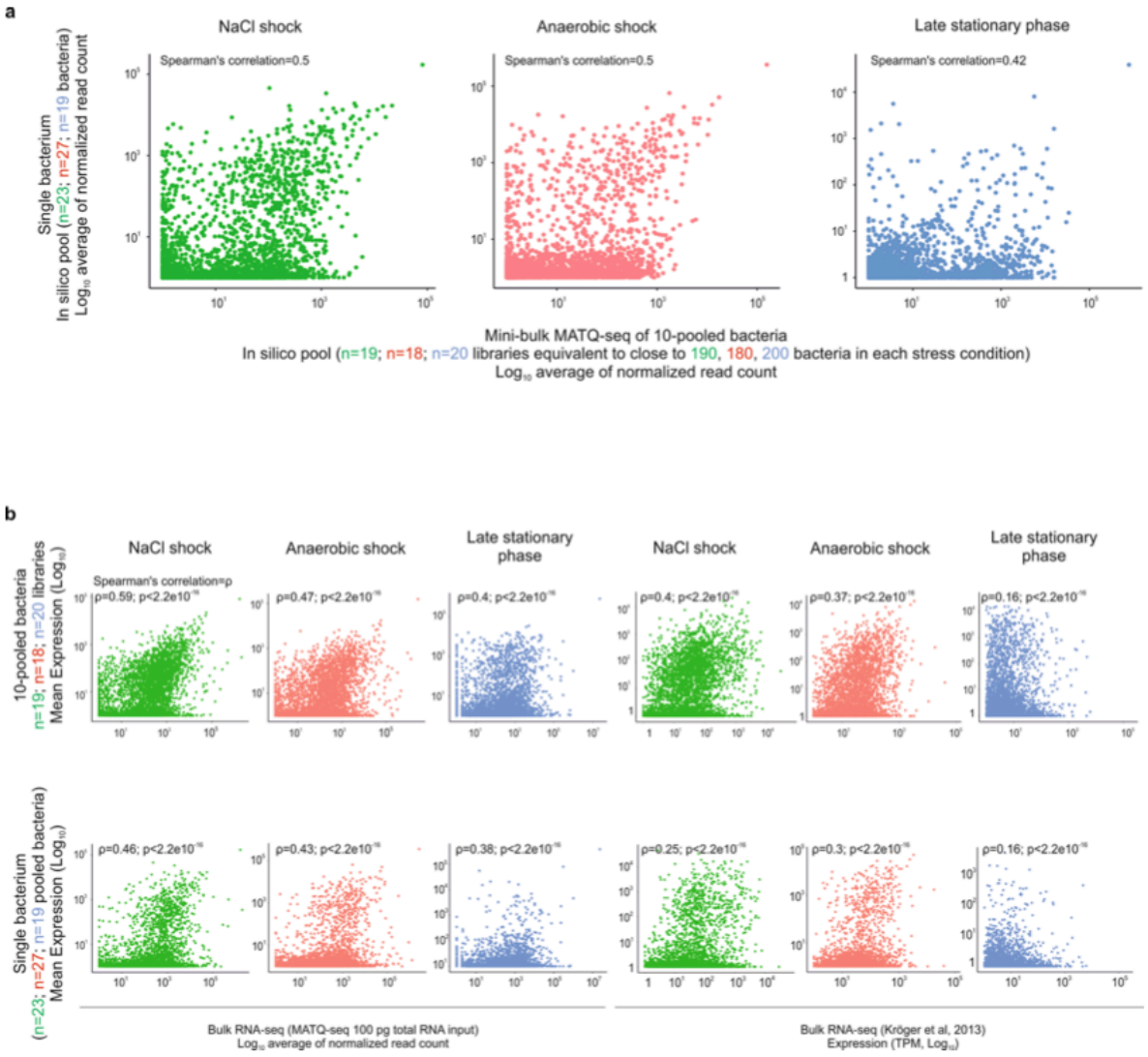

Fig. 12: Correlation between single-cell and bulk RNA-seq in the different growth conditions. **a)** Scatter plots display the correlation between single-cell libraries and in silico pooled 10-cell libraries in the respective conditions. **b)** Scatter plots depict the correlation between 10-pooled libraries (upper panel) and single-cell libraries (lower panel) of this study to two independent bulk RNA-seq experiments in the respective conditions with the associated Spearman's correlation coefficient and the associated p-values. Inhouse with MATQ-seq protocol (left panel) or Kröger et al. 2013[172] (right panel). Imdahl et al. 2020 Nature Microbiology[139]

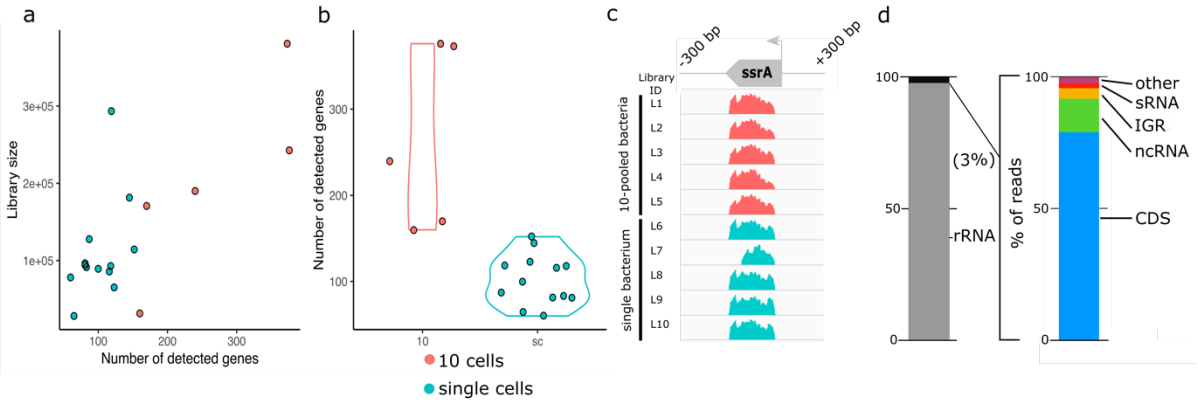

Fig. 13: MATQ-seq transcriptome capture of *Pseudomonas aeruginosa*. a) Library size plotted against the number of detected genes for single-bacteria libraries (blue) and 10-pooled bacteria (red) b) Violin plots depicting the number of detected genes per 10-pooled or single-bacteria library c) Coverage plot of uniquely aligned genes mapped to the ssrA gene d) Percentage of reads mapping to the RNA species. Imdahl et al Nature Microbiology[139].

3.2 Human intestinal tissue model reveals heterogeneity and dynamics of *Salmonella* infection

3.2.1 A primary human intestinal epithelial tissue model (hITM)

Aiming to create an in vitro model of human intestinal tissue a generic workflow was generated by Thomas Däullary of the chair of tissue engineering and regenerative medicine, Würzburg. In brief, starting with the explantation of crypts from a human donor, intestinal epithelial cells were expanded in an organoid culture before they were seeded to a decellularized porcine small intestinal submucosal (SIS) matrix and finally differentiated to a mature tissue model. On that basis a consolidated validation was performed which included histochemistry, immunofluorescence microscopy, raster electron microscopy and comprehensive transcriptome profiling via single-cell RNA sequencing (Fig. 14).

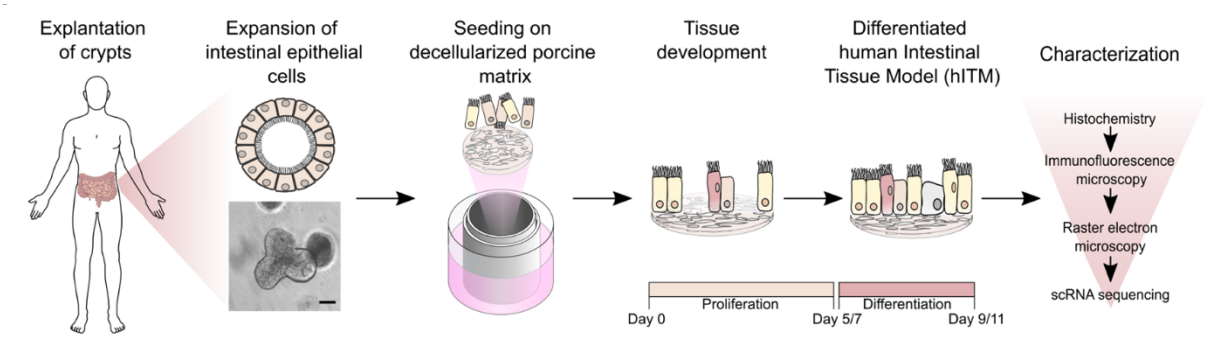

Fig. 14: Development of a human intestinal tissue model and applied characterization methods *(Däullary T. & Imdahl F. unpublished 2021)*

The microscopic analysis of the differentiated tissue model revealed a continuous monolayer of well polarized cells. Histochemical staining with Alcian blue makes the tissue organization visible, showing nuclei (dark pink), cytoplasm (pink) and mucins (blue) (Fig.15 *left*). In an immunofluorescence approach, Nuclei were further dyed with DAPI highlighting the concatenated structure of the layer (Fig.15). The epithelial membrane antigen MUC-1 lines the apical surface of epithelial cells in the intestine and other tissues, Vil-1 is playing a major role in microvilli formation which are also apically located, together staining against these proteins prove the correct orientation of the tissue model and the presence of indispensable cellular structures (Fig.15) [252]. Transmission electron microscopy (TEM) and raster electron microscopy (REM) allow a closer view on the tissue monolayer clearly showing cell boundaries and the

brush-like microvilli structure at the apical surface of the cells (Fig.15). Even tight junctions become visible (Fig.15 *white arrows*).

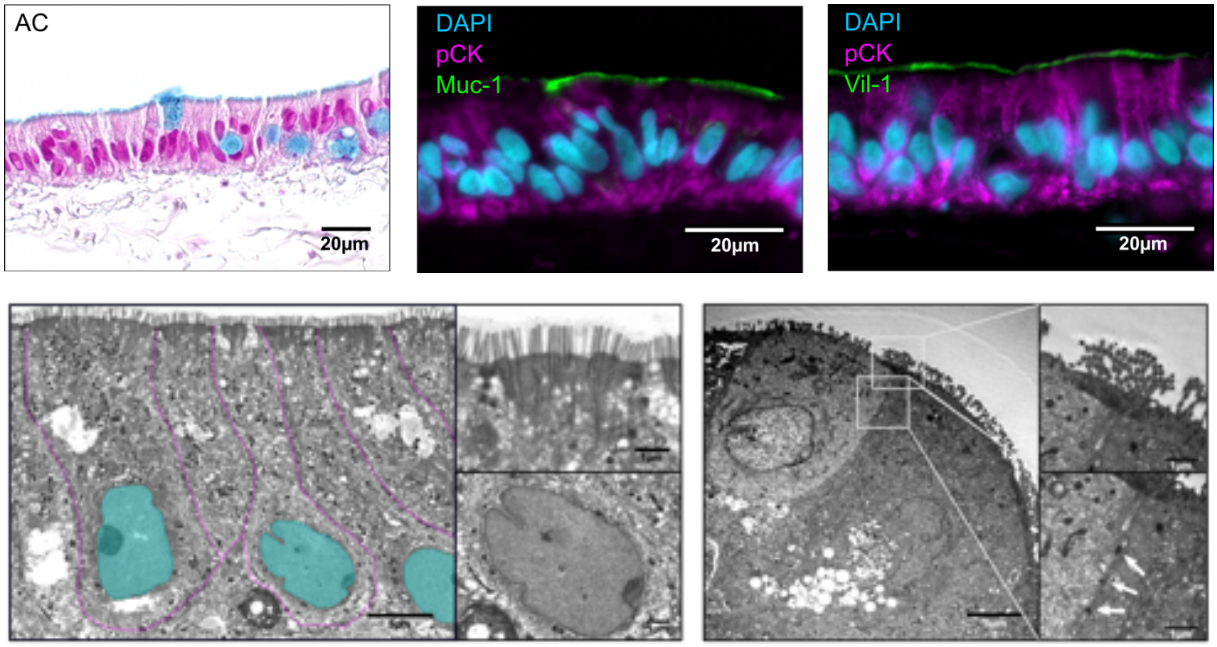

Fig. 15: Microscopic characterization of hITM **a)** Histochemistry (Alcian blue) and immunofluorescence microscopy of hITM staining for DAPI, Pan-cytokeratin (pCK), Mucin 1 (Muc-1), Villin 1(Vil-1) **b)** TEM and REM images of hITM. Purple lines depict the cell boundaries, cyan areas represent the nuclei and white arrows highlight the tight junctions. Imaging performed by Thomas Däullary (*Däullary T. & Imdahl F. unpublished 2021)*

Together fluorescence and electron microscopic imaging in combination with histochemistry could reveal a polarized monolayer containing tight-junctions, micro villi and differentiated intestinal epithelial markers, in good agreement to the real human intestine. To further investigate the cell composition of the tissue and to identify the entirety of characteristic cell-types, droplet-based single-cell RNA sequencing was performed.

For greater traceability, two replicates of fully differentiated hITMs were hash-tagged with total-seq antibodies after enzyme-based dissociation, subsequently pooled and loaded to the 10x controller (Fig. 16a) In total 11,727 cells were sequenced with a median of 2,424 genes detected per cell. Transcriptome clustering produced a unique profile allowing the annotation, of individual clusters to major intestinal cell types, based on distinct marker genes (Fig. 16b). A clear demarcation between proliferative and non-proliferative or differentiated cells could be observed in the clustering. The largest population was represented by progenitor cells, followed by immature enterocytes and secretory cells (Fig. 16c). Mature enterocytes and proliferating

cells showed roughly the same proportion of analyzed cells. Although in small quantities also stem-cells and M-like cells could be found in the tissue by scRNA-seq.

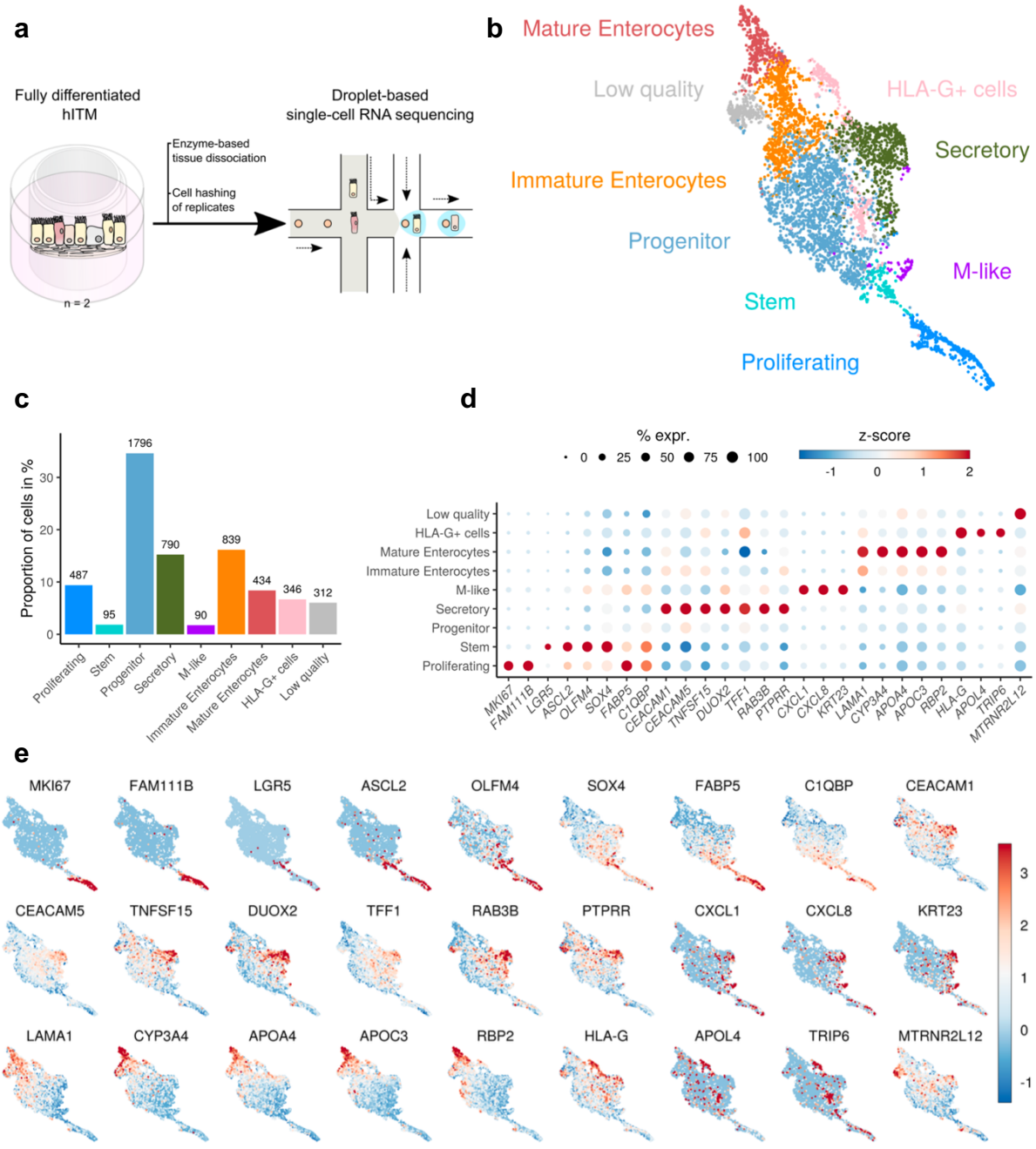

Fig. 16: Droplet-based scRNA-seq of hITM via 10x Genomics **a)** Experimental scheme of droplet-based single-cell RNA sequencing via 10x genomics with previous cell hashing. **b)** Leiden clustering and annotation of transcriptomic data **c)** Proportions of different cell types represented in the cluster **d)** Dot-plot depicting the enriched marker genes for the respective cell types (colour represents the z-score, dot size the rel. % of expression) **e)** Selected marker genes for each sub cluster plotted z-normalized to the embedding. Sc-analysis was performed by Oliver Dietrich. *(Däullary T. & Imdahl F. unpublished 2021)*

Furthermore, a cluster of HLA-G positive cells was observed besides the unavoidable cluster of low-quality cells. It can be stated that characteristic key cell types could be found via single cell transcriptomics. On closer inspection of the transcriptome clustering, a certain transition, starting from proliferating cells over stem and progenitor cells to immature and mature Enterocytes came into view (Fig. 16b). The annotation of clusters was performed based on marker genes known form literature which were plotted individually on the clustering (Fig.16e). Marker genes enriched in a specific cell type based on previously described genes and differentially expressed ones are depicted in a dot plot (Fig.16d).

Summarizing the results of the performed hITM validation experiments, we can state that the differentiated model well depicts the in vivo intestine featuring typical cell phenotypes regarding orientation, polarization, containment of microvilli and all major intestinal cell types.

3.2.2 hITM as an in vitro *Salmonella* infection model

As introductory discussed, there are several advantages of human in vitro test systems over cell-culture or mouse models. Most of the studies on which our knowledge of *Salmonella* infection is based come from animal models, primarily mouse models. Since *Salmonella* pathogenesis show fundamental deviations in mice compared to human, its reflection of the human infection is limited[253]. The opportunity to observe the infection process on a human intestinal tissue holds the potential to elucidate infection processes along with underlying mechanisms of host and pathogen.

To assess hITM as a valid instrument to study infections, a generally applicable workflow was developed. Either constitutively GFP expressing *Salmonella,* or the genetically engineered fluorescence-dilution strain were used as inoculum and added to the tissue covering medium. After synchronization via centrifugation cells were incubated and extracellular bacteria were killed through a gentamicin treatment. After several determined time-points the tissue was either fixed for microscopy/imaging or harvested and dissociated to a single-cell solution, which was subsequently analyzed with FACS (Fig. 17a). In a one cell per well approach infected cells were sorted, lysed, and forwarded to Smart-seq2 library preparation and Illumina sequencing.

