Review

Daniel Ryan, Gianluca Prezsa and Alexander J. Westermann*

An RNA-centric view on gut Bacteroidetes

<https://doi.org/10.1515/hsz-2020-0230>

Received June 26, 2020; accepted August 21, 2020; published online September 24, 2020

Abstract: Bacteria employ noncoding RNAs to maintain cellular physiology, adapt global gene expression to fluctuating environments, sense nutrients, coordinate their interaction with companion microbes and host cells, and protect themselves against bacteriophages. While bacterial RNA research has made fundamental contributions to biomedicine and biotechnology, the bulk of our knowledge of RNA biology stems from the study of a handful of aerobic model species. In comparison, RNA research is lagging in many medically relevant obligate anaerobic species, in particular the numerous commensal bacteria comprising our gut microbiota. This review presents a guide to RNA-based regulatory mechanisms in the phylum Bacteroidetes, focusing on the most abundant bacterial genus in the human gut, *Bacteroides* spp. This includes recent case reports on riboswitches, an mRNA leader, *cis*- and *trans*-encoded small RNAs (sRNAs) in *Bacteroides* spp., and a survey of CRISPR-Cas systems across Bacteroidetes. Recent work from our laboratory now suggests the existence of hundreds of noncoding RNA candidates in *Bacteroides thetaiotaomicron*, the emerging model organism for functional microbiota research. Based on these collective observations, we predict mechanistic and functional commonalities and differences between *Bacteroides* sRNAs and those of other model bacteria, and outline open questions and tools needed to boost Bacteroidetes RNA research.

Keywords: *Bacteroides*; CRISPR-Cas; GibS; microbiota; noncoding RNA; small RNA.

Intestinal Bacteroidetes thrive in a dynamic microenvironment

The human gut is broadly subdivided into the small and large intestine. Compared with the small intestine, the large bowel presents a relatively mild environment to colonizing microorganisms due in part to the relatively higher pH levels (pH 5.5–7), lower oxygen tension, and a reduced immunogenic milieu that favors balance over clearance. However, the nutritional environment, which is dominated by complex polysaccharides that cannot be readily absorbed in the small intestine, imposes strong metabolic pressure on colon-resident bacteria. Nevertheless, thanks to the evolution of sophisticated enzymatic repertoires to catabolize these carbon sources and metabolic co-dependencies, the large intestine represents one of the densest microbial ecosystems in nature. Of the 10^{11} – 10^{12} bacteria per gram of fecal content that constitute the colon microbiota (Knight and Girling 2003), obligate anaerobic species comprise the largest fraction with two dominant phyla, the Gram-positive Firmicutes and Gram-negative Bacteroidetes (Huttenhower et al. 2012).

The Bacteroidetes are non-motile, non-spore forming, rod-shaped bacteria. Although Bacteroidetes species also occur outside the gut, here we focus on intestinal members. Among them, *Bacteroides* spp. constitute the most abundant bacterial genus in the human gut, where they contribute to the release of energy from dietary fiber and represent a major source of short-chain fatty acids. Biogeographically speaking, *Bacteroides* are enriched in the lumen and outer mucus layer along the colon (Donaldson et al. 2016) (Figure 1). Some mucin-degrading species, such as *Bacteroides fragilis*, can also colonize colonic crypts where they modulate the host immune system (Lee et al. 2013; Round et al. 2011). *Bacteroides* spp. appear in neonates at about four to six days after birth with relative abundance depending on the mode of delivery, diet, and gestational age and stably persist in the gut for a lifetime. Previous studies linked *Bacteroides* abundance in the human intestine with a lower risk of developing obesity (Ley

***Corresponding author: Alexander J. Westermann**, Helmholtz Institute for RNA-based Infection Research (HIRI), Helmholtz Centre for Infection Research (HZI), Josef-Schneider-Str. 2/D15, D-97080, Würzburg, Germany; and Institute of Molecular Infection Biology (IMIB), University of Würzburg, Josef-Schneider-Str. 2/D15, D-97080, Würzburg, Germany, E-mail: alexander.westermann@uni-wuerzburg.de. <https://orcid.org/0000-0003-3236-0169>