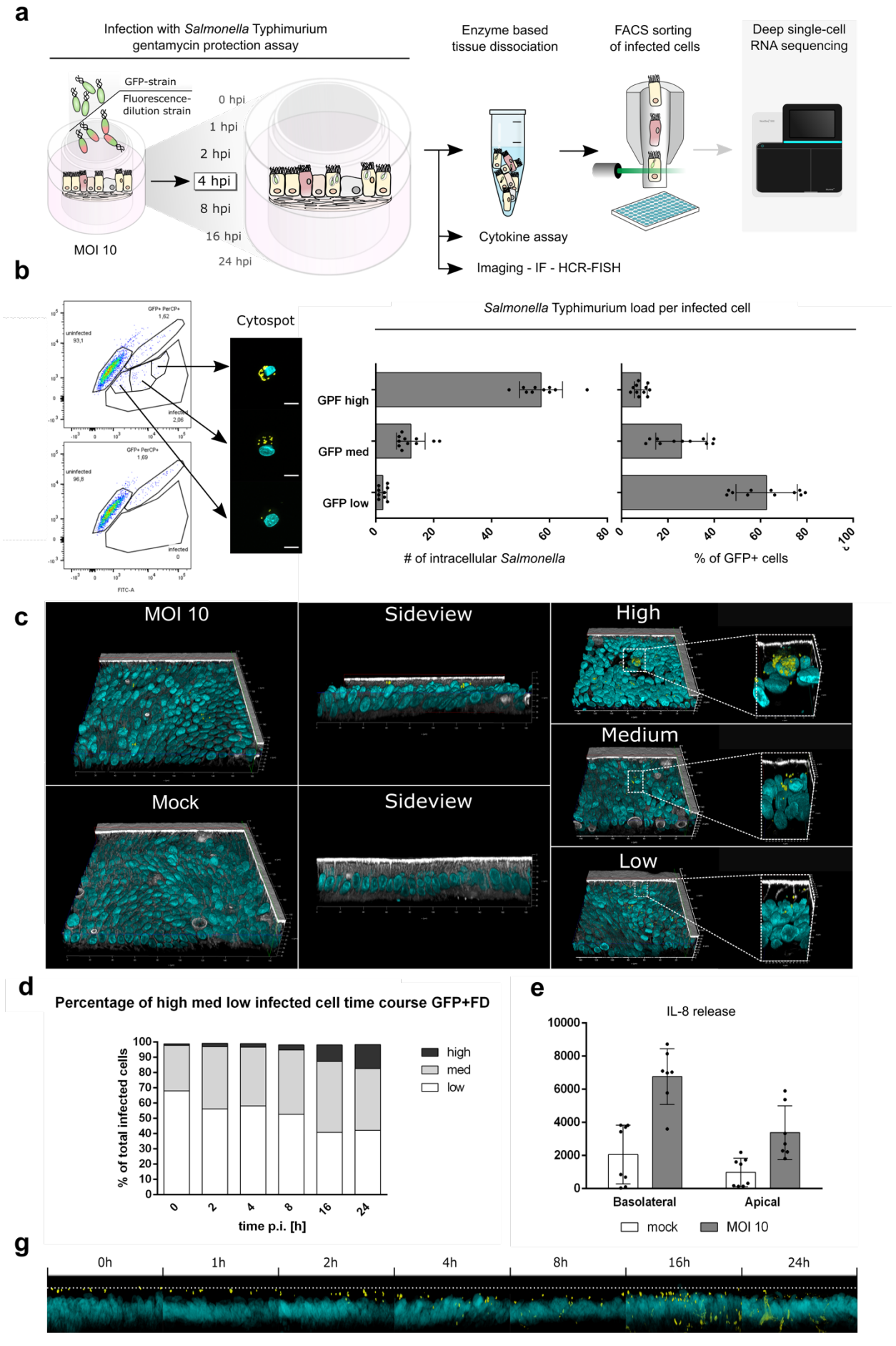
a
Infection with Salmonella Typhimurium
gentamycin protection assay
Enzyme based
tissue dissociation
FACS sorting
of infected cells
Deep single-cell
RNA sequencing
GFP-strain
Fluorescence-
dilution strain
0 hpi
1 hpi
2 hpi
4 hpi
8 hpi
16 hpi
24 hpi
MOI 10
Cytokine assay
Imaging - IF - HCR-FISH
b
Cytospot
Salmonella Typhimurium load per infected cell
GPF high
GFP med
GFP low
# of intracellular Salmonella
% of GFP+ cells
c
MOI 10
Sideview
High
Medium
Mock
Sideview
Low
d
Percentage of high med low infected cell time course GFP+FD
% of total infected cells
time p.i. [h]
high
med
low
e
IL-8 release
Basolateral
Apical
mock
MOI 10
g
0h
1h
2h
4h
8h
16h
24h

Fig. 17: *Salmonella* Infection of hITM **a)** Oerall workflow of hITM *Salmonella* infection at different time points and subsequent validation of the complete tissue via imaging and cytokine release measurements, of the dissociated tissue via FACS, and sequencing. **b**) FACS gating strategy to capture infected cells according to their infectious load (High, Medium and Low), represented by single-cell fluorescence microscopy (cyan = nuceli, yello = *Salmonella*), quantification displayed in bar plots. **c)** Fluorescenc microscopy of whole mount tissue model *Salmonella* infected (yellow) vs Mock (cyan = DAPI-staining of nuceli, grey = phaloidin staining of F-actin). Infectious load (High, Medium and Low) shown in the tissue. **d)** Distribution of cells with low (white), med (grey), and high (black) infectious load over time. **e)** IL-8 release measured in the apical and basolateral medium after 4h of infection vs. uninfected model **f)** Migration of *Salmonella* (yellow) through the hITM over time. Dashed line marks the apical cell membrane. Imaging was performed by Thomas Däullary. *(Däullary T. & Imdahl F. unpublished 2021)*

Analyzing the tissue cells which were infected with constitutively GFP expressing *Salmonella* and a MOI of 10 via FACS, revealed an average infection rate of 4-6%, over all time-points. Considering the GFP intensity, some cells were containing a higher bacterial load than others. On that basis three populations were defined, sorted and fixed. Individual cells of each population were subsquently fluorescence-microscopically analysed and the infectious load was determined (Fig. 17b). Showing an overall lower abundance (~10%) in the tissue, highly infected cells (high) carried on average close to 60 bacteria. Tissue cells in the medium population contaied between 8 and 20 *Salmonella* and show an intermediate abundance (~28%) of while the modestly infected 'low'-population had an ifectious load of 1-3 and an abundance of about 62% (Fig. 17b). To demonstrate that these phenomenons are not only confined to the single-cell solution, the full tissue model was comprehensively scanned for cells showing high, medium or low infectious loads via fluorescence microscopy (Fig. 17c). Observing the infection over a time course between 0h and 24h, reported an increasing number of highly infected cells, while tissue cells with a low and medium amount of intracellular *Salmonella* decreased over time (Fig. 17d). Applying fluorescent microscopic imaging to the infected tissue over time revealed a view on *Salmonella* migration. While partly located still in the membrane at the 0h time point a clear progression of migration from the apical to the basolateral side of the tissue could be observed over time. Strikinly in the late time points (16h, 24h) *Salmonella* induced filaments became visible (Fig. 17e). Interactions of *Salmonella* and the luminal intestinal surface trigger an accute inflammatory response of the epithelial cells wich is next to other proinflammatory molecules mainly mediated by interleukin-8 (IL-8)[254]. Aiming for an infection-model which is mimicking the in vivo tissue as good as possible, the cytokine release related to infection was measured 4h p.i. in the apical and basolateral medium (Fig. 17e). As

peviously described for *Salmonella* infected cells the IL-8 release occurs mainly basolaterally[254,255] mirroring what could be observed in the tissue model. Stringing together fluorescent microscopy images according to their timepoints nicely compiles the progression of infection and *Salmonella* migration up to the formation of *Salmonella* induced filaments (SIFs)(Fig. 17g).

Having investigated *S.* Typhimurium infection within the tissue model via fluorescent imaging, FACS anlysis and chemokine probing, it could be stated that hITM is a suitable and valid model to study human intestinal infection processes.

3.2.3 hITM infection time-course across different bacterial burdens

At six different time-points between 0h and 24h post infection, infected hITM cells have been systematically FACS sorted, according to their infectious burden, in a 96-well plate containing lysis buffer. Due to the different abundancies oft cells containing a low-, medium-, or high number of bacteria, the proportions variate. For each time point 30 cells with a low infectious load, 20 cells with a medium load, and 10 cells with a high load were sorted and processed with Smart-Seq2 (Fig. 18a). After the removal of 6 outliers, 354 cells were plotted in a UMAP which delineated five different clusters. Besides the inevitable accumulation of low-quality cells expressing mainly mitochondrial gene, the cluster of proliferative cells splits off the most (Fig. 18b). Allocating the cells according to the different timepoints already gave a hint on differentially expressed genes along progression of infection (Fig. 18c). A temporal direction from top to bottom could be inferred especially observing the late timepoints. The heatmap (Fig. 18d) confirmed the assumption of differential gene expression over time displaying distinct patterns for each timepoint. Associating the cells in the UMAP to their individual infectious load, the distribution looked random on the first sight (Fig. 18e). However, the different partition size of the processed cells must be considered. Meaningful, a striking large proportion of highly infected cells could be identified in the low-quality cluster, most probably marking dying cells. The corresponding heatmap divided by infection content, shows a differential gene expression, across the infectious condition. While medium and high share most highly expressed genes, the cells which contain a low infectious burden show a clearly different expression pattern (Fig. 18f). Selected marker genes projected on the embedding unveil specific cell types and their respective positions (Fig. 18g). The marker genes *PSRC1, FAM111B* and *MKI67* mainly tag the proliferative cluster where also *FABP5* is located.

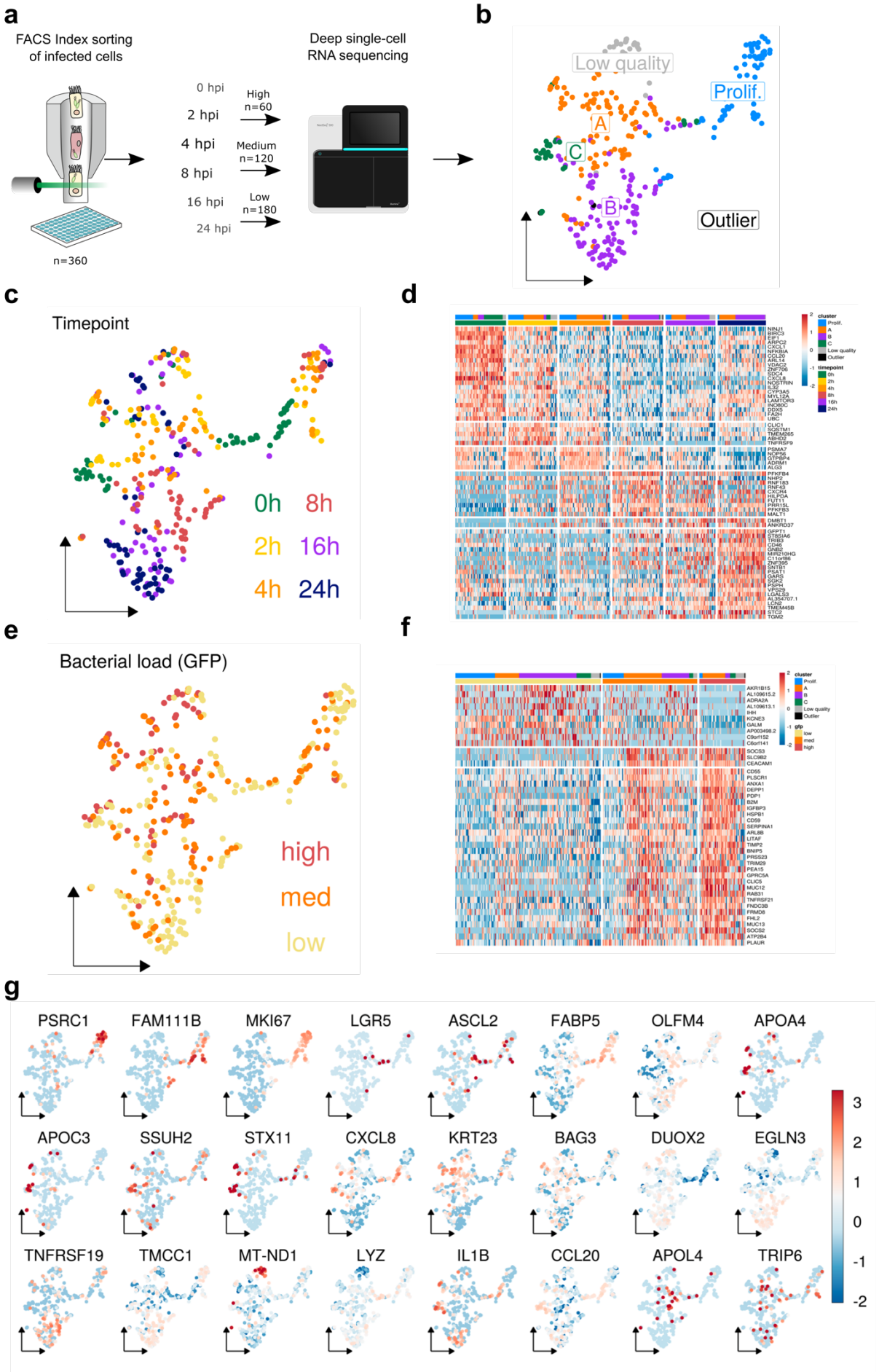
a
FACS Index sorting
of infected cells
n=360
0 hpi
2 hpi
4 hpi
8 hpi
16 hpi
24 hpi
High
n=60
Medium
n=120
Low
n=180
Deep single-cell
RNA sequencing
b
Low quality
Prolif.
A
C
B
Outlier
c
Timepoint
0h
8h
2h
16h
4h
24h
d
e
Bacterial load (GFP)
high
med
low
f
g
PSRC1
FAM111B
MKI67
LGR5
ASCL2
FABP5
OLFM4
APOA4
APOC3
SSUH2
STX11
CXCL8
KRT23
BAG3
DUOX2
EGLN3
TNFRSF19
TMCC1
MT-ND1
LYZ
IL1B
CCL20
APOL4
TRIP6
3
2
1
0
-1
-2

Fig. 18: HITM infection time course **a)** Infected cells were sorted according to their bacterial burden at six different time points between 0 and 24h post infection. **b)** UMAP from 15 PCs of log-counts from 5000 genes selected by variance with ERCCs – unsupervised clustering. **c)** UMAP as is (b) highlighting the different time-points of infection. **d)** Heatmap display the differential gene expression across the time course. **e)** UMAP as in (b) depicting the different infectious loads; high (red) medium (orange) and low (yellow). **f)** Heatmap shows the differential gene expression according to the bacterial load of the cells. **g)** Expression delineation of selected marker genes. All plots and heatmaps were generated by Oliver Dietrich. *(Däullary T. & Imdahl F. unpublished 2021)*

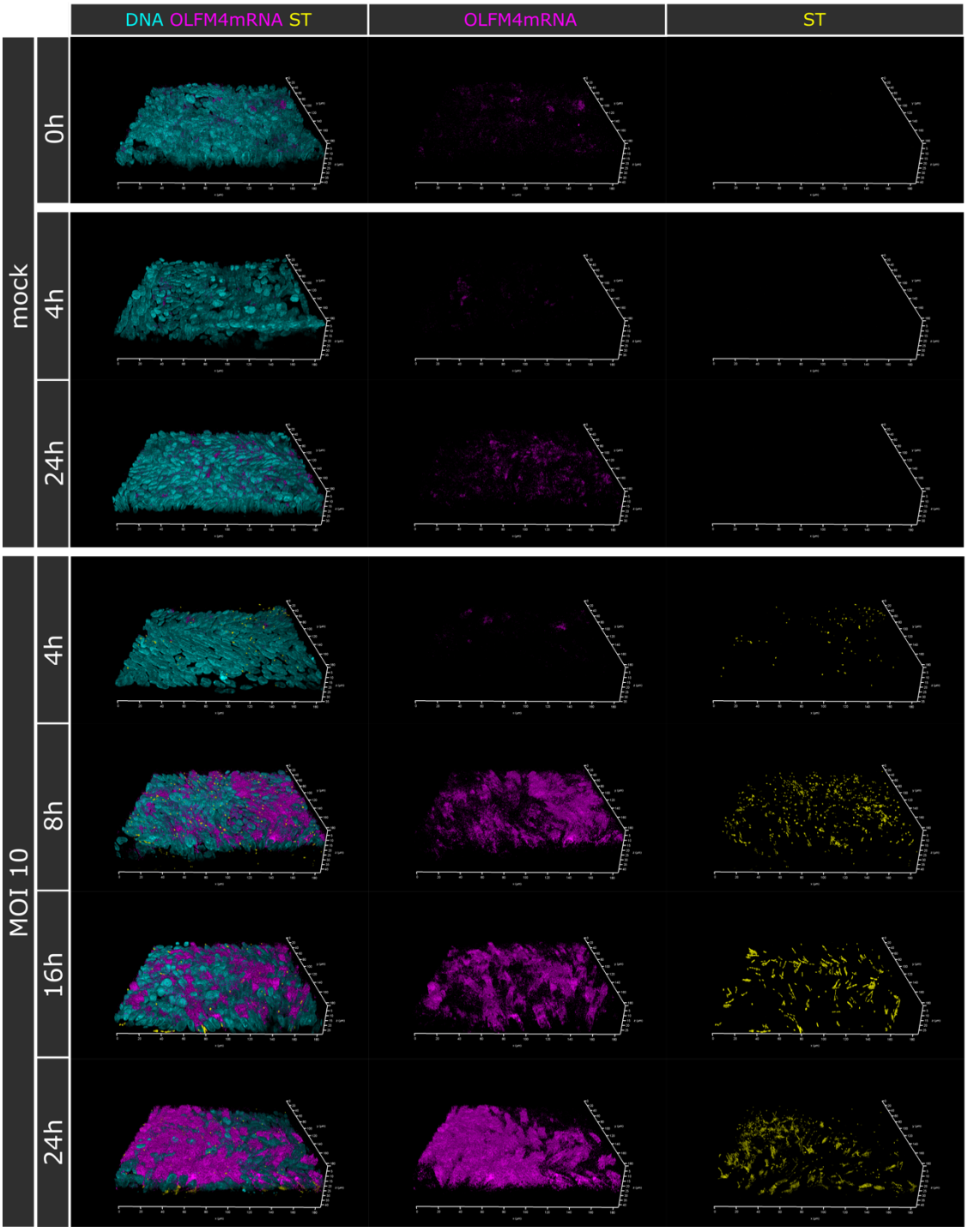

Fig. 19: HCR-FISH Imaging of *Salmonella* infected hITM with focus on *OLFM4* expression over time.

The mitochondrial gene *MT-ND1* is clearly enriched in the low-quality cluster. *APOA4* and *APOC3* which are associated with mature enterocytes, accumulate in cluster (C). Secretory cells are represented by high *DUOX2* expression. M-like cells are indicated by the presence of *CXCL8* or *KRT23* but didn't show a contiguous cluster. The most robust stem cell marker in the small intestine might be *LGR5,* which is complemented by *ASCL2*. Located at the transition from the proliferative cluster to cluster A. With the attempt to unite the different dimensions of this experiment and a special focus on stem cells, one can state that compared to their general abundance a clear accumulation of *LGR5*$^{+}$cells could be seen at the 0h timepoint. Additionally, most of those cells do not feature any proliferation markers such as *MKI67* and show just a low infectious load. In the later timepoints it seems as if the stem-cell population vanishes. *OlFM4*, in some studies also used as a stem cell marker, does not cluster to the other stem cells, but seems to play a role during infection[256]. Due to the lack of adequate antibodies for immunofluorescence microscopy we sought for another method. Hybridization chain reaction in combination with quantitative and high-resolution RNA fluorescence in situ hybridization provides clear evidence of a gene being expressed. Applying this technique equipped with probes for the *OLFM4* gene to our infected tissue model revealed a dramatic increase of expression over time and infection compared to the uninfected model (Fig.19).

4. Discussion:

4.1 Bacterial RNA-seq at a single-cell resolution

In our study we managed to achieve a transcriptomic single-cell resolution of bacteria under different infection relevant conditions. With the choice of the pertinent and well-studied pathogen *Salmonella* Typhimurium, we established a strong background for the validation of our study. The applied stress conditions, showing highly divergent gene expression, were selected according to the work of Kröger et al., 2013 which at the same time served as benchmark study[172]. An additional internal bulk RNA-seq reference could equally confirm the correctness of our work. Several steps in the workflow are crucial to achieve access to the genuine transcriptome of a bacterium. Initially important is the first processing of the bacteria whose specific feature is the particularly short transcriptomic half-life[257]. Using a mild fixation with 1:1 PBS diluted RNA*later* stabilization solution we try to extend the shelf life of the bacterial RNA integrity, knowing that the pure solution is heavily affecting the FACS profiles of the analysed samples. A second essential step is the bacterial isolation which must be precise and robust, as the first quality check comes late in the further protocol. Incorrectly sorted bacteria can dry out on the walls of reaction tubes or bias the results. Extensive testing and pedantic calibration of the FACS allowed us sorting bacteria confidently (Fig. 8). Most difficult and not yet at the end of research is the lysis buffer, which at the same time is urged to sufficiently lyse the thick bacterial cell wall, but not to harm the precious RNA content. We found that a too high concentration of lysozyme is interfering with the reverse transcriptase in the later protocol. The biggest issue, however, is the lack of generalizability. As the buffer is tailored to a specific bacterium, establishment work is required when applying it to other species. This effect is of course compounded by the different cell wall compositions of Gram-negative and Gram-positive bacteria. Among the other recently published methods for single-bacteria RNA sequencing, MATQ-seq is the only well-based method and convinces at the same time with a high gene detection sensitivity[83,84]. This unique feature underlined by the comparatively low drop-out rate allows especially research with rare populations. The split-pool-barcoding-based methods are working with stochasticity and therefore need a very high number of bacteria. The loss of cells over the procedure, however, is immense (up to 90%). Nevertheless, the massive parallel processing reduces costs and the ability of treating both, Gram-negatives, and Gram-positives simultaneously, holds great potential for microbiome

research. A general hurdle when conducting scRNA-seq in bacteria is the ribosomal RNA which accounts for more than 95% of the transcript. While eukaryotic RNA-seq protocols circumvent rRNA by selectively amplifying only polyadenylated mRNA transcripts via poly(A)-capture, bacterial mRNA is missing this feature. Kuchina et al., therefore in their method, added a poly-A tail to every transcript and depleted rRNA via 5S, 16S and 23S blocking primers according to the method of Wangsanuwat et al., 2020[84,258]. Since a rRNA depletion would have a massive reductive effect on the sequencing costs, an integration into the MATQ-seq protocol would be extremely beneficial. As its not intended to change the initial nature of the protocol, the depletion would have to be carried out towards the end of the library construction. Depletion of Abundant Sequences by Hybridization (DASH) utilizes the endonuclease Cas9 to specifically target rRNA transcripts and represents a promising approach applicable to our existing protocol[259,260].

4.2 Outlook and future challenges of bacterial scRNA-seq

ScRNA-seq has revolutionized our understanding of cellular functions and heterogeneity in many research areas and has rightly developed to one of the core technologies of our time. The logical continuation is the extension of the application towards microbes. Unlike for eukaryotic microbes, for bacterial scRNA-seq the kick-off has only recently occurred. Already a few approaches have achieved the goal to successfully uncovering the transcriptome of single bacteria and observing their heterogeneity[82–84]. However, all these methods are quite laborious and require very skilled staff, not to mention the excessive costs compared to eukaryotes. A robust, reproduceable, and easy-to-use high-throughput method integrating microfluidic platforms such as 10X Chromium, would be a great benefit to the field and is currently being investigated by several research groups.

Furthermore, in context of deeper understanding of host-pathogen interactions, it would be of great interest to look at the transcriptome of the invading bacterium and the referring host cell simultaneously. A cornerstone for this was already laid in 2012 by Westermann et al. with the invention of dual RNA-seq[165]. However, dynamics and heterogeneity of infection remain hidden behind the averaging bulk approach. But this very detail, might represent the missing piece of puzzle in order to investigate pathogenic infection- or host defence mechanisms.

4.3 Development of a validation pipeline for the human intestinal tissue model

To comprehensively characterize and validate the human intestinal tissue model, a genuine workflow was developed. With increasing detail, starting from brightfield microscopy over immunohistochemistry and electron microscopy up to scRNA-seq, general and specific features of the model were elucidated. In order to come as close as possible to the objective of recreating the native tissue *in vitro*, several key characteristics of the human small intestine must be fulfilled. Tissue orientation, polarization and membrane integrity are obvious features that should be complied with, and which can be easily observed microscopically. Electron microscopy can reveal details such as microvilli or tight junctions. Selected cell types and the typical mucus layer can be proven by antibody staining and immunofluorescence microscopy. The cellular composition, including all present cell types and their respective abundance in the tissue, can be revealed by high-throughput scRNA-seq. Applying this combination of methods to the human intestinal tissue model, disclosed a functional, polarized tissue model, containing major cell types of the *in vivo* small intestine.

The transcriptomic footprints of some cell types, however, were not detectable in our scRNA-seq data. Besides the generally low occurrence of Paneth and Goblet cells compared to other cell types in the small intestine, the absence of these cell types may have multifactorial reasons[261,262]. Located at the bottom of intestinal crypts Paneth cells are specialized in the secretion of antimicrobials like lysozyme, alpha-defensines or phospholipase-A2 and play an important role for the homeostasis of intestinal epithelial cells[204,263,264]. While well studied in murine organoids (Johan van Es), it remains challenging to sustain Paneth cell populations in human intestinal organoids[261,265]. One reason could be active Notch signalling which regulates the intestinal stem cell homeostasis and regeneration. Thus, it prevents cell differentiation and maintains the stem cell integrity via induction of downstream target genes, such as Olfactomedin 4 (*Olfm4*)[266–268]. In addition to the protein gradients that are necessary for the differentiation of special cell types, recent studies show that the geometry and mechanical properties of the crypts also play a decisive role for the differentiation of Paneth and Goblet cells[234,269].

Nevertheless, other cell types and their respective abundances are in good agreement with the native small intestine *in vivo*.

Overall, it can be stated that the developed characterization pipeline illuminates the cellular facets of the engineered tissue and thus can confirm a reproducible and comprehensive human

tissue model mimicking the characteristics of the real tissue. This sophisticated workflow has a strong validity and could act as a blueprint for the characterization of other tissue models.

4.4 Human intestinal infection model reveals new insights into *Salmonella* infection

The great benefit of a human tissue model is its applicability towards infection. As mentioned above, the widely used animal models do not necessarily reflect the situation in humans, due to a different physiology on the host side or a varying pathogenicity on the pathogen site. The here described hITM mirrors the human small intestinal physiology and therefore allows conclusions about the infection process, without having ethical restrictions. Using the extensively studied Gram-negative, facultative intracellular, pathogen *Salmonella* Typhimurium, we created a strong background for benchmarking. However, despite replicating the previously described infection pathways of *Salmonella* in hITM, we were also able to make observations that deviate from the general scientific opinion. Thus it is known that *Salmonellae* invades epithelial cells, M cells, macrophages or dendritic cells, of which M cells are described as the favoured traverse through the epithelial barrier[270]. When reviewing the infected cells in our time course experiment, also other cell types also come to light. Most strikingly, stem cells are a preferred destination of invading *Salmonella* compared to their overall abundance. This can mainly be observed in the early time points while in later timepoints the stem cell population seem to vanish. This could hint to a potential *Salmonella* induced remodelling of the transcriptomic signature of the host cell. Besides that, we found a massive increase of Olfactomedin 4 (*OLFM4*) upon infection and time. To validate and verify the data obtained by scRNA-seq of this vicarious hit, HCR-FISH was applied on the infected hITM probing for *OLFM4*. We could identify varying expression and spatial distribution of the Olfactomedin4 which is a target gene of NF-κB and negatively regulates NOD-induced NF-κB signalling and thus down-regulates innate immunity[271]. Based on the knowledge of *OLFM4* in the context of *H. pylori* infection and the here observed strengthened expression over time, we can hypothesise that the gene plays a role in the regulation of the epithelium specific inflammatory- and immune response in *Salmonella* infected small intestine. However, deeper investigation would be needed to identify the actual underlying mechanism, which is far beyond the scope of this work. Nevertheless, hits like this confirm the power of our deep scRNA-seq time course experiment.

4.5 Outlook

Human tissue models already have a great impact in the field of infection research and modern organoid cultures are a fairly accurate representation of the *in vivo* situation. However, reaching out for the goal to engineer a standardized *in vitro* model fully reflecting the diverse and functional human gut, demands still several improvements. Especially working with biological scaffolds makes the highly important standardization an ambitious challenge. Therefor common standards need to be defined and valid evaluation methods need to be developed. Besides standardization, as a long-term goal, multiple layers of complexity, such as vascularization, integration of microbiome and immune cells or improved, guided morphogenesis, that have all been already shown independently, should be combined to one functional model. Nevertheless, this technology is predestined of having a major impact on future medicine. Exemplary, patient-derived organoid cultures can help to model pathogenic mechanisms and disease progression and organoids derived from immortal pluripotent stem cells can be a tool to study genetic diseases or tissue development. In the near future we will still have a diversity of models tailored to distinct research questions. Planned experiments must be preceded by an assessment and evaluation process in order to make the right choice of model.

ScRNA-seq represents an ideal analytical method not only for the validation of tissue models but also to study infection. The variety of different approaches, on the one hand increasing the throughput massively and on the other hand focussing on smaller and smaller details with an unseen resolution, makes this technology so powerful. While still in development, we encounter limitations on all sides that need to be overcome. One neglected area was the field of regulatory noncoding RNA molecules such as miRNAs, lncRNAs and circRNAs, and observing mRNA modifications such as alternative splicing or RNA methylation. Recent progresses in epitranscriptomics could elucidate the importance of RNA editing and noncoding RNA regulations, but not yet on a single-cell level[272]. Another boundary we come up against is the spatio-temporal resolution of scRNA-seq data. Spatial transcriptomics has been offering completely new possibilities for a few years now, enabling to localise gene expression accurately. Temporal resolution however is mostly based on algorithms which are trying to retrieve a cells trajectory. A very smart recent approach allows the identification of newly

synthesised RNA in comparison to old RNA to actively record transcriptional activity of thousands of genes in a single cell[273].

Even though enormous progress has already been made, the biggest task will be to put all the information into one big picture. Increasing amounts of data are pushing us to the limits of human comprehension. With the actual approaches which aim to increase the throughput and to sequence even higher amounts of cells, comes the demand of bigger flow cells and more computing power. At the same time computational tools are flourishing trying to make the masses of data understandable for us.

Retrospectively, the speed at which the field has developed since the first publication by Tang *et al*.[100] is almost unbelievable and it is all the more exciting to see what we will achieve in the next 10 years.

5. Appendix:

Table 21: Microbial single-cell RNA-seq. Overview of the methods. Imdahl & Saliba Current opinion in Microbiology 2020

	Protozoa		**Yeast**			**Prokaryotes**				
	SMART-seq	**10x Chromium**	**10x Chromium**	**SCRB-seq**	**YSC-seq**	**CEL-seq2**	**MATQ-seq**	**PETRI-seq**	**Micro-SPLiT**	**PatH-Cap**
PMID	PMID: 30333624	PMID: 29094698	PMID: 32420869	PMID: 30718845	PMID: 30718850	PMID: 29073931	PMID: 32807892	PMID: 32451472	PMID: 33335020	PMID: 31848386
Organism	*Trypanosoma brucei, Plasmodium malariae*	*Plasmodium falciparum, P. berhei, P. knowlesi*	*Saccharomyces cerevisiae*	*Schizosaccharomyces pombe*	*Saccharomyces cerevisiae*	*Salmonella* Typhimurium	*Salmonella* Typhimurium, *Pseudomonas aeruginosa*	*Staphylococcus aureus, Escherichia coli*	*Bacillus subtilis, Escherichia coli*	*Pseudomonas aeruginosa*
Single-cell Isolation	FACS, microfluidics	Droplet based	Droplet based	Dissection Microscopy	FACS	FACS, microfluidics	FACS	Not needed	Not needed	FACS of infected host cell
Lysis	Takara Lyisis buffer	In Drop Lysis	Zymolyase and heat - In drop lysis	Heat based Lysis: 10 min 98°C	Zymolase based	TE and NP-40 based Lysisbuffer	Lysozyme based	Lyzozyme and Lysostaphin based Permeabilization	Tween-20 and Lysozyme based Permeabilization	TE and NP-40 based Lysisbuffer
Second strand Synthesis	TSO	TSO	TSO	TSO	TSO	Rnase H und DNA pol I	dC-tailing and T4 DNA pol	TSO	TSO	Rnase H und DNA pol I

Multiplexing	Library PCR-with barcoded primers	Barcoded RT-Primers	Barcoded RT-Primers	Barcoded RT-primers	Tn5 mediated tagmentation and barcode integration	Barcoded RT-primers	Library PCR-with barcoded primers	Ligation of barcoded RT-primers	Ligation of barcoded RT-primers	Barcoded RT-primers
Pooling before Library	No	Yes	Yes	No	No	Yes	No	Yes	Yes	Yes
Amplification strategy	PCR	PCR	PCR	PCR	PCR	In vitro transcription	PCR	PCR	PCR	In vitro transcription
UMI integration	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Gene Coverage	Full length	3'	3'	3'	5' strand specific	3' strand specific	Full length	3'	3'	3' strand specific
Number of cells per Assay	10^2	10^3 - 10^4	>10^3	>10^3	>10^2	10^2	10^2	>10^4	>10^4	10^3
Sensitivity	~ 1400 genes/cell	~ 1500 genes/cell	500 to 1300	~ 1000 genes/cell	~ 3400 genes/cell	~ 470 bacterial transcripts/ 0.01 pg of RNA	~ 170 genes/cell	~ 50 UMI transcripts/ cell	~ 300 mRNA traqnscripsts/cell	~ 120 genes/ 1-3 cells

Table 22: Overview of sequenced Salmonella SL1344 libraries displaying the library-ID (column 1), the associated growth condition (column 2) and the number of bacteria analysed (Column 3). Five outliers were removed from further analysis and their library ID are labeled with a grey background.

Library-ID	Condition	Number of cells
M1A1	Late Stationary Phase	1
M1A12	Late Stationary Phase	10
M1A2	Late Stationary Phase	1
M1A3	Late Stationary Phase	1
M1A6	Late Stationary Phase	10
M1A7	Late Stationary Phase	10
M1B1	Late Stationary Phase	1
M1B2	Late Stationary Phase	1
M1B3	Late Stationary Phase	1
M1B4	Late Stationary Phase	1
M1B7	Late Stationary Phase	10
M1C1	Late Stationary Phase	1
M1C4	Late Stationary Phase	1
M1C5	Late Stationary Phase	10
M1D2	Late Stationary Phase	1
M1D6	Late Stationary Phase	10
M1E11	Late Stationary Phase	10
M1E2	Late Stationary Phase	1
M1E3	Late Stationary Phase	1
M1E5	Late Stationary Phase	10
M1F11	Late Stationary Phase	10
M1F2	Late Stationary Phase	1
M1F3	Late Stationary Phase	1
M1F4	Late Stationary Phase	1
M1F5	Late Stationary Phase	10
M1F6	Late Stationary Phase	10
M1F7	Late Stationary Phase	10
M1F9	Late Stationary Phase	10
M1G2	Late Stationary Phase	1
M1G3	Late Stationary Phase	1
M1G6	Late Stationary Phase	10
M1H11	Late Stationary Phase	10
M1H2	Late Stationary Phase	1
M1H3	Late Stationary Phase	1
M1H5	Late Stationary Phase	10

M2A10	NaCl Shock	1
M2A11	Anaerobic Shock	1
M2A12	Anaerobic Shock	1
M2A2	Late Stationary Phase	10
M2A3	Late Stationary Phase	10
M2A5	NaCl Shock	10
M2A7	Anaerobic Shock	10
M2A8	NaCl Shock	1
M2A9	NaCl Shock	1
M2B10	NaCl Shock	1
M2B11	Anaerobic Shock	1
M2B12	Anaerobic Shock	1
M2B3	Late Stationary Phase	10
M2B5	NaCl Shock	10
M2B6	Anaerobic Shock	10
M2B7	Anaerobic Shock	10
M2B8	NaCl Shock	1
M2B9	NaCl Shock	1
M2C10	Anaerobic Shock	1
M2C11	Anaerobic Shock	1
M2C3	NaCl Shock	10
M2C4	NaCl Shock	10
M2C6	Anaerobic Shock	10
M2C7	NaCl Shock	1
M2C8	NaCl Shock	1
M2C9	NaCl Shock	1
M2D10	Anaerobic Shock	1
M2D11	Anaerobic Shock	1
M2D3	NaCl Shock	10
M2D4	NaCl Shock	10
M2D6	Anaerobic Shock	10
M2D7	NaCl Shock	1
M2D8	NaCl Shock	1
M2D9	NaCl Shock	1
M2E10	Anaerobic Shock	1
M2E11	Anaerobic Shock	1
M2E3	NaCl Shock	10
M2E4	NaCl Shock	10
M2E5	NaCl Shock	10
M2E6	Anaerobic Shock	10
M2E7	NaCl Shock	1

M2E8	NaCl Shock	1
M2E9	NaCl Shock	1
M2F10	Anaerobic Shock	1
M2F11	Anaerobic Shock	1
M2F3	NaCl Shock	10
M2F4	NaCl Shock	10
M2F6	Anaerobic Shock	10
M2F7	NaCl Shock	1
M2F8	NaCl Shock	1
M2F9	NaCl Shock	1
M2G10	Anaerobic Shock	1
M2G11	Anaerobic Shock	1
M2G2	Late Stationary Phase	10
M2G3	NaCl Shock	10
M2G4	NaCl Shock	10
M2G5	NaCl Shock	10
M2G6	Anaerobic Shock	10
M2G7	NaCl Shock	1
M2G8	NaCl Shock	1
M2G9	NaCl Shock	1
M2H10	Anaerobic Shock	1
M2H11	Anaerobic Shock	1
M2H3	NaCl Shock	10
M2H4	NaCl Shock	10
M2H5	NaCl Shock	10
M2H7	NaCl Shock	1
M2H8	NaCl Shock	1
M2H9	NaCl Shock	1
M3A1	NaCl Shock	10
M3A2	NaCl Shock	10
M3A3	NaCl Shock	10
M3A5	Anaerobic Shock	1
M3A6	Anaerobic Shock	1
M3A7	Anaerobic Shock	10
M3B4	Anaerobic Shock	10
M3B5	Anaerobic Shock	1
M3C4	Anaerobic Shock	10
M3C5	Anaerobic Shock	1
M3C6	Anaerobic Shock	10
M3C7	Anaerobic Shock	10
M3D4	Anaerobic Shock	10

M3D5	Anaerobic Shock	1	
M3E4	Anaerobic Shock	10	
M3E5	Anaerobic Shock	1	
M3E6	Anaerobic Shock	10	
M3E7	Anaerobic Shock	10	
M3F4	Anaerobic Shock	1	
M3F5	Anaerobic Shock	1	
M3F6	Anaerobic Shock	10	
M3G4	Anaerobic Shock	1	
M3G5	Anaerobic Shock	1	
M3G6	Anaerobic Shock	10	
M3H4	Anaerobic Shock	1	
M3H5	Anaerobic Shock	1	
M3H6	Anaerobic Shock	10	
	Late Stationary Phase	NaCl Shock	Anaerobic Shock
10-pooled bacteria	20	20	20
Outliers removed:		-1	-2
Single bacterium	19	24	28
Outliers removed:		-1	-1

Table 23: Mapping statistics for the 10-pooled bacteria (SL1344) libraries that passed the quality filter. Columns 1 and 2 recapitulate the library unique ID and the associated growth condition, respectively. Column 3 to 12 give the percentage for each transcript categories.

Library name	Condition	rRNA	tRNA	ncRNA	sRNA	protein_coding	pseudogene	antisense	ribozyme	antitoxin	IGR
M1A12	Late Stationary Phase	93,829	0,017	1,846	0,025	4,168	0,033	0,001	0	0	0,081
M1A6	Late Stationary Phase	94,959	0,016	2,822	0,038	2,132	0,006	0	0,015	0	0,011
M1A7	Late Stationary Phase	95,758	0,013	1,852	0,033	2,248	0,005	0	0,072	0	0,019
M1B7	Late Stationary Phase	96,311	0,019	1,722	0,006	1,869	0,004	0,001	0,022	0	0,045
M1C5	Late Stationary Phase	94,352	0,015	2,294	0,009	3,094	0,005	0	0,228	0	0,002
M1D6	Late Stationary Phase	96,636	0,008	0,984	0,098	2,223	0,003	0	0,002	0	0,045
M1E11	Late Stationary Phase	91,591	0,045	2,113	0,01	6,228	0,004	0	0,001	0,001	0,008
M1E5	Late Stationary Phase	94,995	0,054	1,838	0,012	3,045	0,005	0	0,014	0	0,037
M1F11	Late Stationary Phase	93,049	0,006	2,853	0,034	3,987	0,006	0	0,001	0	0,065
M1F5	Late Stationary Phase	97,126	0,007	1,085	0,051	1,278	0,222	0	0,034	0,004	0,192
M1F6	Late Stationary Phase	96,681	0,027	1,324	0,018	1,87	0,006	0	0,029	0	0,044
M1F7	Late Stationary Phase	94,38	0,016	3,313	0,026	2,204	0,018	0,001	0,001	0	0,04
M1F9	Late Stationary Phase	94,193	0,018	2,229	0,006	3,493	0,007	0	0	0	0,054
M1G6	Late Stationary Phase	94,906	0,021	1,819	0,071	3,042	0,074	0	0	0	0,066
M1H11	Late Stationary Phase	82,581	0,019	2,581	0,024	14,719	0,009	0	0	0	0,067
M1H5	Late Stationary Phase	93,956	0,004	2,921	0,008	2,773	0,004	0	0,187	0	0,147
M2A2	Late Stationary Phase	94,057	0,004	2,616	0,013	3,236	0,007	0	0,042	0	0,023
M2A3	Late Stationary Phase	91,189	0,011	3,658	0,022	5,052	0,012	0	0,008	0,001	0,048
M2A5	NaCl Shock	95,093	0,007	0,84	0,022	3,908	0,025	0	0,009	0	0,096
M2A7	Anaerobic Shock	93,247	0,027	1,078	0,017	5,534	0,004	0,007	0,014	0	0,072
M2B3	Late Stationary Phase	86,273	0,05	2,54	0,021	10,789	0,034	0	0,001	0,001	0,289
M2B5	NaCl Shock	94,679	0,03	0,708	0,041	4,362	0,037	0	0,032	0,001	0,109
M2B6	Anaerobic Shock	95,421	0,01	0,842	0,009	3,64	0,005	0	0,005	0	0,069

M2C3	NaCl Shock	94,365	0,02	0,827	0,034	4,577	0,011	0	0,021	0	0,144
M2C4	NaCl Shock	93,297	0,022	0,685	0,043	5,669	0,04	0	0,039	0,01	0,193
M2C6	Anaerobic Shock	93,883	0,003	0,89	0,006	5,092	0,005	0	0,021	0	0,101
M2D3	NaCl Shock	94,531	0,004	0,55	0,025	4,699	0,039	0	0,002	0	0,148
M2D4	NaCl Shock	89,63	0,004	0,342	0,008	9,922	0,007	0,002	0,019	0	0,066
M2D6	Anaerobic Shock	95,394	0,004	0,821	0,008	3,712	0,008	0	0,012	0	0,041
M2E3	NaCl Shock	94,545	0,009	0,737	0,022	4,552	0,005	0	0,003	0	0,126
M2E4	NaCl Shock	94,798	0,02	0,748	0,029	4,243	0,011	0	0,013	0	0,138
M2E5	NaCl Shock	94,47	0,028	0,782	0,011	4,518	0,023	0	0,011	0	0,155
M2E6	Anaerobic Shock	95,455	0,008	0,839	0,042	3,568	0,009	0	0,001	0	0,078
M2F3	NaCl Shock	94,848	0,035	0,666	0,006	4,307	0,004	0	0,016	0,003	0,115
M2F4	NaCl Shock	94,886	0,023	0,768	0,038	4,096	0,012	0	0,043	0	0,134
M2F6	Anaerobic Shock	94,592	0,008	0,804	0,02	4,449	0,005	0	0,038	0	0,083
M2G2	Late Stationary Phase	94,004	0,023	2,174	0,022	3,68	0,009	0	0	0	0,088
M2G4	NaCl Shock	93,278	0,017	0,898	0,026	5,588	0,024	0	0,005	0	0,163
M2G5	NaCl Shock	90,71	0,072	0,767	0,032	8,207	0,044	0,001	0	0	0,167
M2G6	Anaerobic Shock	93,511	0,023	0,844	0,024	5,504	0,007	0	0	0	0,086
M2H3	NaCl Shock	95,139	0,029	0,632	0,032	3,975	0,006	0	0,001	0	0,184
M2H4	NaCl Shock	93,723	0,006	0,898	0,027	5,157	0,02	0	0,039	0	0,129
M2H5	NaCl Shock	90,496	0,074	0,75	0,025	8,368	0,008	0,004	0,016	0	0,258
M3A1	NaCl Shock	93,709	0,005	0,588	0,023	5,504	0,009	0,002	0,007	0	0,152
M3A2	NaCl Shock	94,138	0,027	0,559	0,015	5,123	0,021	0,009	0	0	0,108
M3A3	NaCl Shock	92,983	0,007	0,658	0,016	6,19	0,007	0,005	0,004	0	0,129
M3A7	Anaerobic Shock	93,853	0,01	0,395	0,009	5,656	0,015	0	0,033	0,001	0,027
M3B4	Anaerobic Shock	95,151	0,002	1,304	0,003	3,495	0,003	0	0,019	0	0,022
M3C6	Anaerobic Shock	93,293	0,02	0,541	0,012	6,06	0,02	0	0,019	0,002	0,033
M3C7	Anaerobic Shock	93,848	0,01	0,401	0,008	5,609	0,023	0	0,066	0,001	0,034
M3D4	Anaerobic Shock	89,852	0,014	2,534	0,029	7,478	0,003	0	0,001	0	0,089

M3E4	Anaerobic Shock	94,42	0,003	1,375	0,007	4,076	0,014	0	0,045	0,001	0,058
M3E6	Anaerobic Shock	93,928	0,01	0,497	0,011	5,486	0,018	0	0,035	0,001	0,015
M3E7	Anaerobic Shock	92,928	0,015	0,489	0,015	6,463	0,018	0	0,052	0,002	0,018
M3F6	Anaerobic Shock	92,951	0,016	0,462	0,017	6,453	0,022	0	0,054	0,002	0,021
M3G6	Anaerobic Shock	94,751	0,022	2,896	0,019	2,234	0,016	0	0,004	0	0,057
M3H6	Anaerobic Shock	94,318	0,01	0,541	0,02	4,967	0,015	0	0,097	0,001	0,03

	rRNA10	tRNA10	ncRNA10	sRNA	CDS	pseudogene	antisense	rybozyme	antitoxin	IGR
Average	**93,701**	**0,018**	**1,325**	**0,023**	**4,799**	**0,018**	**0,001**	**0,025**	**0,001**	**0,088**
St_dev_P	2,350	0,015	0,877	0,016	2,296	0,030	0,002	0,041	0,001	0,063
St_dev_S	2,371	0,015	0,885	0,017	2,316	0,031	0,002	0,041	0,002	0,063

Table 24: Mapping statistics for the single bacteria (SL1344) libraries that passed the quality filter. Columns 1 and 2 recapitulate the library unique ID and the associated growth condition, respectively. Column 3 to 12 give the percentage for each transcript categories.