Daniel Ryan and Gianluca Prezsa, Helmholtz Institute for RNA-based Infection Research (HIRI), Helmholtz Centre for Infection Research (HZI), Josef-Schneider-Str. 2/D15, D-97080, Würzburg, Germany, E-mail: daniel.ryan@helmholtz-hiri.de (D. Ryan), E-mail: gianluca.prezsa@helmholtz-hiri.de (G. Prezsa). <https://orcid.org/0000-0003-4261-2702> (D. Ryan). <https://orcid.org/0000-0002-0032-4369> (G. Prezsa)

et al. 2006) or colorectal cancer (Lee et al. 2018), but also with inflammatory disorders (Bloom et al. 2011) – yet often cause and consequence are difficult to disentangle. The commensals *Bacteroides thetaiotaomicron* and *B. fragilis* – the latter also being an opportunistic pathogen outside the gut (Goldstein, 1996) – are gaining increasing attention as model organisms for functional microbiota research. This is due to their prevalence, impact on host physiology and metabolism, and the relative ease with which they can be cultured and genetically manipulated (Bacic and Smith 2008). Their study has already revealed molecular processes that form the basis for successful colonization of the large intestine.

For instance, to stably colonize the colon, *Bacteroides* spp. evolved multiple arrays of paralogous gene clusters known as polysaccharide utilization loci (PULs) that allow them to feed on complex diet- and host-derived polysaccharides. Generally, PULs encode cell envelope-spanning complexes consisting of glycolytic enzymes and outer membrane proteins such as SusCD homologs (for starch utilization system; historically the first described PUL system (Reeves et al. 1997)) that are required for glycan binding (SusD) and import (SusC). The regulation of PULs occurs at the transcriptional level by several mechanisms. SusR-like regulators (D’Elia and Salysers 1996) and hybrid two-component systems (HTCSs) (Sonnenburg et al., 2006, 2010) both combine sugar-sensing and gene regulatory

functions into a single polypeptide and activate the transcription of specific PUL operons. In parallel, PUL transcription is controlled by extracytoplasmic function sigma/anti-sigma factor pairs whose dissociation is induced in the presence of appropriate glycans, releasing the cognate sigma factor to activate transcription of its target operon (Martens et al. 2008).

Bacteroides genomes further contain multiple polysaccharide biosynthesis loci for capsule formation. Capsular polysaccharides (CPSs) are essential for gut colonization as they contribute to cross-talk with the host epithelium and determine bacterial susceptibility to bacteriophage attack (Liu et al. 2008). CPS expression is regulated by invertible promoters and mediated by site-specific recombinases (Coyne et al. 2003). Additionally, CPS expression is regulated co-transcriptionally by specialized NusG-like proteins of the UpxY family (where “x” designates the cognate CPS locus). Upon binding specifically to sequences within the nascent 5’ UTR of CPS operons, UpxY interacts with the transcribing RNA polymerase to prevent premature termination, thus allowing complete transcription of these 11–23 kb loci (Chatzidaki-Livanis et al. 2009). In turn, UpxZ (produced from a gene downstream of *upxY* in a CPS operon) inhibits the anti-termination activity of UpxY proteins of different CPSs, thereby giving rise to a regulatory hierarchy independent of DNA inversions (Chatzidaki-Livanis et al. 2010).

Figure 1: Dynamics associated with the intestinal niches occupied by *Bacteroidetes*. The large intestine is associated with fluctuations in nutrient availability (1) and oxygen concentration. Colon-colonizing bacteria need to adapt to these changes in order to efficiently compete with neighboring microbes. Additionally, they defend themselves against antimicrobial compounds released by co-colonizing bacteria (2) and phage attacks (3), and cross-talk with the immune system of the host (4). Gut-associated *Bacteroidetes* occur in the lumen of the large bowel, but some species may also attach to mucosal surfaces at the host epithelium. As glycan generalists, *Bacteroides* spp. can feed on both, diet- or host-derived polysaccharides. They are bile-resistant and survive transient increases in oxygen levels. Individual entities are not drawn to scale.