Library name	Condition	rRNA	tRNA	ncRNA	sRNA	protein_coding	pseudogene	antisense	ribozyme	antitoxin	IGR
M1A1	Late Stationary Phase	93,394	0,184	0,803	0,109	5,459	0,031	0,007	0,001	0,002	0,01
M1A2	Late Stationary Phase	92,263	0,064	0,287	0,019	7,321	0,028	0,001	0,002	0,001	0,013
M1A3	Late Stationary Phase	95,074	0,008	1,721	0,015	3,142	0,012	0,001	0	0	0,027
M1B1	Late Stationary Phase	92,514	0,047	2,258	0,007	5,156	0,014	0	0,001	0,001	0,002
M1B2	Late Stationary Phase	87,187	0,062	0,747	0,128	11,737	0,054	0,001	0,028	0,001	0,055
M1B3	Late Stationary Phase	77,692	0,145	0,813	0,099	21,059	0,116	0,001	0,009	0,003	0,062
M1B4	Late Stationary Phase	94,498	0,012	1,267	0,015	3,873	0,017	0,006	0,304	0	0,009
M1C2	Late Stationary Phase	89,903	0,155	0,232	0,069	9,567	0,038	0,001	0,004	0,003	0,027
M1C4	Late Stationary Phase	82,813	0,04	0,145	0,043	16,923	0,021	0	0,001	0,001	0,013
M1D2	Late Stationary Phase	91,337	0,029	0,242	0,038	8,275	0,039	0,001	0,002	0,002	0,034
M1E2	Late Stationary Phase	90,277	0,025	4,186	0,046	5,346	0,018	0	0,019	0,001	0,084
M1E3	Late Stationary Phase	93,822	0,022	0,208	0,04	5,859	0,031	0	0,005	0,001	0,012
M1F2	Late Stationary Phase	95,121	0,034	1,901	0,082	2,826	0,012	0,001	0	0	0,023
M1F3	Late Stationary Phase	88,718	0,039	0,239	0,026	10,914	0,048	0,001	0,005	0,001	0,01
M1F4	Late Stationary Phase	75,371	0,073	1,368	0,73	22,32	0,104	0	0,007	0,003	0,024
M1G2	Late Stationary Phase	92,906	0,022	3,225	0,01	3,79	0,014	0	0	0,001	0,032
M1G3	Late Stationary Phase	94,744	0,027	2,53	0,005	2,665	0,011	0	0,012	0,001	0,003
M1H2	Late Stationary Phase	91,732	0,018	4,31	0,009	3,708	0,017	0	0	0,001	0,204
M1H3	Late Stationary Phase	90,681	0,021	2,265	0,015	6,971	0,037	0	0,001	0,001	0,01
M2A10	NaCl Shock	85,222	0,021	1,378	0,052	13,054	0,031	0	0,001	0,001	0,24
M2A11	Anaerobic Shock	95,09	0,01	0,204	0,093	4,569	0,019	0,001	0,004	0	0,01
M2A12	Anaerobic Shock	94,612	0,006	1,018	0,034	4,052	0,009	0	0,001	0	0,266
M2A8	NaCl Shock	91,549	0,065	0,867	0,059	7,258	0,014	0,001	0,019	0	0,167

M2A9	NaCl Shock	93,636	0,116	0,607	0,024	5,518	0,01	0	0	0	0,088
M2B10	NaCl Shock	86,392	0,01	0,37	0,007	13,038	0,023	0	0,001	0	0,159
M2B11	Anaerobic Shock	90,119	0,013	0,272	0,124	9,423	0,023	0	0,009	0,001	0,016
M2B12	Anaerobic Shock	95,159	0,006	0,756	0,067	3,971	0,01	0	0,001	0	0,029
M2B8	NaCl Shock	94,233	0,006	0,545	0,009	5,059	0,017	0	0,046	0	0,085
M2B9	NaCl Shock	95,804	0,005	1,399	0,018	2,731	0,004	0	0	0	0,038
M2C11	Anaerobic Shock	92,781	0,004	1,511	0,052	5,403	0,009	0	0,002	0	0,238
M2C7	NaCl Shock	90,626	0,007	1,312	0,069	7,542	0,006	0	0,033	0	0,404
M2C8	NaCl Shock	93,382	0,005	0,656	0,023	5,803	0,01	0	0	0	0,12
M2C9	NaCl Shock	93,963	0,006	0,203	0,168	5,579	0,008	0	0	0	0,073
M2D10	Anaerobic Shock	94,631	0,009	0,765	0,018	4,52	0,013	0,001	0,005	0	0,038
M2D11	Anaerobic Shock	90,511	0,008	1,621	0,019	7,692	0,025	0	0,001	0	0,122
M2D7	NaCl Shock	94,269	0,018	0,23	0,02	5,312	0,01	0	0	0	0,139
M2D8	NaCl Shock	95,039	0,074	0,843	0,008	3,941	0,007	0,001	0	0	0,087
M2D9	NaCl Shock	93,174	0,006	0,088	0,059	6,604	0,008	0	0,001	0	0,06
M2E10	Anaerobic Shock	94,206	0,006	0,625	0,035	4,985	0,012	0,005	0	0	0,125
M2E11	Anaerobic Shock	91,837	0,011	1,294	0,054	6,53	0,007	0	0,001	0	0,265
M2E7	NaCl Shock	93,529	0,056	0,729	0,01	5,59	0,013	0,002	0,004	0,001	0,066
M2E8	NaCl Shock	93,51	0,009	0,815	0,018	5,51	0,005	0	0,033	0	0,099
M2E9	NaCl Shock	95,161	0,007	0,94	0,028	3,668	0,004	0	0	0	0,192
M2F10	Anaerobic Shock	92,162	0,069	0,399	0,023	7,039	0,009	0	0,001	0	0,297
M2F11	Anaerobic Shock	89,647	0,007	1,335	0,046	8,816	0,006	0	0,097	0	0,046
M2F7	NaCl Shock	90,82	0,115	1,434	0,076	7,278	0,047	0,002	0,009	0,001	0,22
M2F8	NaCl Shock	89,559	0,03	1,145	0,208	8,94	0,039	0,018	0,022	0,001	0,038
M2F9	NaCl Shock	92,219	0,052	3,418	0,248	4,026	0,03	0	0,001	0,003	0,004
M2G10	Anaerobic Shock	94,633	0,005	0,565	0,039	4,712	0,006	0,001	0,001	0	0,037
M2G11	Anaerobic Shock	87,384	0,025	1,703	0,017	10,847	0,007	0	0	0	0,016
M2G7	NaCl Shock	95,5	0,004	1,189	0,005	3,287	0,006	0	0	0,001	0,008

M2G8	NaCl Shock	94,549	0,008	0,296	0,067	5,03	0,022	0	0,003	0,001	0,025
M2G9	NaCl Shock	91,898	0,003	0,676	0,07	7,291	0,008	0	0,017	0	0,037
M2H10	Anaerobic Shock	94,559	0,02	0,925	0,05	4,354	0,015	0	0,019	0	0,058
M2H11	Anaerobic Shock	92,483	0,005	0,701	0,007	6,779	0,007	0,003	0,001	0	0,013
M2H8	NaCl Shock	93,629	0,004	0,901	0,009	5,341	0,014	0	0	0	0,101
M2H9	NaCl Shock	95,211	0,004	0,92	0,017	3,802	0,008	0,001	0,018	0	0,02
M3A5	Anaerobic Shock	96,326	0,004	1,404	0,026	2,211	0,005	0,001	0,003	0	0,02
M3A6	Anaerobic Shock	96,04	0,005	1,144	0,019	2,777	0,004	0	0	0	0,011
M3B5	Anaerobic Shock	96,134	0,002	1,157	0,015	2,459	0,005	0	0,15	0	0,077
M3C5	Anaerobic Shock	94,883	0,004	1,903	0,093	3,095	0,005	0	0	0	0,017
M3D5	Anaerobic Shock	96,854	0,004	1,128	0,008	1,981	0,003	0	0,002	0	0,02
M3E5	Anaerobic Shock	95,849	0,002	1,412	0,005	2,722	0,003	0	0,003	0	0,004
M3F4	Anaerobic Shock	97,069	0,004	1,134	0,056	1,697	0,004	0	0,02	0	0,015
M3F5	Anaerobic Shock	95,068	0,002	1,246	0,006	3,616	0,004	0	0	0	0,058
M3G4	Anaerobic Shock	96,839	0,004	0,59	0,023	2,508	0,006	0	0,006	0	0,023
M3G5	Anaerobic Shock	95,396	0,005	1,426	0,04	3,127	0,005	0	0	0	0,001
M3H4	Anaerobic Shock	97,777	0,002	0,645	0,01	1,529	0,002	0	0,013	0	0,022
M3H5	Anaerobic Shock	96,775	0,002	1,254	0,014	1,796	0,003	0	0,051	0	0,104

	rRNA1	tRNA	ncRNA	sRNA	CDS	pseudogene	antisense	rybozyme	antitoxine	IGR
Average	**92,576**	**0,028**	**1,128**	**0,055**	**6,106**	**0,018**	**0,001**	**0,014**	**0,000**	**0,074**
stdev_P	4,016	0,038	0,861	0,094	4,013	0,020	0,002	0,042	0,001	0,084
stdev_S	4,045	0,039	0,867	0,095	4,042	0,020	0,002	0,042	0,001	0,085

Table 25: (Related to Fig. 7 and c; upper heatmap (b) and panels (c)). Differentially expressed genes for 10-pooled bacteria and comparison of our study with the reference bulk RNA-seq dataset from Kröger et al. 2013 (PMID: 24331466). Gene ID and gene name of differentially expressed genes of 10-pooled bacteria in the three growth conditions (Figure 3b) are listed in column 1 and 2, respectively (color code in Column 2: blue: 'Late Stationary Phase'; green: 'Salt (NaCl) shock'; red: 'Anaerobic shock'). Genes are ordered as in the heatmap. For each gene differentially expressed between Salt (NaCl) and Anaerobic shocks, column 3 and 4 recapitulate the gene expression extracted from Kröger et al. under Salt (NaCl) shock and Anaerobic shock, respectively. The ratio and log10 transformed ratio of (NaCl/Anaerobic) of the latter expression values from bulk-RNA-seq from Kröger et al. 2013 are calculated in columns 5 and 6, respectively.

Gene ID	Gene name	Expression values Salt (NaCl) Shock (Kröger et al) (TPM)	Expression values Anaerobic Shock (Kröger et al.) (TPM)	Ratio Salt(NaCl) /Anaerobic	Ratio Salt(NaCl) /Anaerobic (Log10)
SL1344_1619	pspB				
SL1344_2376	mntH				
SL1344_1618	pspC				
SL1344_0462	rpmE2				
SL1344_3761	ilvN				
SL1344_2773	csiD				
SL1344_P1_0022	traC				
SL1344_1620	pspA				
SL1344_P2_0074	traJ				
SL1344_0649					
SL1344_1063	putP				
SL1344_P2_0081	shfB				
SL1344_2756	fljB				
SL1344_3300	arcB				
SL1344_P2_0082	shfB2				
SL1344_1223					
SL1344_0899	lolA				
SL1344_1006	rmf				
SL1344_2287	nuoL				
SL1344_3936	fadB				
SL1344_0936	ompF				
EBG00001133906	Bacteria_small_SRP				
SL1344_1409	pntA				
SL1344_4056	katG				
SL1344_4487	yjjG				
SL1344_1763	yebN				
SL1344_4180	pspG				
SL1344_2131	mrp				

SL1344_2860	iacP				
SL1344_1148	ndh				
SL1344_2059	rfbP				
SL1344_3301	yhcC				
SL1344_4216	nrfD				
SL1344_0573	fepA				
SL1344_2239	rcsB				
SL1344_3106					
SL1344_4465					
EBG00000241440	IsrM				
SL1344_1616	pspE				
SL1344_2731					
SL1344_3746					
SL1344_1415	asr				
SL1344_2847	hilC				
SL1344_2838	hypE				
SL1344_1828	ruvB				
SL1344_4499	stjB				
SL1344_2043	sopA				
EBG00001133739	tmRNA				
SL1344_4346					
EBG00001133849	RNaseP_bact_a				
SL1344_4217	nrfE				
SL1344_1508	narZ				
SL1344_1696	ychP				
SL1344_4507	nadR				
SL1344_0348					
SL1344_4520	sthB				
SL1344_2786	ygaM	367	73,3	5,0068	0,6996
SL1344_1915	gcpA	7,9	6,9	1,1449	0,0588
SL1344_4224	phnO	42,7	13,5	3,163	0,5001
SL1344_3030	gcvH	107,1	524,4	0,2042	-0,6899
SL1344_0818	ybiV(1)	139,7	21,1	6,6209	0,8209
EBG00001133861	P26	NA	NA	NA	NA
SL1344_2297	nuoA	228	109,9	2,0746	0,3169
SL1344_P1_0080	ccdA	NA	NA	NA	NA
SL1344_0081		444,8	11,7	38,0171	1,58
SL1344_2792	nrdE	29,6	2,3	12,8696	1,1096
SL1344_4228	basS	18	25,9	0,695	-0,158
SL1344_0354		690	16,8	41,0714	1,6135
SL1344_1914	mngB	4,6	2,3	2	0,301
SL1344_0431		23,7	4,6	5,1522	0,712
SL1344_2304	yfbU	75	102,1	0,7346	-0,134

SL1344_1019	yccV	596,4	80,8	7,3812	0,8681
SL1344_0426	phnX	1	2,3	0,4348	-0,3617
SL1344_3471	yhgF	32	10,1	3,1683	0,5008
SL1344_1288	ydiQ	0	0,6	0	-Inf
SL1344_3397	rplN	1856,6	405,5	4,5785	0,6607
SL1344_1202		1183,3	62,5	18,9328	1,2772
SL1344_2370	yfdZ	56	51,7	1,0832	0,0347
SL1344_2796	proX	762	3,3	230,9091	2,3634
SL1344_2625	pheA	54,3	11,9	4,563	0,6593
SL1344_2436	tal	38,3	8,4	4,5595	0,6589
SL1344_0433	cyoE	232	43,9	5,2847	0,723
SL1344_0459	ybaY	266,1	44,8	5,9397	0,7738
SL1344_4369	cybC	539,4	225	2,3973	0,3797
SL1344_0461	ylaB	23,5	38,7	0,6072	-0,2166
SL1344_1732	ycgB	42,3	28,5	1,4842	0,1715
SL1344_1304	sufB	18,2	5,3	3,434	0,5358
SL1344_2660	smpB	240,3	106,4	2,2585	0,3538
SL1344_0868	ybjP	59,8	45,6	1,3114	0,1177
SL1344_2182	fruK	3	8,5	0,3529	-0,4523
SL1344_3472	feoA	52,3	15	3,4867	0,5424
SL1344_0212	yaeH	363,7	1003,5	0,3624	-0,4408
SL1344_3202	yqjE	159,7	200	0,7985	-0,0977
SL1344_1132	fabG	358,4	225,1	1,5922	0,202
SL1344_2648	rimM	523	175,2	2,9852	0,475
SL1344_3492	glpE	242,8	82,5	2,943	0,4688
SL1344_0386	yaiE	185,4	223,4	0,8299	-0,081
SL1344_0763	ybhC	17,5	16,8	1,0417	0,0177
SL1344_2183	fruB	2,3	3,5	0,6571	-0,1823
SL1344_0737	aroG	67,3	71,6	0,9399	-0,0269
SL1344_2108	baeR	8,5	9,3	0,914	-0,0391
SL1344_2486	yfgA	111,2	71	1,5662	0,1948
SL1344_3237	yraM	28,9	46	0,6283	-0,2019
SL1344_4456	hsdM	7,4	36,8	0,2011	-0,6966
SL1344_4450		13,3	8,8	1,5114	0,1794
SL1344_3955	typA	75,3	40	1,8825	0,2747
SL1344_3282	yrbC	89,5	58,6	1,5273	0,1839
SL1344_1245	nadE	139,7	81,4	1,7162	0,2346
SL1344_2464	ppk	70,5	56,8	1,2412	0,0938
SL1344_3396	rplX	3303,7	608,5	5,4293	0,7347
SL1344_4503	lplA	22,3	15,7	1,4204	0,1524
SL1344_4227	proP	1160,2	58	20,0034	1,3011
SL1344_0498	ybbO	24,8	21,7	1,1429	0,058
SL1344_3394	rpsN	1616,8	309	5,2324	0,7187
SL1344_2747		24,7	34,9	0,7077	-0,1501

SL1344_2118	fbaB	24	12,2	1,9672	0,2939
SL1344_2811	mltB	49,2	25,1	1,9602	0,2923
SL1344_2938	rumA	61,3	113,3	0,541	-0,2668
EBG00000241426	STnc700	NA	NA	NA	NA
SL1344_1774	prc	59,7	51,1	1,1683	0,0676
SL1344_P1_0055	parA	NA	NA	NA	NA
SL1344_4230	yjdB	65,6	23,2	2,8276	0,4514
SL1344_3838	atpB	208,8	211,6	0,9868	-0,0058
SL1344_P2_0010	yafB	NA	NA	NA	NA
SL1344_0871		13,3	15,5	0,8581	-0,0665
SL1344_1496	sfcA	20,8	23,6	0,8814	-0,0548
SL1344_0866	artI	32,1	46	0,6978	-0,1563
SL1344_3452	damX	63,6	72,6	0,876	-0,0575
SL1344_3505	glgB	38,5	45,7	0,8425	-0,0745
SL1344_3494		4,4	22,7	0,1938	-0,7126
SL1344_4076	yijC	158,9	144,3	1,1012	0,0419
SL1344_0089	apaH	51,8	38,7	1,3385	0,1266
SL1344_1537		33,1	125,7	0,2633	-0,5795
SL1344_0047	ileS	69,4	62,4	1,1122	0,0462
SL1344_1243		84,5	10,6	7,9717	0,9016
SL1344_3041	yggE	106,1	50,8	2,0886	0,3199
SL1344_3034	pepP	59,2	64,4	0,9193	-0,0366
SL1344_2280	elaB	254,5	147,8	1,7219	0,236
SL1344_3890	hemY	59,8	67,1	0,8912	-0,05
SL1344_2607	srmB	42	27,9	1,5054	0,1776
SL1344_3201	yqjC	788,3	229	3,4424	0,5369
SL1344_3708	spoT	86,3	74,3	1,1615	0,065
SL1344_1666	ispZ	49,9	44,7	1,1163	0,0478
SL1344_3393	rpsH	1802,1	318,1	5,6652	0,7532
SL1344_3523	ugpB	39,3	29,6	1,3277	0,1231
SL1344_1767	yobF	2068,5	2360,9	0,8761	-0,0574
SL1344_0044	rpsT	3440,4	967,7	3,5552	0,5509
SL1344_1493	osmC	260,7	25,4	10,2638	1,0113
SL1344_3029	gcvP	24,8	114,2	0,2172	-0,6632
SL1344_2290	nuoI	85,7	88,6	0,9673	-0,0145
SL1344_3875	trxA	1663,3	477	3,487	0,5425
SL1344_0442	clpP	509,3	218,4	2,332	0,3677
SL1344_4013	cdh-a	134,8	23,6	5,7119	0,7568
SL1344_0045	yaaY	86,8	50,9	1,7053	0,2318
SL1344_4008	cpxR	56,3	46,8	1,203	0,0803
SL1344_1237	xthA	21,1	26	0,8115	-0,0907
SL1344_0187	dksA	412,6	225	1,8338	0,2633
SL1344_1887	fliB	35,5	24,7	1,4372	0,1575
SL1344_0658	phoL	223	82,1	2,7162	0,434

SL1344_2273	pmrD	1077,3	329,7	3,2675	0,5142
SL1344_0627	mrdB	22,8	19,5	1,1692	0,0679
SL1344_3462	igaA	72,4	56,5	1,2814	0,1077
SL1344_3419	yheO	114,1	85,4	1,3361	0,1258
SL1344_0937	asnCa	80,5	87,9	0,9158	-0,0382
SL1344_0192	fhuA	49,5	7,9	6,2658	0,797
SL1344_2649	rps16	779,5	143,9	5,417	0,7338
SL1344_1196		52	6,7	7,7612	0,8899
SL1344_2267	pmrF	49	17,7	2,7684	0,4422
SL1344_0225	yaeT	181,2	134,8	1,3442	0,1285
SL1344_3184	dnaG	113,2	46,3	2,4449	0,3883
SL1344_1646	topA	87,8	32,6	2,6933	0,4303
SL1344_1633	rnb	36,9	24,8	1,4879	0,1726
SL1344_0436	cyoB	185,5	21,4	8,6682	0,9379
SL1344_3667	secB	451,1	348,6	1,294	0,1119
SL1344_2203	yejK	20,3	19,8	1,0253	0,0108
SL1344_3410	bfr	420,1	228,3	1,8401	0,2648
SL1344_0434	cyoD	333,1	49,7	6,7022	0,8262
SL1344_0172	yadF	468,8	89,8	5,2205	0,7177
SL1344_2483	yfgM	76	51	1,4902	0,1732
SL1344_2245	ubiG	96,5	49,1	1,9654	0,2934
SL1344_1577	ldhA	108,1	42,7	2,5316	0,4034
SL1344_4311	yjfN	20,7	249,9	0,0828	-1,0818
SL1344_3501	glgP	13,9	97,9	0,142	-0,8478
SL1344_3382	pez	994	310,2	3,2044	0,5057
SL1344_1755	sdaA	20,7	46,6	0,4442	-0,3524
SL1344_3938	yigZ	93,4	67,1	1,392	0,1436
SL1344_3862	ilvM	137,1	43,2	3,1736	0,5016
SL1344_1571	hrpA	24,5	20	1,225	0,0881
SL1344_3947	polA	31,9	48,6	0,6564	-0,1828
SL1344_2289	nuoJ	45,1	53,2	0,8477	-0,0717
SL1344_4325	priB	496,6	145,9	3,4037	0,532
SL1344_0191	mrcB	45,3	38,5	1,1766	0,0706
SL1344_4489	osmY	2451,9	88,2	27,7993	1,444
SL1344_2937	relA	34,8	79,9	0,4355	-0,361
SL1344_0656	ybeX	129,7	67,7	1,9158	0,2824
SL1344_2963	ygdI	1685,4	492,6	3,4214	0,5342
SL1344_1396	ydgA	44,6	17,8	2,5056	0,3989
SL1344_0119	mraZ	405	288,2	1,4053	0,1478
SL1344_1220	yeaG	25,4	28,6	0,8881	-0,0515
SL1344_0444	lon	358	118,8	3,0135	0,4791
SL1344_3391	rl18	1798,9	388,7	4,628	0,6654
SL1344_1226	yeaA	290,2	98,7	2,9402	0,4684
SL1344_0788	ybhO	10,2	2,6	3,9231	0,5936

SL1344_3277	ispB	116,2	92,8	1,2522	0,0977
SL1344_0170	gcd	60,9	7,5	8,12	0,9096
SL1344_4181	qor	31,5	36,7	0,8583	-0,0664
SL1344_1422		162,3	2,2	73,7727	1,8679
SL1344_2482	yfgL	154,9	114,4	1,354	0,1316
SL1344_1863	otsB	130,6	5,9	22,1356	1,3451
SL1344_2296	nuoB	156,3	117,5	1,3302	0,1239
SL1344_3463	yrfG	118,2	43,3	2,7298	0,4361
SL1344_1876	uvrC	26,7	44,9	0,5947	-0,2257
SL1344_1253	katE	23,5	20,3	1,1576	0,0636
SL1344_1982		15	2,6	5,7692	0,7611
SL1344_1753	pabB	32	24,7	1,2955	0,1125
SL1344_0443	clpX	611	318,7	1,9172	0,2827
SL1344_2068	rfbG	311,5	129,1	2,4129	0,3825
SL1344_3587	yhjS	88	49,2	1,7886	0,2525
SL1344_0808	ompX	2073,6	547,7	3,786	0,5782
SL1344_0997	pepN	28,6	59,2	0,4831	-0,316
SL1344_3676	rfaD	167,1	137,4	1,2162	0,085
SL1344_3668	grxC	362,5	285,4	1,2701	0,1039
SL1344_1394	ydgJ	22,8	33,7	0,6766	-0,1697
SL1344_4312	yjfO	100,1	999,5	0,1002	-0,9993
SL1344_2232	eco	214,5	37,2	5,7661	0,7609
SL1344_2393	cysK	184,8	64	2,8875	0,4605
SL1344_0468	acrB	69,6	42,3	1,6454	0,2163
SL1344_3199	yqjA	69,7	73,6	0,947	-0,0236
SL1344_P1_0072		NA	NA	NA	NA
SL1344_2190	yeiU	118,6	24,2	4,9008	0,6903
SL1344_4060	ptsA	18	23,3	0,7725	-0,1121
SL1344_4005	yiiM	22,7	52,7	0,4307	-0,3658
SL1344_3614	yiaG	218,4	180	1,2133	0,084
SL1344_2048	hisG	18,6	4,9	3,7959	0,5793
SL1344_2381	gltX	42,5	41,6	1,0216	0,0093
SL1344_3112		164	134,2	1,2221	0,0871
SL1344_3696	dfp	75,5	56,8	1,3292	0,1236
SL1344_3191	oat	7,3	4,6	1,587	0,2006
SL1344_2932	pyrG	177,8	188,8	0,9417	-0,0261
SL1344_1737		1090	871,1	1,2513	0,0974
SL1344_4040	hslU	485,8	74,1	6,556	0,8166
SL1344_1131	fabD	109,8	52,3	2,0994	0,3221
SL1344_0659	miaB	44,5	25,7	1,7315	0,2384
SL1344_0714	sdhC	413,3	62,6	6,6022	0,8197
SL1344_0384	yaiA	564,5	104,1	5,4227	0,7342
SL1344_1489		5,3	2,4	2,2083	0,3441
SL1344_1683	galU	112,1	70,3	1,5946	0,2027

SL1344_3196	ygjR	7,7	47,7	0,1614	-0,792
SL1344_0469	acrA	97,6	52,8	1,8485	0,2668
SL1344_1057	yccJ	153,2	770	0,199	-0,7012
SL1344_3707	rpoZ	450,1	157,2	2,8632	0,4569
SL1344_2437	tktB	22,5	9,5	2,3684	0,3745
SL1344_4041	hslV	670,4	57,5	11,6591	1,0667
SL1344_0467	ybaJ	901,6	399,2	2,2585	0,3538
SL1344_2105	yegO	3	3,4	0,8824	-0,0544
SL1344_4326	rpsR	562,8	86,4	6,5139	0,8138
SL1344_0529	ppiB	575,3	169,3	3,3981	0,5312
SL1344_1733	fadR	80,9	141,1	0,5734	-0,2416
SL1344_3453	aroB	74	72,5	1,0207	0,0089
SL1344_0616	pagP	593,3	74,7	7,9424	0,9
SL1344_2594		926,6	729	1,2711	0,1042
SL1344_2391	zipA	381,2	307,6	1,2393	0,0932
SL1344_3276	rplU	1855,4	443,3	4,1854	0,6217
SL1344_1062	putA	9,9	3,5	2,8286	0,4516
SL1344_2515	csiE	16,3	43,4	0,3756	-0,4253
SL1344_4029		15,6	21,8	0,7156	-0,1453
SL1344_1377	slyB	628,1	272	2,3092	0,3635
SL1344_1372	sodCb	39,8	58	0,6862	-0,1635
SL1344_P2_0012	cib	NA	NA	NA	NA
SL1344_3053	yggG	29	28,8	1,0069	0,003
SL1344_3161	ygiB	152,4	107,3	1,4203	0,1524
SL1344_2241	gyrA	127,6	60,4	2,1126	0,3248
SL1344_4190	uvrA	35,2	18,5	1,9027	0,2794
SL1344_1820	zwf	34,4	38,3	0,8982	-0,0466
SL1344_1058	wrbA	101,3	216,8	0,4673	-0,3304
SL1344_4004	sodA	468,1	123,2	3,7995	0,5797
SL1344_3554	pitA	19,8	43,4	0,4562	-0,3408
SL1344_0480	htpG	1057,3	109,6	9,6469	0,9844
SL1344_3454	aroK	171,1	165,1	1,0363	0,0155
SL1344_0160		783,7	104,5	7,4995	0,875
SL1344_3353	accC	94	71,7	1,311	0,1176
SL1344_1490		8,5	2,3	3,6957	0,5677
SL1344_3407	rplC	1142	208	5,4904	0,7396
SL1344_2615	yfiQ	14,4	53,4	0,2697	-0,5692
SL1344_2292	nuoG	86,1	104,3	0,8255	-0,0833
SL1344_2628	aroF	75,4	26,9	2,803	0,4476
SL1344_1384	tppB	49,2	26,6	1,8496	0,2671
SL1344_3346	mreB	76,9	81,5	0,9436	-0,0252
SL1344_3255	pnp	80,4	79,7	1,0088	0,0038
SL1344_3392	rplF	1327,5	290,3	4,5729	0,6602
SL1344_0617	cspE	2444,1	5682,8	0,4301	-0,3664

SL1344_0879	ybjX	306,5	63,5	4,8268	0,6837
SL1344_4496	deoB	150,1	89,9	1,6696	0,2226
SL1344_1905	fliM	32,4	30,4	1,0658	0,0277
SL1344_3596	dppA	40,8	98,4	0,4146	-0,3823
SL1344_0715	sdhD	299,7	64,8	4,625	0,6651
SL1344_0446	cypD	63,7	56,4	1,1294	0,0529
SL1344_4463	cstAb	41,4	83,5	0,4958	-0,3047
SL1344_0312	pepD	46,8	127,8	0,3662	-0,4363
SL1344_2266	yfbE	43,1	12,9	3,3411	0,5239
SL1344_2785	ygaC	289,8	11,8	24,5593	1,3902
SL1344_0235	ldcC	23,7	25,6	0,9258	-0,0335
SL1344_4177	zur	60,4	57,5	1,0504	0,0214
SL1344_3473	feoB	24	14,7	1,6327	0,2129
SL1344_0152	aceE	283,1	41,1	6,8881	0,8381
SL1344_3383	rpsD	1077,6	414,4	2,6004	0,415
SL1344_4260	dipZ	27,4	11,3	2,4248	0,3847
SL1344_1002	uup	22,9	13,3	1,7218	0,236
SL1344_4044	priA	42,9	26,9	1,5948	0,2027
SL1344_4267	groEL	1361,7	282,1	4,827	0,6837
SL1344_3015	idi	88,9	47,1	1,8875	0,2759
SL1344_3720	rhuM	28,2	24,7	1,1417	0,0576
SL1344_3385	rpsM	1974,7	530,9	3,7195	0,5705
SL1344_0802	ybiO	14,7	3,5	4,2	0,6232
SL1344_2237	ompC	739	835	0,885	-0,053
SL1344_4030	tpiA	247,6	647,8	0,3822	-0,4177
SL1344_3806	rpmH	7727,5	849,3	9,0987	0,959
SL1344_3876	rho	276,8	158,2	1,7497	0,243
SL1344_4405	valS	57,7	32,3	1,7864	0,252
SL1344_4157	pgi	69,2	105,9	0,6534	-0,1848
SL1344_4069	ppc	43,4	66,7	0,6507	-0,1866
SL1344_0918	rpsA	698,5	237,6	2,9398	0,4683
SL1344_2067	rfbH	421,7	130,1	3,2414	0,5107
SL1344_3984	fdoG	86	65,4	1,315	0,1189
SL1344_4090	rplJ	1267,1	576,3	2,1987	0,3422
SL1344_3408	rpsJ	1883,8	302,8	6,2213	0,7939
SL1344_1680	adh	21,2	435,8	0,0486	-1,313
SL1344_2395	ptsI	170,9	386,3	0,4424	-0,3542
SL1344_3412	tufA	441,9	129,7	3,4071	0,5324
SL1344_0005	yaaA	62,3	50,7	1,2288	0,0895
SL1344_3675	kbl	52,8	143,3	2,7140	0,4336
SL1344_0910	pflB	47,5	439,7	9,2568	0,9665
SL1344_1429	dmsA1	4,1	15,1	3,6829	0,5662
SL1344_1503	ompD	1394,7	1950,4	1,3984	0,1456

SL1344_4348	fbp	126,9	69,3	0,5461	-0,2627
SL1344_3310	nanT	4	234,3	58,5750	1,7677
SL1344_0596	ahpC	341,4	166,5	0,4877	-0,3118
SL1344_0136	secA	66,8	73,7	1,1033	0,0427
SL1344_3850	rbsC	13,8	22,9	1,6594	0,2200
SL1344_4169	malM	17,3	176	10,1734	1,0075
SL1344_4405	valS	57,7	32,3	0,5598	-0,2520
SL1344_4166	malE	16,9	175,1	10,3609	1,0154
SL1344_4168	lamB	15	179,9	11,9933	1,0789
SL1344_4457	hsdR	11,1	75,1	6,7658	0,8303
SL1344_0752	galT	9	55,6	6,1778	0,7908
SL1344_3265	glmM	28,3	55,5	1,9611	0,2925
SL1344_3848	rbsD	46,3	29,6	0,6393	-0,1943
SL1344_4263	aspA	188,7	2469,5	13,0869	1,1168
SL1344_3081	ansB	21	1041,3	49,5857	1,6954
SL1344_P2_0049	nikB	NA	NA	NA	NA
SL1344_1025	yccA	84,5	35	0,4142	-0,3828
SL1344_3450	rpe	18,1	20,2	1,1160	0,0477
SL1344_4108	yjaG	163,2	100,5	0,6158	-0,2106
SL1344_3217	tdcA	20,9	1556,3	74,4641	1,8719
SL1344_0666	nagB	35,1	86,1	2,4530	0,3897
SL1344_0487	ushA	44,6	56,7	1,2713	0,1042
SL1344_1592	ydaA	110	243,6	2,2145	0,3453
SL1344_1224A		9,4	61	6,4894	0,8122
SL1344_3851	rbsB	71,1	368,1	5,1772	0,7141
SL1344_2306	ackA	149,2	143,4	0,9611	-0,0172
SL1344_4280	frdA	10	460,4	46,0400	1,6631
SL1344_0309	dbh	37,6	15	0,3989	-0,3991
SL1344_3256	rpsO	4654,6	518	0,1113	-0,9536
SL1344_0751	galK	24,2	80,1	3,3099	0,5198
SL1344_3110	uxuB	6,4	45,6	7,1250	0,8528
SL1344_2458	yfgD	41,2	44	1,0680	0,0286
SL1344_1822	pykA	23,3	202,3	8,6824	0,9386
SL1344_2041	phsA	3,4	146	42,9412	1,6329
SL1344_1056	agp	9,5	106,6	11,2211	1,0500
SL1344_2950	sdaC	43,9	292	6,6515	0,8229
SL1344_4221	SC4B5,11c	7	69,2	9,8857	0,9950
SL1344_4186	yjbQ	65,8	64	0,9726	-0,0120
SL1344_2003	cbiH	1,2	81,2	67,6667	1,8304
SL1344_2255	glpC	11,6	579,4	49,9483	1,6985
SL1344_1582	ynaF	142,2	1095,5	7,7039	0,8867
SL1344_2347	fabB	129,9	316,5	2,4365	0,3868
SL1344_0557	ybdG	89,2	22,2	0,2489	-0,6040
SL1344_2521	cadA	7,7	782,2	101,5844	2,0068

SL1344_2098	dcd	29,6	22,6	0,7635	-0,1172
SL1344_P2_0006	yadA	NA	NA	NA	NA
SL1344_4164	malF	3,1	18,6	6,0000	0,7782
SL1344_3215	tdcC	7,8	1264,6	162,1282	2,2099
SL1344_4361	iolC	4,6	2,9	0,6304	-0,2004
SL1344_4079	btuB	44,4	55,2	1,2432	0,0946
SL1344_2663		0,7	1	1,4286	0,1549
SL1344_2254	glpB	6	435,9	72,6500	1,8612
SL1344_2253	glpA	7,4	526,2	71,1081	1,8519
SL1344_3852	rbsK	17,4	67,4	3,8736	0,5881
SL1344_4279	frdB	16,2	570,7	35,2284	1,5469
SL1344_0162	kdgT	48,2	90,8	1,8838	0,2750
SL1344_0683	speF	1,9	626,6	329,7895	2,5182
SL1344_3309	nanE	3,2	165,7	51,7813	1,7142
SL1344_1073	ycdX	61	34	0,5574	-0,2539
SL1344_2644		31,4	164,4	5,2357	0,7190
SL1344_2314		11,5	84,7	7,3652	0,8672
SL1344_0625	dacA	87,3	39,7	0,4548	-0,3422
SL1344_2307	pta	68	69,6	1,0235	0,0101
SL1344_1368	nemA	31,3	27,7	0,8850	-0,0531
SL1344_3652	mtlR	11,8	42,7	3,6186	0,5585
SL1344_2014	pduA	0	296,5	inf	inf
SL1344_2311		6,5	557,3	85,7385	1,9332
SL1344_4278	frdC	16,4	475,7	29,0061	1,4625
SL1344_0902	dmsA	1,7	241,1	141,8235	2,1517
SL1344_3859	yifE	293,6	836,1	2,8478	0,4545
SL1344_1428	dmsA2	0,8	7,2	9,0000	0,9542
SL1344_2831	hycC	1,1	2,1	1,9091	0,2808
SL1344_2011	cbiA	1,7	105,9	62,2941	1,7944
SL1344_3212	tdcG	5,5	571,4	103,8909	2,0166
SL1344_2310		4,1	406,2	99,0732	1,9960
SL1344_3299	yhbL	101,6	87,9	0,8652	-0,0629
SL1344_2312		8,7	444,9	51,1379	1,7087
SL1344_4382	nrdD	4,8	270,2	56,2917	1,7504
SL1344_2016	dhaB	0	156,3	inf	inf
SL1344_0019		3,1	2,7	0,8710	-0,0600
SL1344_3222	garL	6,9	221,2	32,0580	1,5059
SL1344_4237	fumB	6,9	183,7	26,6232	1,4253
SL1344_3208	yhaK	12,9	24,6	1,9070	0,2803
SL1344_4238	dcuB	3,9	249	63,8462	1,8051
SL1344_3124	hypO	13,1	162,1	12,3740	1,0925
SL1344_0750	galM	30,1	133,5	4,4352	0,6469
SL1344_3213	tdcE	4,5	792,4	176,0889	2,2457
SL1344_3120	hybD	11,7	200,4	17,1282	1,2337

SL1344_3674	tdh	98,2	280,9	2,8605	0,4564
SL1344_1761		22,4	489,5	21,8527	1,3395
SL1344_2260	yfaW	5,5	12,9	2,3455	0,3702
SL1344_3216	tdcB	16,7	1611,5	96,4970	1,9845
SL1344_3214	tdcD	9,7	974,6	100,4742	2,0021
SL1344_0805	glnH	154,2	86,6	0,5616	-0,2506
SL1344_2529	yfhD	12,6	8,8	0,6984	-0,1559
SL1344_0682	potE	1,7	751,9	442,2941	2,6457
SL1344_0118	fruR	38,3	128,1	3,3446	0,5244

Table 26: (Related to Fig. 7 and c; lower heatmap (b) and panels (c)). Differentially expressed genes for single bacteria and comparison of our study with reference bulk RNA-seq dataset from Kröger et al. 2013 (PMID: 24331466). Gene ID and gene name of differentially expressed genes of single bacteria in the three growth conditions (Figure 3b) are listed in column 1 and 2, respectively (color code in Column 2: blue: 'Late Stationary Phase' ; green: 'Salt (NaCl) shock'; red: 'Anaerobic shock'). Genes are ordered as in the heatmap. For each gene differentially expressed between Salt (NaCl) and Anaerobic shocks, column 3 and 4 recapitulate the gene expression extracted from Kröger et al. under Salt (NaCl) shock and Anaerobic shock, respectively. Missing values in the benchmark dataset (Kröger et al. 2013) are indicated with 'NA'. The ratio and log10 transformed ratio of (NaCl/Anaerobic) of the latter expression values from bulk-RNA-seq from Kröger et al. 2013 are calculated in columns 5 and 6, respectively.

Gene ID	Gene name	Expression values Salt (NaCl) Shock (Kröger et al) (TPM)	Expression values Anaerobic Shock (Kröger et al.) (TPM)	Ratio Salt(NaCl) /Anaerobic	Ratio Salt(NaCl) /Anaerobic (Log10)
SL1344_3301	yhcC				
SL1344_3356	prmA				
SL1344_2860	iacP				
SL1344_2847	hilC				
SL1344_1508	narZ				
SL1344_2650	ffh				
EBG00000241440	IsrM				
SL1344_4123	metH				
SL1344_3746					
SL1344_1948					
SL1344_2043	sopA				
SL1344_3106					
SL1344_2795	proW				
SL1344_4520	sthB				
SL1344_4217	nrfE				
SL1344_3709	spoU	44,2	49,6	0,8911	-0,0501

SL1344_3574	yhjE	13,5	6,5	2,0769	0,3174
SL1344_2963	ygdI	1685,4	492,6	3,4214	0,5342
SL1344_0002	thrA	34,6	13,2	2,6212	0,4185
SL1344_3843	asnCb	102,5	30,8	3,3279	0,5222
SL1344_3390	rpsE	1034,9	199,1	5,1979	0,7158
SL1344_0172	yadF	468,8	89,8	5,2205	0,7177
SL1344_4004	sodA	468,1	123,2	3,7995	0,5797
SL1344_3408	rpsJ	1883,8	302,8	6,2213	0,7939
SL1344_4093	rpoC	91	71	1,2817	0,1078
EBG00001133793	t44	NA	NA	NA	NA
SL1344_0718	sucA	73,2	75,1	0,9747	-0,0111
SL1344_3894	cyaA	155,3	92,2	1,6844	0,2264
SL1344_3890	hemY	59,8	67,1	0,8912	-0,05
SL1344_0217	rpsB	1213	178,7	6,7879	0,8317
SL1344_4172	plsB	56,2	73,3	0,7667	-0,1154
SL1344_2295	nuoC	120,4	106,2	1,1337	0,0545
SL1344_1581	nifJ	13,4	6,8	1,9706	0,2946
SL1344_4348	fbp	126,9	69,3	1,8312	0,2627
SL1344_1170	purB	17,2	7	2,4571	0,3904
SL1344_4285	psd	53,5	28,4	1,8838	0,275
SL1344_1442	dcp	10,3	19,2	0,5365	-0,2705
SL1344_2280	elaB	254,5	147,8	1,7219	0,236
SL1344_0134	lpxC	528,9	347	1,5242	0,183
SL1344_3668	grxC	362,5	285,4	1,2701	0,1039
SL1344_3833	atpG	131,9	175,4	0,752	-0,1238
SL1344_1441	ydfG	135,3	124,2	1,0894	0,0372
SL1344_0417	ispA	47,9	66,3	0,7225	-0,1412
SL1344_3932	yigC	136,3	89,6	1,5212	0,1822
SL1344_2293	nuoF	71,1	93,1	0,7637	-0,1171
SL1344_3581	yhjL	35,1	18,9	1,8571	0,2688
SL1344_4177	zur	60,4	57,5	1,0504	0,0214
SL1344_0972		1,7	3,9	0,4359	-0,3606
SL1344_0437	cyoA	448,4	37,3	12,0214	1,08
SL1344_4301	rnr	59,1	82,1	0,7199	-0,1428
SL1344_3669	yibN	280,5	195,9	1,4319	0,1559
SL1344_1323	orf319	37,6	28,1	1,3381	0,1265
SL1344_3393	rpsH	1802,1	318,1	5,6652	0,7532
SL1344_1856	cheA	228	174,7	1,3051	0,1156
SL1344_1121	flgL	272,4	227,3	1,1984	0,0786
SL1344_3392	rplF	1327,5	290,3	4,5729	0,6602
SL1344_2068	rfbG	311,5	129,1	2,4129	0,3825
SL1344_3401	rpsC	1473	206,8	7,1228	0,8527
SL1344_2267	pmrF	49	17,7	2,7684	0,4422
SL1344_4092	rpoB	94	71,9	1,3074	0,1164

SL1344_0012	dnaK	1100,5	157,9	6,9696	0,8432
SL1344_4343	ytfN	77,9	58,3	1,3362	0,1259
SL1344_0441	tig	328,9	115	2,86	0,4564
SL1344_3160	tolC	130,9	96,8	1,3523	0,1311
SL1344_3350		11	15,2	0,7237	-0,1405
SL1344_0154	lpdA	471,4	88,5	5,3266	0,7264
SL1344_0566	nfnB	75,8	47,8	1,5858	0,2002
SL1344_3196	ygjR	7,7	47,7	0,1614	-0,792
SL1344_3317	rplM	1220,7	212,6	5,7418	0,759
SL1344_0160		783,7	104,5	7,4995	0,875
SL1344_0152	aceE	283,1	41,1	6,8881	0,8381
SL1344_3405	rplW	1435,2	242,6	5,9159	0,772
SL1344_P2_0096	pilK	NA	NA	NA	NA
SL1344_2563	gipA	370,8	525,9	0,7051	-0,1518
SL1344_1057	yccJ	153,2	770	0,199	-0,7012
SL1344_2649	rps16	779,5	143,9	5,417	0,7338
SL1344_P1_0072		NA	NA	NA	NA
SL1344_P1_0055	parA	NA	NA	NA	NA
EBG00001133868	StyR-44	NA	NA	NA	NA
SL1344_4167	malK	5,2	63,4	0,082	-1,0861
EBG00000241426	STnc700	NA	NA	NA	NA
SL1344_3394	rpsN	1616,8	309	5,2324	0,7187
SL1344_2395	ptsI	170,9	386,3	0,4424	-0,3542
SL1344_3891	hemX	103,7	89,3	1,1613	0,0649
SL1344_3596	dppA	40,8	98,4	0,4146	-0,3823
SL1344_3069	yggJ	38,5	45,7	0,8425	-0,0745
SL1344_3505	glgB	38,5	45,7	0,8425	-0,0745
SL1344_1240	astD	2,2	1,7	1,2941	0,112
SL1344_3413	fusA	767,4	204,3	3,7562	0,5748
SL1344_3395	rplE	3185,3	609,7	5,2244	0,718
SL1344_0214	dapD	186,5	190,1	0,9811	-0,0083
SL1344_3045	pgk	113,8	324,6	0,3506	-0,4552
SL1344_4085	tufB	388,3	310,6	1,2502	0,097
SL1344_4347	ppa	964	526	1,8327	0,2631
SL1344_4324	rpsF	330,1	84,7	3,8973	0,5908
SL1344_3694	rpmB	3495,1	595,6	5,8682	0,7685
SL1344_1680	adh	21,2	435,8	0,0486	-1,313
SL1344_2464	ppk	70,5	56,8	1,2412	0,0938
SL1344_1168	phoQ	143,1	57	2,5105	0,3998
SL1344_1732	ycgB	42,3	28,5	1,4842	0,1715
SL1344_0802	ybiO	14,7	3,5	4,2	0,6232
SL1344_0480	htpG	1057,3	109,6	9,6469	0,9844
SL1344_3984	fdoG	86	65,4	1,315	0,1189
SL1344_3926	aarF	101,2	86,7	1,1672	0,0672

SL1344_3387	prlA		1823	486,1	3,7503	0,5741
SL1344_0868	ybjP		59,8	45,6	1,3114	0,1177
SL1344_3391	rl18		1798,9	388,7	4,628	0,6654
SL1344_1234	gdhA		48,2	39,4	1,2234	0,0876
SL1344_0866	artI		32,1	46	0,6978	-0,1563
SL1344_1245	nadE		139,7	81,4	1,7162	0,2346
SL1344_2183	fruB		2,3	3,5	0,6571	-0,1823
SL1344_1253	katE		23,5	20,3	1,1576	0,0636
SL1344_2698			511,9	620,4	0,8251	-0,0835
SL1344_0653	gltJ		22,3	3,9	5,7179	0,7572
SL1344_0712	gltA		638,9	108,2	5,9048	0,7712
SL1344_1813	opdB		21,3	17,8	1,1966	0,078
SL1344_0808	ompX		2073,6	547,7	3,786	0,5782
SL1344_P1_0084	repA2		NA	NA	NA	NA
SL1344_3876	rho		276,8	158,2	1,7497	0,243
SL1344_0617	cspE		2444,1	5682,8	0,4301	-0,3664
SL1344_0918	rpsA		698,5	237,6	2,9398	0,4683
SL1344_0459	ybaY		266,1	44,8	5,9397	0,7738
SL1344_3493	glpD		1131,1	698,3	1,6198	0,2095
SL1344_2300	yfbQ		9,4	16,4	0,5732	-0,2417
SL1344_3478	yhgI		435,5	43,6	9,9885	0,9995
SL1344_0929	mukF		43,4	29,3	1,4812	0,1706
SL1344_4267	groEL		1361,7	282,1	4,827	0,6837
SL1344_2268	yfbG		25,4	12,5	2,032	0,3079
SL1344_1328	ssaC		7,1	0,2	35,5	1,5502
SL1344_4339	ytfK		6191,4	1596,9	3,8771	0,5885
SL1344_2436	tal		38,3	8,4	4,5595	0,6589
SL1344_4526	yjjY		8,5	26,3	0,3232	-0,4905
SL1344_1267	thrS		222,2	198,1	1,1217	0,0499
SL1344_3312	yhcK		43,5	56,2	0,774	-0,1112
SL1344_0792	ybhS		69,5	32,2	2,1584	0,3341
SL1344_0159	acnB		190,3	69,2	2,75	0,4393
SL1344_0336			99,8	25,4	3,9291	0,5943
SL1344_3824	pstS		481,4	12,6	38,2063	1,5821
SL1344_4069	ppc		43,4	66,7	0,6507	-0,1866
SL1344_0735	ybgR		18,5	22,8	0,8114	-0,0908
SL1344_3412	tufA		441,9	129,7	3,4071	0,5324
SL1344_2364	pgtA		5,2	15,1	0,3444	-0,463
SL1344_0005	yaaA		62,3	50,7	1,2288	0,0895
SL1344_0149			4,6	4,3	1,0698	0,0293
SL1344_2641			1,1	0,9	1,2222	0,0872
SL1344_3675	kbl		52,8	143,3	0,3685	-0,4336
EBG00001133739	tmRNA	NA	NA	NA	NA	

SL1344_1888	fliC	2532,9	3334	1,3163	0,1193
SL1344_3556	uspA	1048,7	1147	1,0937	0,0389
SL1344_2394	ptsH	200,1	399,4	1,9960	0,3002
SL1344_2756	fljB	29,9	34,1	1,1405	0,0571
SL1344_4168	lamB	15	179,9	11,9933	1,0789
SL1344_2237	ompC	739	835	1,1299	0,0530
SL1344_2785	ygaC	289,8	11,8	0,0407	-1,3902
SL1344_2246	nrdA	66,5	102,6	1,5429	0,1883
SL1344_0753	galE	23,8	73,4	3,0840	0,4891
SL1344_3875	trxA	1663,3	477	0,2868	-0,5425
SL1344_3850	rbsC	13,8	22,9	1,6594	0,2200
SL1344_3223	garD	1,4	42,7	30,5000	1,4843
SL1344_3625	xylB	3,9	11,8	3,0256	0,4808
SL1344_2283		230,3	527,6	2,2909	0,3600
SL1344_2517	glyA	72,3	156,7	2,1674	0,3359
SL1344_0997	pepN	28,6	59,2	2,0699	0,3160
SL1344_3023	ygfY	99,6	126,2	1,2671	0,1028
SL1344_3403	rpsS	2990,4	1010,8	0,3380	-0,4711
SL1344_2521	cadA	7,7	782,2	101,5844	2,0068
SL1344_0752	galT	9	55,6	6,1778	0,7908
SL1344_1113	flgD	52,4	27,8	0,5305	-0,2753
SL1344_2252	glpT	63,7	535,1	8,4003	0,9243
SL1344_1854	cheM	264,7	267,8	1,0117	0,0051
SL1344_2105	yegO	3	3,4	1,1333	0,0544
SL1344_0654	gltI	59,4	12,5	0,2104	-0,6769
SL1344_4090	rplJ	1267,1	576,3	0,4548	-0,3422
SL1344_1707	ipk	69,2	64,7	0,9350	-0,0292
SL1344_0209	dgt	34,5	21,8	0,6319	-0,1994
SL1344_4221	SC4B5,11c	7	69,2	9,8857	0,9950
SL1344_2310		4,1	406,2	99,0732	1,9960
SL1344_0529	ppiB	575,3	169,3	0,2943	-0,5312
SL1344_0751	galK	24,2	80,1	3,3099	0,5198
SL1344_4237	fumB	6,9	183,7	26,6232	1,4253
SL1344_2931	eno	215,7	300,7	1,3941	0,1443
SL1344_3467	pckA	77,6	476,5	6,1405	0,7882
SL1344_3169	yqiC	167,6	163,2	0,9737	-0,0116
SL1344_4238	dcuB	3,9	249	63,8462	1,8051
SL1344_0193	fhuC	11,1	3,6	0,3243	-0,4890
SL1344_4263	aspA	188,7	2469,5	13,0869	1,1168
SL1344_3031	gcvT	53,4	174,7	3,2715	0,5148
SL1344_3112		164	134,2	0,8183	-0,0871
SL1344_1902	fliJ	13,4	23,1	1,7239	0,2365
SL1344_3066	galP	35,8	170,6	4,7654	0,6781
SL1344_4280	frdA	10	460,4	46,0400	1,6631

SL1344_1114	flgE	72,7	28,1	0,3865	-0,4128
SL1344_4262	dcuA	23,7	353	14,8945	1,1730
SL1344_3577	kdgK	39,2	120,3	3,0689	0,4870
SL1344_1537		33,1	125,7	3,7976	0,5795
SL1344_1708	prs	50,9	48,2	0,9470	-0,0237
SL1344_2311		6,5	557,3	85,7385	1,9332
SL1344_0683	speF	1,9	626,6	329,7895	2,5182
SL1344_0612	citC	0,3	1,9	6,3333	0,8016
SL1344_2253	glpA	7,4	526,2	71,1081	1,8519
SL1344_3851	rbsB	71,1	368,1	5,1772	0,7141
SL1344_3905	uvrD	41,8	31,6	0,7560	-0,1215
SL1344_3215	tdcC	7,8	1264,6	162,1282	2,2099
SL1344_2041	phsA	3,4	146	42,9412	1,6329
SL1344_3124	hypO	13,1	162,1	12,3740	1,0925
SL1344_3121	hybC	14,3	131	9,1608	0,9619
SL1344_1503	ompD	1394,7	1950,4	1,3984	0,1456
SL1344_0610	citE	0	1,1	inf	inf
SL1344_3214	tdcD	9,7	974,6	100,4742	2,0021

6. Bibliography:

---

1. World Health Organization. WHO Coronavirus (COVID-19) Dashboard. https://covid19.who.int/table.
2. Morens, D. M. & Fauci, A. S. Emerging Pandemic Diseases: How We Got to COVID-19. *Cell* **182**, 1077–1092 (2020).
3. WHO | Press release. *WHO* (2013).
4. Shuman, E. K. & Malani, P. N. Infectious Diseases Mortality in the United States. *JAMA* **319**, 1205 (2018).
5. Crump, J. A., Luby, S. P. & Mintz, E. D. The global burden of typhoid fever. *Bull. World Health Organ.* **82**, 346–53 (2004).
6. DE CESARE, A., SHELDON, B. W., SMITH, K. S. & JAYKUS, L.-A. Survival and Persistence of Campylobacter and Salmonella Species under Various Organic Loads on Food Contact Surfaces†. *J. Food Prot.* **66**, 1587–1594 (2003).
7. Kumar, Sg. *et al*. Antimicrobial resistance in India: A review. *J. Nat. Sci. Biol. Med.* (2013) doi:10.4103/0976-9668.116970.
8. Jones, K. E. *et al*. Global trends in emerging infectious diseases. *Nature* **451**, 990–993 (2008).
9. Ahmed, N., Dobrindt, U., Hacker, J. & Hasnain, S. E. Genomic fluidity and pathogenic bacteria: applications in diagnostics, epidemiology and intervention. *Nat. Rev. Microbiol.* **6**, 387–394 (2008).
10. Keim, P. S. & Wagner, D. M. Humans and evolutionary and ecological forces shaped the phylogeography of recently emerged diseases. *Nature Reviews Microbiology* (2009) doi:10.1038/nrmicro2219.
11. Sandegren, L. & Andersson, D. I. Bacterial gene amplification: implications for the evolution of antibiotic resistance. *Nat. Rev. Microbiol.* **7**, 578–588 (2009).
12. Hede, K. Antibiotic resistance: An infectious arms race. *Nature* (2014) doi:10.1038/509S2a.
13. Miethke, M. *et al*. Towards the sustainable discovery and development of new antibiotics. *Nat. Rev. Chem.* **5**, 726–749 (2021).
14. Towse, A. *et al*. Time for a change in how new antibiotics are reimbursed: Development of an insurance framework for funding new antibiotics based on a policy

of risk mitigation. *Health Policy* **121**, 1025–1030 (2017).

15. Plackett, B. Why big pharma has abandoned antibiotics. *Nature* **586**, S50–S52 (2020).
16. Damase, T. R. *et al*. The Limitless Future of RNA Therapeutics. *Front. Bioeng. Biotechnol.* **9**, 161 (2021).
17. Eng, S.-K. *et al*. *Salmonella* : A review on pathogenesis, epidemiology and antibiotic resistance. *Front. Life Sci.* **8**, 284–293 (2015).
18. Majowicz, S. E. *et al*. The global burden of nontyphoidal Salmonella gastroenteritis. *Clin. Infect. Dis.* **50**, 882–9 (2010).
19. Brenner, F. W., Villar, R. G., Angulo, F. J., Tauxe, R. & Swaminathan, B. Salmonella nomenclature. *J. Clin. Microbiol.* **38**, 2465–7 (2000).
20. Popoff, M.-Y., Bockemühl, J. & Brenner, F. W. Supplement 1998 (no. 42) to the Kauffmann-White scheme. *Res. Microbiol.* **151**, 63–65 (2000).
21. House, D., Bishop, A., Parry, C., Dougan, G. & Wain, J. Typhoid fever: pathogenesis and disease. *Curr. Opin. Infect. Dis.* **14**, 573–8 (2001).
22. Parry, C. M., Hien, T. T., Dougan, G., White, N. J. & Farrar, J. J. Typhoid fever. *N. Engl. J. Med.* **347**, 1770–82 (2002).
23. Hohmann, E. L. Nontyphoidal salmonellosis. *Clin. Infect. Dis.* **32**, 263–9 (2001).
24. Silva, J. *et al*. Campylobacter spp. as a Foodborne Pathogen: A Review. *Front. Microbiol.* **2**, 200 (2011).
25. CH, C. *et al*. The emergence in Taiwan of fluoroquinolone resistance in Salmonella enterica serotype choleraesuis. *N. Engl. J. Med.* **346**, (2002).
26. HYEON, J.-Y. *et al*. Prevalence, Antibiotic Resistance, and Molecular Characterization of Salmonella Serovars in Retail Meat Products. *J. Food Prot.* **74**, 161–166 (2011).
27. HARRIS, J. C. Fecal Leukocytes in Diarrheal Illness. *Ann. Intern. Med.* **76**, 697 (1972).
28. Winter, S. E. *et al*. Gut inflammation provides a respiratory electron acceptor for Salmonella. *Nature* **467**, 426–9 (2010).
29. Stecher, B. *et al*. Salmonella enterica Serovar Typhimurium Exploits Inflammation to Compete with the Intestinal Microbiota. *PLoS Biol.* **5**, e244 (2007).
30. Jones, B. D., Ghori, N. & Falkow, S. Salmonella typhimurium initiates murine infection by penetrating and destroying the specialized epithelial M cells of the Peyer's patches. *J. Exp. Med.* **180**, 15–23 (1994).
31. Niedergang, F., Sirard, J. C., Blanc, C. T. & Kraehenbuhl, J. P. Entry and survival of Salmonella typhimurium in dendritic cells and presentation of recombinant antigens do

not require macrophage-specific virulence factors. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 14650–5 (2000).

32. Martínez-Moya, M., de Pedro, M. A., Schwarz, H. & García-del Portillo, F. Inhibition of Salmonella intracellular proliferation by non-phagocytic eucaryotic cells. *Res. Microbiol.* **149**, 309–18 (1998).
33. Dalebroux, Z. D. & Miller, S. I. Salmonellae PhoPQ regulation of the outer membrane to resist innate immunity. *Curr. Opin. Microbiol.* **17**, 106–13 (2014).
34. Haraga, A., Ohlson, M. B. & Miller, S. I. Salmonellae interplay with host cells. *Nat. Rev. Microbiol.* **6**, 53–66 (2008).
35. Steele-Mortimer, O. *et al*. The invasion-associated type III secretion system of Salmonella enterica serovar Typhimurium is necessary for intracellular proliferation and vacuole biogenesis in epithelial cells. *Cell. Microbiol.* **4**, 43–54 (2002).
36. Zhou, D., Chen, L. M., Hernandez, L., Shears, S. B. & Galán, J. E. A Salmonella inositol polyphosphatase acts in conjunction with other bacterial effectors to promote host cell actin cytoskeleton rearrangements and bacterial internalization. *Mol. Microbiol.* **39**, 248–59 (2001).
37. Dunn, J. D. & Valdivia, R. H. Uncivil engineers: Chlamydia, Salmonella and Shigella alter cytoskeleton architecture to invade epithelial cells. *Future Microbiol.* **5**, 1219–32 (2010).
38. Hernandez, L. D., Hueffer, K., Wenk, M. R. & Galán, J. E. Salmonella modulates vesicular traffic by altering phosphoinositide metabolism. *Science* **304**, 1805–7 (2004).
39. Norris, F. A., Wilson, M. P., Wallis, T. S., Galyov, E. E. & Majerus, P. W. SopB, a protein required for virulence of Salmonella dublin, is an inositol phosphate phosphatase. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 14057–9 (1998).
40. Tahoun, A. *et al*. Salmonella transforms follicle-associated epithelial cells into M cells to promote intestinal invasion. *Cell Host Microbe* **12**, 645–56 (2012).
41. Bakowski, M. A. *et al*. The phosphoinositide phosphatase SopB manipulates membrane surface charge and trafficking of the Salmonella-containing vacuole. *Cell Host Microbe* **7**, 453–62 (2010).
42. McGhie, E. J., Brawn, L. C., Hume, P. J., Humphreys, D. & Koronakis, V. Salmonella takes control: effector-driven manipulation of the host. *Curr. Opin. Microbiol.* **12**, 117–24 (2009).
43. Nichols, C. D. & Casanova, J. E. Salmonella-Directed Recruitment of New Membrane to Invasion Foci via the Host Exocyst Complex. *Curr. Biol.* **20**, 1316–1320 (2010).

44. Powers, T. R. *et al*. Intracellular niche-specific profiling reveals transcriptional adaptations required for the cytosolic lifestyle of Salmonella enterica. *PLOS Pathog*. **17**, e1009280 (2021).
45. Abrahams, G. L., Müller, P. & Hensel, M. Functional Dissection of SseF, a Type III Effector Protein Involved in Positioning the Salmonella-Containing Vacuole. *Traffic* **7**, 950–965 (2006).
46. Waterman, S. R. & Holden, D. W. Functions and effectors of the Salmonella pathogenicity island 2 type III secretion system. *Cell. Microbiol*. **5**, 501–11 (2003).
47. Malik-Kale, P. *et al*. Salmonella – At Home in the Host Cell. *Front. Microbiol*. **2**, 125 (2011).
48. Knodler, L. A. *et al*. Dissemination of invasive Salmonella via bacterial-induced extrusion of mucosal epithelia. *Proc. Natl. Acad. Sci. U. S. A*. **107**, 17733–8 (2010).
49. Ackermann, M. A functional perspective on phenotypic heterogeneity in microorganisms. *Nat. Rev. Microbiol*. **13**, 497–508 (2015).
50. Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic gene expression in a single cell. *Science* **297**, 1183–6 (2002).
51. Weigel, W. A. & Dersch, P. Phenotypic heterogeneity: a bacterial virulence strategy. *Microbes Infect*. **20**, 570–577 (2018).
52. Ackermann, M. A functional perspective on phenotypic heterogeneity in microorganisms. *Nat. Rev. Microbiol*. **13**, 497–508 (2015).
53. Bumann, D. Heterogeneous host-pathogen encounters: act locally, think globally. *Cell Host Microbe* **17**, 13–9 (2015).
54. Raj, A. & van Oudenaarden, A. Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences. *Cell* **135**, 216–226 (2008).
55. Davis, K. M., Mohammadi, S. & Isberg, R. R. Community behavior and spatial regulation within a bacterial microcolony in deep tissue sites serves to protect against host attack. *Cell Host Microbe* **17**, 21–31 (2015).
56. Stewart, M. K. & Cookson, B. T. Non-genetic diversity shapes infectious capacity and host resistance. *Trends Microbiol*. **20**, 461–6 (2012).
57. Nuss, A. M. *et al*. A Precise Temperature-Responsive Bistable Switch Controlling Yersinia Virulence. *PLOS Pathog*. **12**, e1006091 (2016).
58. Gollan, B., Grabe, G., Michaux, C. & Helaine, S. Bacterial Persisters and Infection: Past, Present, and Progressing. *Annu. Rev. Microbiol*. **73**, 359–385 (2019).
59. Abshire, K. Z. & Neidhardt, F. C. Growth rate paradox of Salmonella typhimurium

within host macrophages. *J. Bacteriol.* **175**, 3744–8 (1993).

60. Helaine, S. *et al.* Internalization of Salmonella by Macrophages Induces Formation of Nonreplicating Persisters. *Science (80-.).* **343**, 204–208 (2014).
61. S Helaine, E. K. Bacterial persisters: formation, eradication, and experimental systems. *Trends Microbiol.* **22**, 417–424 (2014).
62. A Harms, E. M. K. G. Mechanisms of bacterial persistence during stress and antibiotic exposure. *Science (80-.).* **354**, aaf4268 (2016).
63. JE Michiels, B. V. den B. N. V. J. M. Molecular mechanisms and clinical implications of bacterial persistence. *Drug Resist. Updat.* **29**, 76–89 (2017).
64. Monack, D. M. Helicobacter and salmonella persistent infection strategies. *Cold Spring Harb. Perspect. Med.* **3**, a010348 (2013).
65. Fisher, R. A., Gollan, B. & Helaine, S. Persistent bacterial infections and persister cells. *Nat. Rev. Microbiol.* **15**, 453–464 (2017).
66. Hill, P. W. S. & Helaine, S. Antibiotic Persisters and Relapsing Salmonella enterica Infections. in *Persister Cells and Infectious Disease* 19–38 (Springer International Publishing, 2019). doi:10.1007/978-3-030-25241-0_2.
67. Fisher, R. A. *et al.* Salmonella persisters undermine host immune defenses during antibiotic treatment. **1160**, 1156–1160 (2018).
68. Cohen, N. R., Lobritz, M. A. & Collins, J. J. Microbial persistence and the road to drug resistance. *Cell Host Microbe* **13**, 632–42 (2013).
69. Levin-Reisman, I. *et al.* Antibiotic tolerance facilitates the evolution of resistance. *Science (80-.).* **355**, 826–830 (2017).
70. Avershina, E., Shapovalova, V. & Shipulin, G. Fighting Antibiotic Resistance in Hospital-Acquired Infections: Current State and Emerging Technologies in Disease Prevention, Diagnostics and Therapy. *Front. Microbiol.* **12**, 2044 (2021).
71. Bakkeren, E. *et al.* Salmonella persisters promote the spread of antibiotic resistance plasmids in the gut. *Nature* **573**, 276–280 (2019).
72. Moyed, H. S. & Bertrand, K. P. hipA, a newly recognized gene of Escherichia coli K-12 that affects frequency of persistence after inhibition of murein synthesis. *J. Bacteriol.* **155**, 768–775 (1983).
73. Stapels, D. A. C. *et al. Salmonella persisters undermine host immune defenses during antibiotic treatment.* http://science.sciencemag.org/.
74. Helaine, S. *et al.* Dynamics of intracellular bacterial replication at the single cell level. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 3746–51 (2010).

75. Shah, D. *et al*. Persisters: a distinct physiological state of E. coli. (2006) doi:10.1186/1471-2180-6-53.
76. Wang, Z., Gerstein, M. & Snyder, M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet*. **10**, 57–63 (2009).
77. E Maisonneuve, M. C.-C. K. G. (p)ppGpp controls bacterial persistence by stochastic induction of toxin–antitoxin activity. *Cell* **154**, 1140–1150 (2013).
78. Amato, S. The role of metabolism in bacterial persistence. *Front. Microbiol*. **5**, 70 (2014).
79. T Dörr, M. V. K. L. Ciprofloxacin causes persister formation by inducing the TisB toxin in Escherichia coli. *PLoS Biol*. **8**, e1000317 (2010).
80. Saliba, A.-E. *et al*. Single-cell RNA-seq ties macrophage polarization to growth rate of intracellular Salmonella. *Nat. Microbiol*. **2**, 16206 (2017).
81. AL Spoering, K. L. Biofilms and planktonic cells of Pseudomonas aeruginosa have similar resistance to killing by antimicrobials. *J. Bacteriol*. **183**, 6746–6751 (2001).
82. Imdahl, F., Vafadarnejad, E., Homberger, C., Saliba, A.-E. & Vogel, J. Single-cell RNA-sequencing reports growth-condition-specific global transcriptomes of individual bacteria. *Nat. Microbiol*. **5**, 1202–1206 (2020).
83. Blattman, S. B., Jiang, W., Oikonomou, P. & Tavazoie, S. Prokaryotic single-cell RNA sequencing by in situ combinatorial indexing. *Nat. Microbiol*. **5**, 1192–1201 (2020).
84. Kuchina, A. *et al*. Microbial single-cell RNA sequencing by split-pool barcoding. *Science* **371**, (2021).
85. Gest, H. The discovery of microorganisms by Robert Hooke and Antoni Van Leeuwenhoek, fellows of the Royal Society. *Notes Rec. R. Soc. Lond*. **58**, 187–201 (2004).
86. Arendt, D. *et al*. The origin and evolution of cell types. *Nat. Rev. Genet*. **17**, 744–757 (2016).
87. Orkin, S. H. Diversification of haematopoietic stem cells to specific lineages. *Nat. Rev. Genet*. **1**, 57–64 (2000).
88. Mosmann, T. R., Cherwinski, H., Bond, M. W., Giedlin, M. A. & Coffman, R. L. Two types of murine helper T cell clone. I. Definition according to profiles of lymphokine activities and secreted proteins. 1986. *J. Immunol*. **175**, 5–14 (2005).
89. Kelso, A. *et al*. Heterogeneity in Lymphokine Profiles of CD4 $^{+}$ and CD8 $^{+}$ T Cells and Clones Activated in vivo and in vitro. *Immunol. Rev*. **123**, 85–114 (1991).
90. Garbis, S., Lubec, G. & Fountoulakis, M. Limitations of current proteomics

technologies. *J. Chromatogr. A* **1077**, 1–18 (2005).

91. Westermann, A. J. & Vogel, J. Cross-species RNA-seq for deciphering host–microbe interactions. *Nat. Rev. Genet.* **22**, 361–378 (2021).
92. Svensson, V., Vento-Tormo, R. & Teichmann, S. A. Exponential scaling of single-cell RNA-seq in the past decade. *Nat. Protoc.* **13**, 599–604 (2018).
93. Westermann, A. J. & Vogel, J. Cross-species RNA-seq for deciphering host–microbe interactions. *Nat. Rev. Genet.* (2021) doi:10.1038/s41576-021-00326-y.
94. Westermann, A. J., Gorski, S. A. & Vogel, J. Dual RNA-seq of pathogen and host. *Nat. Rev. Microbiol.* **10**, 618–630 (2012).
95. 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 (2014).
96. Penaranda, C. & Hung, D. T. Single-Cell RNA Sequencing to Understand Host-Pathogen Interactions. *ACS Infect. Dis.* **5**, 336–344 (2019).
97. Wang, Z., Gerstein, M. & Snyder, M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **10**, 57–63 (2009).
98. Costa, V., Aprile, M., Esposito, R. & Ciccodicola, A. RNA-Seq and human complex diseases: recent accomplishments and future perspectives. *Eur. J. Hum. Genet.* **21**, 134–142 (2013).
99. Bernard, G., Pathmanathan, J. S., Lannes, R., Lopez, P. & Bapteste, E. Microbial Dark Matter Investigations: How Microbial Studies Transform Biological Knowledge and Empirically Sketch a Logic of Scientific Discovery. *Genome Biol. Evol.* **10**, 707–715 (2018).
100. Tang, F. *et al*. mRNA-Seq whole-transcriptome analysis of a single cell. *Nat. Methods* **6**, 377–382 (2009).
101. Ramsköld, D. *et al*. Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat. Biotechnol.* **30**, 777–82 (2012).
102. Islam, S. *et al*. Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. *Genome Res.* **21**, 1160–7 (2011).
103. Hashimshony, T., Wagner, F., Sher, N. & Yanai, I. CEL-Seq: Single-Cell RNA-Seq by Multiplexed Linear Amplification. *Cell Rep.* **2**, 666–673 (2012).
104. Klein, A. M. *et al*. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* **161**, 1187–1201 (2015).
105. Macosko, E. Z. *et al*. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* **161**, 1202–1214 (2015).

106. Zheng, G. X. Y. *et al*. Massively parallel digital transcriptional profiling of single cells. *Nat. Commun*. **8**, 14049 (2017).
107. Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. *Nature* **562**, 367–372 (2018).
108. Consortium, T. T. S. & Quake, S. R. The Tabula Sapiens: a single cell transcriptomic atlas of multiple organs from individual human donors. *bioRxiv* 2021.07.19.452956 (2021) doi:10.1101/2021.07.19.452956.
109. Isakova, A., Neff, N. & Quake, S. R. Single cell profiling of total RNA using Smart-seq-total. *bioRxiv* 2020.06.02.131060 (2020) doi:10.1101/2020.06.02.131060.
110. Salmen, F. *et al*. Droplet-based Single-cell Total RNA-seq Reveals Differential Non-Coding Expression and Splicing Patterns during Mouse Development. *bioRxiv* 2021.09.15.460240 (2021) doi:10.1101/2021.09.15.460240.
111. Sheng, K., Cao, W., Niu, Y., Deng, Q. & Zong, C. Effective detection of variation in single-cell transcriptomes using MATQ-seq. *Nat. Methods* (2017) doi:10.1038/nmeth.4145.
112. Seelbinder, B. *et al*. Triple RNA-Seq Reveals Synergy in a Human Virus-Fungus Co-infection Model. *Cell Rep*. **33**, 108389 (2020).
113. Imdahl, F. & Saliba, A.-E. Advances and challenges in single-cell RNA-seq of microbial communities. *Curr. Opin. Microbiol*. **57**, 102–110 (2020).
114. Stark, R., Grzelak, M. & Hadfield, J. RNA sequencing: the teenage years. *Nat. Rev. Genet*. **20**, 631–656 (2019).
115. Sorek, R. & Cossart, P. Prokaryotic transcriptomics: a new view on regulation, physiology and pathogenicity. *Nat. Rev. Genet*. **11**, 9–16 (2010).
116. Bhattacharyya, R. P. *et al*. Simultaneous detection of genotype and phenotype enables rapid and accurate antibiotic susceptibility determination. *Nat. Med*. **25**, 1858–1864 (2019).
117. Picelli, S. *et al*. Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc*. **9**, 171–81 (2014).
118. Regev, A. *et al*. The Human Cell Atlas. *Elife* **6**, (2017).
119. Rajewsky, N. *et al*. LifeTime and improving European healthcare through cell-based interceptive medicine. *Nature* **587**, 377–386 (2020).
120. Tanay, A. & Regev, A. Scaling single-cell genomics from phenomenology to mechanism. *Nature* **541**, 331–338 (2017).
121. Gilbert, J. A. *et al*. Current understanding of the human microbiome. *Nat. Med*. **24**,

392–400 (2018).

122. Sender, R., Fuchs, S. & Milo, R. Are We Really Vastly Outnumbered? Revisiting the Ratio of Bacterial to Host Cells in Humans. *Cell* **164**, 337–40 (2016).
123. Turnbaugh, P. J. *et al*. The Human Microbiome Project. *Nature* **449**, 804–810 (2007).
124. Iliev, I. D. & Leonardi, I. Fungal dysbiosis: immunity and interactions at mucosal barriers. *Nat. Rev. Immunol.* **17**, 635–646 (2017).
125. Kundu, P., Blacher, E., Elinav, E. & Pettersson, S. Our Gut Microbiome: The Evolving Inner Self. *Cell* **171**, 1481–1493 (2017).
126. Thion, M. S. *et al*. Microbiome Influences Prenatal and Adult Microglia in a Sex-Specific Manner. *Cell* **172**, 500-516.e16 (2018).
127. Fulde, M. *et al*. Neonatal selection by Toll-like receptor 5 influences long-term gut microbiota composition. *Nature* **560**, 489–493 (2018).
128. Schirmer, M. *et al*. Dynamics of metatranscription in the inflammatory bowel disease gut microbiome. *Nat. Microbiol.* **3**, 337–346 (2018).
129. Freed, N. E. *et al*. A Simple Screen to Identify Promoters Conferring High Levels of Phenotypic Noise. *PLoS Genet.* **4**, e1000307 (2008).
130. Balaban, N. Q., Merrin, J., Chait, R., Kowalik, L. & Leibler, S. Bacterial persistence as a phenotypic switch. *Science* **305**, 1622–5 (2004).
131. Rosenthal, A. Z. *et al*. Metabolic interactions between dynamic bacterial subpopulations. *Elife* **7**, (2018).
132. Single-cell microbiology. *Nat. Biotechnol.* **34**, 1077–1077 (2016).
133. Hatzenpichler, R., Krukenberg, V., Spietz, R. L. & Jay, Z. J. Next-generation physiology approaches to study microbiome function at single cell level. *Nat. Rev. Microbiol.* **18**, 241–256 (2020).
134. Kreibich, S. & Hardt, W.-D. Experimental approaches to phenotypic diversity in infection. *Curr. Opin. Microbiol.* **27**, 25–36 (2015).
135. Hug, L. A. *et al*. A new view of the tree of life. *Nat. Microbiol.* **1**, 16048 (2016).
136. Castelle, C. J. & Banfield, J. F. Major New Microbial Groups Expand Diversity and Alter our Understanding of the Tree of Life. *Cell* **172**, 1181–1197 (2018).
137. Rinke, C. *et al*. Insights into the phylogeny and coding potential of microbial dark matter. *Nature* **499**, 431–437 (2013).
138. Woyke, T., Doud, D. F. R. & Schulz, F. The trajectory of microbial single-cell sequencing. *Nat. Methods* **14**, 1045–1054 (2017).
139. Imdahl, F., Vafadarnejad, E., Homberger, C., Saliba, A.-E. & Vogel, J. Single-cell

RNA-sequencing reports growth-condition-specific global transcriptomes of individual bacteria. *Nat. Microbiol.* doi:10.1038/s41564-020-0774-1.

140. Taniguchi, Y. *et al.* Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells. *Science (80-.).* (2010) doi:10.1126/science.1188308.
141. Maamar, H., Raj, A. & Dubnau, D. Noise in gene expression determines cell fate in Bacillus subtilis. *Science* **317**, 526–9 (2007).
142. Marguerat, S. *et al.* Quantitative analysis of fission yeast transcriptomes and proteomes in proliferating and quiescent cells. *Cell* **151**, 671–83 (2012).
143. Bagnoli, J. W. *et al.* Sensitive and powerful single-cell RNA sequencing using mcSCRB-seq. *Nat. Commun.* **9**, 2937 (2018).
144. Ziegenhain, C. *et al.* Comparative Analysis of Single-Cell RNA Sequencing Methods. *Mol. Cell* **65**, 631-643.e4 (2017).
145. Claudi, B. *et al.* Phenotypic variation of salmonella in host tissues delays eradication by antimicrobial chemotherapy. *Cell* (2014) doi:10.1016/j.cell.2014.06.045.
146. Stapels, D. A. C. *et al.* Salmonella persisters undermine host immune defenses during antibiotic treatment. *Science (80-.).* **362**, 1156–1160 (2018).
147. Burton, N. A. *et al.* Disparate impact of oxidative host defenses determines the fate of Salmonella during systemic infection in mice. *Cell Host Microbe* **15**, 72–83 (2014).
148. Roche, B. & Bumann, D. Single-cell reporters for pathogen responses to antimicrobial host attacks. *Curr. Opin. Microbiol.* **59**, 16–23 (2021).
149. Valentini, T. D. *et al.* Bioorthogonal non-canonical amino acid tagging reveals translationally active subpopulations of the cystic fibrosis lung microbiota. *Nat. Commun.* **11**, 2287 (2020).
150. Lee, K. S. *et al.* An automated Raman-based platform for the sorting of live cells by functional properties. *Nat. Microbiol.* **4**, 1035–1048 (2019).
151. Luro, S., Potvin-Trottier, L., Okumus, B. & Paulsson, J. Isolating live cells after high-throughput, long-term, time-lapse microscopy. *Nat. Methods* **17**, 93–100 (2020).
152. Watterson, W. J. *et al.* Droplet-based high-throughput cultivation for accurate screening of antibiotic resistant gut microbes. *Elife* **9**, (2020).
153. Wu, A. R. *et al.* Quantitative assessment of single-cell RNA-sequencing methods. *Nat. Methods* **11**, 41–46 (2014).
154. Reid, A. J. *et al.* Single-cell RNA-seq reveals hidden transcriptional variation in malaria parasites. *Elife* **7**, (2018).

155. Müller, L. S. M. *et al*. Genome organization and DNA accessibility control antigenic variation in trypanosomes. *Nature* **563**, 121–125 (2018).
156. Howick, V. M. *et al*. The Malaria Cell Atlas: Single parasite transcriptomes across the complete Plasmodium life cycle. *Science* **365**, (2019).
157. Poran, A. *et al*. Single-cell RNA sequencing reveals a signature of sexual commitment in malaria parasites. *Nature* **551**, 95–99 (2017).
158. Reuter, C. *et al*. Vector-borne Trypanosoma brucei parasites develop in artificial human skin and persist as skin tissue forms. *bioRxiv* 2021.05.13.443986 (2021) doi:10.1101/2021.05.13.443986.
159. Gasch, A. P. *et al*. Single-cell RNA sequencing reveals intrinsic and extrinsic regulatory heterogeneity in yeast responding to stress. *PLOS Biol*. **15**, e2004050 (2017).
160. Saint, M. *et al*. Single-cell imaging and RNA sequencing reveal patterns of gene expression heterogeneity during fission yeast growth and adaptation. *Nat. Microbiol*. **4**, 480–491 (2019).
161. Nadal-Ribelles, M. *et al*. Sensitive high-throughput single-cell RNA-seq reveals within-clonal transcript correlations in yeast populations. *Nat. Microbiol*. **4**, 683–692 (2019).
162. Jackson, C. A., Castro, D. M., Saldi, G.-A., Bonneau, R. & Gresham, D. Gene regulatory network reconstruction using single-cell RNA sequencing of barcoded genotypes in diverse environments. *Elife* **9**, (2020).
163. Jariani, A. *et al*. A new protocol for single-cell RNA-seq reveals stochastic gene expression during lag phase in budding yeast. *Elife* **9**, (2020).
164. Rosenberg, A. B. *et al*. Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science (80-.)*. (2018) doi:10.1126/science.aam8999.
165. Westermann, A., Westermann, A. J., Gorski, S. A. & Vogel, J. Dual RNA-Seq of pathogen and host Dual RNA-seq of pathogen and host. *Nat. Publ. Gr*. **10**, 618–630 (2012).
166. Penaranda, C. & Hung, D. T. Single-Cell RNA Sequencing to Understand Host–Pathogen Interactions. (2019) doi:10.1021/acsinfecdis.8b00369.
167. Avital, G. *et al*. scDual-Seq: mapping the gene regulatory program of Salmonella infection by host and pathogen single-cell RNA-sequencing. *Genome Biol*. **18**, 200 (2017).
168. Yanai, I. & Hashimshony, T. CEL-Seq2—Single-Cell RNA Sequencing by

Multiplexed Linear Amplification. in 45–56 (Humana, New York, NY, 2019). doi:10.1007/978-1-4939-9240-9_4.

169. Golding, I., Paulsson, J., Zawilski, S. M. & Cox, E. C. Real-Time Kinetics of Gene Activity in Individual Bacteria. *Cell* **123**, 1025–1036 (2005).
170. Marinov, G. K. *et al*. From single-cell to cell-pool transcriptomes: stochasticity in gene expression and RNA splicing. *Genome Res*. **24**, 496–510 (2014).
171. Zong, C., Lu, S., Chapman, A. R. & Xie, X. S. Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. *Science (80-.)*. (2012) doi:10.1126/science.1229164.
172. Kröger, C. *et al*. An infection-relevant transcriptomic compendium for salmonella enterica serovar typhimurium. *Cell Host Microbe* (2013) doi:10.1016/j.chom.2013.11.010.
173. Ramos, T. & Moroni, L. Tissue Engineering and Regenerative Medicine 2019: The Role of Biofabrication—A Year in Review. *Tissue Eng. Part C Methods* **26**, 91–106 (2020).
174. Fatehullah, A., Tan, S. H. & Barker, N. Organoids as an in vitro model of human development and disease. *Nat. Cell Biol*. **18**, 246–54 (2016).
175. Barrila, J. *et al*. Modeling Host-Pathogen Interactions in the Context of the Microenvironment: Three-Dimensional Cell Culture Comes of Age. *Infect. Immun*. **86**, (2018).
176. Schmid, F. F. *et al*. Applicability of a Dual-Arm Robotic System for Automated Downstream Analysis of Epidermal Models. *Appl. Vitr. Toxicol*. **2**, 118–125 (2016).
177. Mekhileri, N. V *et al*. Automated 3D bioassembly of micro-tissues for biofabrication of hybrid tissue engineered constructs. *Biofabrication* **10**, 024103 (2018).
178. Knoll, A., Scherer, T., Poggendorf, I., Lütkemeyer, D. & Lehmann, J. Flexible automation of cell culture and tissue engineering tasks. *Biotechnol. Prog*. **20**, 1825–35 (2004).
179. Lavik, E. & Langer, R. Tissue engineering: current state and perspectives. *Appl. Microbiol. Biotechnol*. **65**, 1–8 (2004).
180. Van Norman, G. A. Limitations of Animal Studies for Predicting Toxicity in Clinical Trials: Is it Time to Rethink Our Current Approach? *JACC Basic to Transl. Sci*. **4**, 845–854 (2019).
181. Doke, S. K. & Dhawale, S. C. Alternatives to animal testing: A review. *Saudi Pharm. J. SPJ Off. Publ. Saudi Pharm. Soc*. **23**, 223–9 (2015).

182. Hay, M., Thomas, D. W., Craighead, J. L., Economides, C. & Rosenthal, J. Clinical development success rates for investigational drugs. *Nat. Biotechnol.* **32**, 40–51 (2014).
183. Diavatopoulos, D. A. *et al.* Bordetella pertussis, the Causative Agent of Whooping Cough, Evolved from a Distinct, Human-Associated Lineage of B. bronchiseptica. *PLoS Pathog.* **1**, e45 (2005).
184. Burch, R.L. and Russell, W. M. S. The Principles of Humane Experimental Technique. in *Medical Journal of Australia* vol. 1 500–500 (John Wiley & Sons, Ltd, 1960).
185. Goodman, B. E. Insights into digestion and absorption of major nutrients in humans. *Adv. Physiol. Educ.* **34**, 44–53 (2010).
186. Nicholl, C. G., Polak, J. M. & Bloom, S. R. The hormonal regulation of food intake, digestion, and absorption. *Annu. Rev. Nutr.* **5**, 213–39 (1985).
187. Patel, G. Oral Delivery of Proteins and Peptides: Concepts and Applications. *Challenges Deliv. Ther. Genomics Proteomics* 481–529 (2011) doi:10.1016/B978-0-12-384964-9.00010-4.
188. Artursson, P., Neuhoff, S., Matsson, P. & Tavelin, S. Passive Permeability and Active Transport Models for the Prediction of Oral Absorption. *Compr. Med. Chem. II* 259–278 (2007) doi:10.1016/B0-08-045044-X/00126-7.
189. Furuse, M., Fujita, K., Hiiragi, T., Fujimoto, K. & Tsukita, S. Claudin-1 and -2: novel integral membrane proteins localizing at tight junctions with no sequence similarity to occludin. *J. Cell Biol.* **141**, 1539–50 (1998).
190. Lechuga, S. & Ivanov, A. I. Disruption of the epithelial barrier during intestinal inflammation: Quest for new molecules and mechanisms. *Biochim. Biophys. Acta - Mol. Cell Res.* **1864**, 1183–1194 (2017).
191. Takeichi, M. Cadherin cell adhesion receptors as a morphogenetic regulator. *Science* **251**, 1451–5 (1991).
192. Martìn-Padura, I. *et al.* Junctional adhesion molecule, a novel member of the immunoglobulin superfamily that distributes at intercellular junctions and modulates monocyte transmigration. *J. Cell Biol.* **142**, 117–27 (1998).
193. Wang, Y. *et al.* Bioengineered Systems and Designer Matrices That Recapitulate the Intestinal Stem Cell Niche. *Cell. Mol. Gastroenterol. Hepatol.* **5**, 440-453.e1 (2018).
194. Barker, N. Adult intestinal stem cells: critical drivers of epithelial homeostasis and regeneration. *Nat. Rev. Mol. Cell Biol.* **15**, 19–33 (2014).
195. Barker, N. *et al.* Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature* **449**, 1003–1007 (2007).

196. Tian, H. *et al*. A reserve stem cell population in small intestine renders Lgr5-positive cells dispensable. *Nature* **478**, 255–259 (2011).
197. Darwich, A. S., Aslam, U., Ashcroft, D. M. & Rostami-Hodjegan, A. Meta-analysis of the turnover of intestinal epithelia in preclinical animal species and humans. *Drug Metab. Dispos.* **42**, 2016–22 (2014).
198. Wang, Y. *et al*. Bioengineered Systems and Designer Matrices That Recapitulate the Intestinal Stem Cell Niche. *Cell. Mol. Gastroenterol. Hepatol.* **5**, 440-453.e1 (2018).
199. Potten, C. S. & Loeffler, M. Stem cells: attributes, cycles, spirals, pitfalls and uncertainties. Lessons for and from the crypt. *Development* **110**, 1001–1020 (1990).
200. Frisch, S. M. & Francis, H. Disruption of epithelial cell-matrix interactions induces apoptosis. *J. Cell Biol.* **124**, 619–26 (1994).
201. van der Flier, L. G. & Clevers, H. Stem Cells, Self-Renewal, and Differentiation in the Intestinal Epithelium. *Annu. Rev. Physiol.* **71**, 241–260 (2009).
202. Date, S. & Sato, T. Mini-Gut Organoids: Reconstitution of the Stem Cell Niche. *Annu. Rev. Cell Dev. Biol.* **31**, 269–289 (2015).
203. Sato, T. *et al*. Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature* **459**, 262–265 (2009).
204. Sato, T. *et al*. Long-term Expansion of Epithelial Organoids From Human Colon, Adenoma, Adenocarcinoma, and Barrett's Epithelium. *Gastroenterology* **141**, 1762–1772 (2011).
205. Farin, H. F. *et al*. Visualization of a short-range Wnt gradient in the intestinal stem-cell niche. *Nature* **530**, 340–343 (2016).
206. Battle, T. E. & Yen, A. Ectopic Expression of CXCR5/BLR1 Accelerates Retinoic Acid- and Vitamin D_3 -Induced Monocytic Differentiation of U937 Cells. *Exp. Biol. Med.* **227**, 753–762 (2002).
207. Haramis, A.-P. G. *et al*. De novo crypt formation and juvenile polyposis on BMP inhibition in mouse intestine. *Science* **303**, 1684–6 (2004).
208. Cheng, H. & Leblond, C. P. Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine I. Columnar cell. *Am. J. Anat.* **141**, 461–479 (1974).
209. Helander, H. F. & Fändriks, L. Surface area of the digestive tract – revisited. *Scand. J. Gastroenterol.* **49**, 681–689 (2014).
210. BROWN, A. L. Microvilli of the human jejunal epithelial cell. *J. Cell Biol.* **12**, 623–7 (1962).

211. Peterson, L. W. & Artis, D. Intestinal epithelial cells: regulators of barrier function and immune homeostasis. *Nat. Rev. Immunol.* **14**, 141–153 (2014).
212. Birchenough, G. M. H., Nystrom, E. E. L., Johansson, M. E. V. & Hansson, G. C. A sentinel goblet cell guards the colonic crypt by triggering Nlrp6-dependent Muc2 secretion. *Science (80-.).* **352**, 1535–1542 (2016).
213. Worthington, J. J., Reimann, F. & Gribble, F. M. Enteroendocrine cells-sensory sentinels of the intestinal environment and orchestrators of mucosal immunity. *Mucosal Immunol.* **11**, 3–20 (2018).
214. Birchenough, G. M. H., Nystrom, E. E. L., Johansson, M. E. V. & Hansson, G. C. A sentinel goblet cell guards the colonic crypt by triggering Nlrp6-dependent Muc2 secretion. *Science (80-.).* **352**, 1535–1542 (2016).
215. Järvi, O. & Keyriläinen, O. ON THE CELLULAR STRUCTURES OF THE EPITHELIAL INVASIONS IN THE GLANDULAR STOMACH OF MICE CAUSED BY INTRAMURAL APPLICATION OF 20-METHYLCHOLANTREN. *Acta Pathol. Microbiol. Scand.* **38**, 72–73 (1956).
216. Verbrugghe, P., Kujala, P., Waelput, W., Peters, P. J. & Cuvelier, C. A. Clusterin in human gut-associated lymphoid tissue, tonsils, and adenoids: localization to M cells and follicular dendritic cells. *Histochem. Cell Biol.* **129**, 311–320 (2008).
217. Sato, A. Tuft cells. *Anat. Sci. Int.* **82**, 187–99 (2007).
218. Rao, A. L. & Sankar, G. G. *Asian journal of pharmaceutical research & health care AJPRHC. Asian Journal of Pharmaceutical Research and Health Care* vol. 1 (2009).
219. Dosh, R. H., Jordan-Mahy, N., Sammon, C. & Le Maitre, C. L. Tissue Engineering Laboratory Models of the Small Intestine. *Tissue Eng. Part B. Rev.* **24**, 98–111 (2018).
220. Engle, M. J., Goetz, G. S. & Alpers, D. H. Caco-2 cells express a combination of colonocyte and enterocyte phenotypes. *J. Cell. Physiol.* **174**, 362–9 (1998).
221. Behrens, I., Stenberg, P., Artursson, P. & Kissel, T. Transport of lipophilic drug molecules in a new mucus-secreting cell culture model based on HT29-MTX cells. *Pharm. Res.* **18**, 1138–45 (2001).
222. Jochems, P., Garssen, J., van Keulen, A., Masereeuw, R. & Jeurink, P. Evaluating Human Intestinal Cell Lines for Studying Dietary Protein Absorption. *Nutrients* **10**, 322 (2018).
223. Dedhia, P. H., Bertaux-Skeirik, N., Zavros, Y. & Spence, J. R. Organoid Models of Human Gastrointestinal Development and Disease. *Gastroenterology* **150**, 1098–1112

(2016).

224. Clevers, H. Searching for adult stem cells in the intestine. *EMBO Mol. Med.* **1**, 255–259 (2009).

225. Clevers, H. Modeling Development and Disease with Organoids. *Cell* **165**, 1586–1597 (2016).

226. Sato, T. & Clevers, H. Growing Self-Organizing Mini-Guts from a Single Intestinal Stem Cell: Mechanism and Applications. *Science (80-.).* **340**, 1190–1194 (2013).

227. Van De Wetering, M., Francies, H. E. & Garnett, M. J. Prospective Derivation of a Living Organoid Biobank of Colorectal Cancer Patients. (2015) doi:10.1016/j.cell.2015.03.053.

228. Blutt, S. E., Crawford, S. E., Ramani, S., Zou, W. Y. & Estes, M. K. Engineered Human Gastrointestinal Cultures to Study the Microbiome and Infectious Diseases. *Cell. Mol. Gastroenterol. Hepatol.* **5**, 241–251 (2018).

229. Heo, I. *et al.* Modelling Cryptosporidium infection in human small intestinal and lung organoids. *Nat. Microbiol.* **3**, 814–823 (2018).

230. Schweinlin, M. *et al.* Development of an Advanced Primary Human *In Vitro* Model of the Small Intestine. *Tissue Eng. Part C Methods* **22**, 873–883 (2016).

231. Andrée, B., Bär, A., Haverich, A. & Hilfiker, A. Small Intestinal Submucosa Segments as Matrix for Tissue Engineering: Review. *Tissue Eng. Part B Rev.* **19**, 279–291 (2013).

232. Behrens, I. & Kissel, T. Do cell culture conditions influence the carrier-mediated transport of peptides in Caco-2 cell monolayers? *Eur. J. Pharm. Sci.* **19**, 433–442 (2003).

233. Kitano, K. *et al.* Bioengineering of functional human induced pluripotent stem cell-derived intestinal grafts. *Nat. Commun.* **8**, 765 (2017).

234. Nikolaev, M. *et al.* Homeostatic mini-intestines through scaffold-guided organoid morphogenesis. *Nature* **585**, 574–578 (2020).

235. Shah, P. *et al.* A microfluidics-based in vitro model of the gastrointestinal human–microbe interface. *Nat. Commun.* **7**, 11535 (2016).

236. Chen, Y. *et al.* Robust bioengineered 3D functional human intestinal epithelium. *Sci. Rep.* **5**, 13708 (2015).

237. Marzorati, M. *et al.* The HMI™ module: a new tool to study the Host-Microbiota Interaction in the human gastrointestinal tract in vitro. *BMC Microbiol.* **14**, 133 (2014).

238. Zeitouni, N. E., Chotikatum, S., von Köckritz-Blickwede, M. & Naim, H. Y. The

impact of hypoxia on intestinal epithelial cell functions: consequences for invasion by bacterial pathogens. *Mol. Cell. Pediatr.* **3**, 14 (2016).

239. Bertassoni, L. E. *et al.* Hydrogel bioprinted microchannel networks for vascularization of tissue engineering constructs. *Lab Chip* **14**, 2202–11 (2014).
240. Papenfort, K. *et al.* Specific and pleiotropic patterns of mRNA regulation by ArcZ, a conserved, Hfq-aependent small RNA. *Mol. Microbiol.* **74**, 139–158 (2009).
241. Patel, A. P. *et al.* Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science (80-.).* (2014) doi:10.1126/science.1254257.
242. Illumina. *Nextera XT DNA Library Prep Kit Reference Guide For Research Use Only. Not for use in diagnostic procedures*. www.illumina.com/company/legal.html. (2018).
243. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
244. Anders, S., Pyl, P. T. & Huber, W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
245. *Package 'factoextra' Type Package Title Extract and Visualize the Results of Multivariate Data Analyses*. https://github.com/kassambara/factoextra/issues (2017).
246. Milo, R. & Phillips, R. *Cell Biology by the Numbers. Cell Biology by the Numbers* (Garland Science, 2015). doi:10.1201/9780429258770.
247. Smirnov, A. *et al.* Grad-seq guides the discovery of ProQ as a major small RNA-binding protein. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 11591–11596 (2016).
248. Chao, Y. *et al.* In Vivo Cleavage Map Illuminates the Central Role of RNase E in Coding and Non-coding RNA Pathways. *Mol. Cell* **65**, 39–51 (2017).
249. Hör, J., Matera, G., Vogel, J., Gottesman, S. & Storz, G. Trans-Acting Small RNAs and Their Effects on Gene Expression in Escherichia coli and Salmonella enterica. *EcoSal Plus* **9**, (2020).
250. Westermann, A. J. *et al.* Dual RNA-seq unveils noncoding RNA functions in host-pathogen interactions. *Nature* **529**, (2016).
251. Brennecke, P. *et al.* Accounting for technical noise in single-cell RNA-seq experiments. *Nat. Methods* **10**, 1093–1095 (2013).
252. Hollingsworth, M. A. & Swanson, B. J. Mucins in cancer: protection and control of the cell surface. *Nat. Rev. Cancer* **4**, 45–60 (2004).
253. Palmer, A. D. & Slauch, J. M. Mechanisms of Salmonella pathogenesis in animal models. *Hum. Ecol. Risk Assess.* **23**, 1877–1892 (2017).
254. Gewirtz, A. T. *et al.* Salmonella typhimurium induces epithelial IL-8 expression via

Ca(2+)-mediated activation of the NF-kappaB pathway. *J. Clin. Invest.* **105**, 79–92 (2000).

255. Krüger, S., Brandt, E., Klinger, M., Kreft, B. & Kreft, B. Interleukin-8 secretion of cortical tubular epithelial cells is directed to the basolateral environment and is not enhanced by apical exposure to Escherichia coli. *Infect. Immun.* **68**, 328–34 (2000).
256. van der Flier, L. G., Haegebarth, A., Stange, D. E., van de Wetering, M. & Clevers, H. OLFM4 Is a Robust Marker for Stem Cells in Human Intestine and Marks a Subset of Colorectal Cancer Cells. *Gastroenterology* **137**, 15–17 (2009).
257. Ow, M. C., Liu, Q. & Kushner, S. R. Analysis of mRNA decay and rRNA processing in Escherichia coli in the absence of RNase E-based degradosome assembly. *Mol. Microbiol.* **38**, 854–866 (2000).
258. Wangsanuwat, C., Heom, K. A., Liu, E., O'Malley, M. A. & Dey, S. S. Efficient and cost-effective bacterial mRNA sequencing from low input samples through ribosomal RNA depletion. *BMC Genomics* **21**, 717 (2020).
259. Gu, W. *et al.* Depletion of Abundant Sequences by Hybridization (DASH): using Cas9 to remove unwanted high-abundance species in sequencing libraries and molecular counting applications. *Genome Biol.* **17**, 41 (2016).
260. Prezza, G. *et al.* Improved bacterial RNA-seq by Cas9-based depletion of ribosomal RNA reads. *RNA* **26**, 1069–1078 (2020).
261. Lueschow, S. R. & McElroy, S. J. The Paneth Cell: The Curator and Defender of the Immature Small Intestine. *Front. Immunol.* **11**, 587 (2020).
262. Conrad, K. & Stöcker, W. Anti-Intestinal Goblet Cell Antibodies. in *Autoantibodies* 425–431 (Elsevier, 2014). doi:10.1016/B978-0-444-56378-1.00050-2.
263. Mei, X., Gu, M. & Li, M. Plasticity of Paneth cells and their ability to regulate intestinal stem cells. *Stem Cell Res. Ther.* **11**, 349 (2020).
264. Porter, E. M., Bevins, C. L., Ghosh, D. & Ganz, T. The multifaceted Paneth cell. *Cell. Mol. Life Sci.* **59**, 156–70 (2002).
265. van Es, J. H. *et al.* Enteroendocrine and tuft cells support Lgr5 stem cells on Paneth cell depletion. *Proc. Natl. Acad. Sci.* **116**, 26599–26605 (2019).
266. Carulli, A. J. *et al.* Notch receptor regulation of intestinal stem cell homeostasis and crypt regeneration. *Dev. Biol.* **402**, 98–108 (2015).
267. Pellegrinet, L. *et al.* Dll1- and Dll4-Mediated Notch Signaling Are Required for Homeostasis of Intestinal Stem Cells. *Gastroenterology* **140**, 1230-1240.e7 (2011).
268. VanDussen, K. L. *et al.* Notch signaling modulates proliferation and differentiation of

intestinal crypt base columnar stem cells. *Development* **139**, 488–497 (2012).

269. Gjorevski, N. *et al*. Tissue geometry drives deterministic organoid patterning. *Science* **375**, eaaw9021 (2022).
270. LaRock, D. L., Chaudhary, A. & Miller, S. I. Salmonellae interactions with host processes. *Nat. Rev. Microbiol.* **13**, 191–205 (2015).
271. Liu, W. *et al*. Olfactomedin 4 down-regulates innate immunity against Helicobacter pylori infection. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 11056–61 (2010).
272. Song, H. *et al*. Epitranscriptomics and epiproteomics in cancer drug resistance: therapeutic implications. *Signal Transduct. Target. Ther.* **5**, 193 (2020).
273. Erhard, F. *et al*. scSLAM-seq reveals core features of transcription dynamics in single cells. *Nature* **571**, 419–423 (2019).

7. List of publications:

Publications associated the present work and beyond:

The listed publications resulted from studies and collaborations during my PhD:

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

Imdahl, F. & Saliba, A.-E. **Advances and challenges in single-cell RNA-seq of microbial communities**. *Curr. Opin. Microbiol.* **57**, 102–110 (2020).

Reuter C., Fabian Imdahl, Laura Hauf, Ehsan Vafadarnejad, Philipp Fey, Tamara Finger, Heike Walles, Antoine-Emmanuel Saliba, Florian Groeber-Becker, Markus Engstler.
Trypanosoma brucei parasites develop in artificial human skin and persist as skin tissue forms. *bioRxiv* 2021.05.13.443986 (2021) doi:10.1101/2021.05.13.443986. In revision.

Sabrina Dähling, Ana Maria Mansilla, Konrad Knöpper, Anika Grafen, Daniel T. Utzschneider, Milas Ugur, Paul G. Whitney, Annabell Bachem, Panagiota Arampatzi, Fabian Imdahl, Tsuneyasu Kaisho, Dietmar Zehn, Frederick Klauschen, Natalio Garbi, Axel Kallies, Antoine-Emmanuel Saliba, Georg Gasteiger, Sammy Bedoui, Wolfgang Kastenmüller
Type 1 conventional dendritic cells maintain and guide the differentiation of precursors of exhausted T cells in distinct cellular niche.
Immunity, Volume 55, Issue 4, 2022, 656-670

Däullary T. & Imdahl F. *et al*. **Humen intetsinal tissue model reveals dynamic and herterogeneity of *Salmonella* Typhimurium infection**. In preparation (2022).

Previous publications:

The listed studies have been published prior to my PhD work and are not related:

Zhang, Y., Lee, C.-W., Wehner, N., Imdahl, F. Svetlana, V. Weiste, C. Dröge-Laser, W., Deeken R. Regulation of Oncogene Expression in T-DNA-Transformed Host Plant Cells. *PLoS Pathog*. **11**, (2015).

8. Contributions by others:

The work presented here was conducted under the supervision of Dr. Antoine-Emmanuel Saliba at Helmholtz-Institute for RNA-based infection research in Würzburg, Germany. Several experiments were conducted by colleagues and collaborators I want to name here:

- Bioinformatic analysis of single-bacteria RNA sequencing was established and performed by Ehsan Vafadarnejad (HIRI, Würzburg).

- Human intestinal tissue model generation was developed, and ready models provided by Thomas Däullary (Chair of tissue engineering and regenerative medicine; Würzburg, Germany)

- Bioinformatic analysis of hITM before and after infection was conducted by Ehsan Vafadarnejad and Oliver Dietrich (HIRI, Würzburg).

- All imaging techniques and procedures were carried out and conducted by Thomas Däullary (Chair of tissue engineering and regenerative medicine; Würzburg, Germany)

- HCR-Fish; Method establishment, probe design and application were performed by Tobias Krammer (HIRI, Würzburg).

10. Affidavit / Eidesstattliche Erklärung:

I hereby confirm that my thesis entitled "*Development of novel experimental approaches to decipher host-pathogen interactions at the single-cell level*" is the result of my own work. I did not receive any help or support from commercial consultants. All sources and materials applied are listed and specified in the thesis.
Furthermore, I confirm that this thesis has not yet been submitted as part of another examination process neither in identical nor similar form.

Hiermit erkläre ich an Eides statt, die Dissertation „*Development of novel experimental approaches to decipher host-pathogen interactions at the single-cell level*" eigenständig d.h. insbesondere selbständig und ohne die Hilfe eines kommerziellen Promotionsberaters angefertigt und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet zu haben.
Ich erkläre außerdem, dass die Dissertation weder in gleicher noch in ähnlicher Form bereits in einem anderen Prüfungsverfahren vorgelegen hat.

Würzburg, 06.05.2022 _____

Fabian Imdahl