RNA landscape of the Bacteroidetes

Complementing protein-mediated regulation of transcription, RNA-mediated gene expression control is widespread in bacteria. Over the past two decades, an astonishing versatility in RNA-centric mechanisms has been uncovered, particularly in model gastrointestinal pathogens. The regulatory RNAs that bring about these control mechanisms are grouped into different classes based on their genomic organization: regulatory elements within 5' untranslated regions (UTRs) of mRNAs, including riboswitches (Breaker 2012) and RNA thermometers (Kortmann and Narberhaus 2012), *cis*-encoded antisense RNAs (Wagner et al. 2002), and *trans*-encoded small RNAs (sRNAs) (Storz et al. 2011; Wagner and Romby 2015). While some riboregulators function autonomously, the activity of others, e.g., that of many sRNAs, depends on assisting proteins (Holmqvist and Vogel 2018).

As inferred from the roles regulatory RNAs play in Proteobacteria, where they often function in the context of stress adaptation and metabolism (Bobrovskyy and Vanderpool 2013; Holmqvist and Wagner 2017), one would expect these molecules to also help commensal gut bacteria to rapidly adapt to their ever-changing microenvironment. Additionally, given the contribution of regulatory RNAs in bacterial pathogens to their interaction with host cells (Svensson and Sharma 2016; Westermann 2018), we postulate riboregulation may also underlie the cross-talk of anaerobic gut commensals with their host and – potentially – companion microbes. However, compared with our firm knowledge of regulatory RNAs in enteric pathogens, little is known about RNA biology in the beneficial bacteria colonizing our gastrointestinal tract and, in particular, in obligate anaerobic Bacteroidetes species. Recent findings from global transcriptomics (Cao et al. 2016; Ryan et al. 2020) now predict hundreds of noncoding RNAs and support the idea of a rich RNA world in *Bacteroides* spp., as will be reviewed in the following sections (summarized in Table 1).

“Housekeeping” RNAs

Amongst the most conserved bacterial noncoding transcripts are certain specialized housekeeping RNAs with functions in the maintenance of key cellular processes. Recent work from our laboratory verified the existence of such transcripts in *B. thetaiotaomicron* (Ryan et al. 2020). The transfer-messenger RNA (tmRNA, a.k.a. SsrA), for example, rescues stalled ribosomes (Withey and Friedman 2003). The two domains of the tmRNA – the tRNA-like

domain and the downstream short open reading frame – can be part of the same molecule or cleaved into two base-pairing parts (Mao et al. 2009). Across the Bacteroidetes, tmRNAs show high sequence conservation and in *B. thetaiotaomicron* RNA-seq and Northern blot data revealed this RNA to be transcribed as a 507 nt precursor that is processed into the ~400 nt mature form containing both domains (Ryan et al. 2020).

The signal recognition particle (SRP) directs target proteins to the membrane. In bacteria, the RNA component of this complex – the 4.5S RNA (~110 nt long; encoded by the *ffs* gene) – fulfills a scaffold function by forming a platform for the association of the proteinaceous SRP constituents (Peluso et al. 2000). Initially identified as candidate sRNA BTnc259 (Ryan et al. 2020), sequence comparison with known 4.5S homologs from other bacteria using Rfam (Kalvari et al. 2018) and BLAST suggested this transcript as the putative 4.5S RNA of *B. thetaiotaomicron* (Prezza and Ryan et al. in prep).

The M1 RNA (encoded by *mpB*) is the catalytic RNA component of the ribonuclease (RNase) P holoenzyme, which is involved in the processing of tRNA, 4.5S RNA, and tmRNA precursor molecules (Altman 2011). While active even by itself under optimized *in vitro* conditions, M1 RNA requires the accessory RnpA protein (BT_3227 in *B. thetaiotaomicron*) for efficient functioning in the bacterial cell. Based on their secondary structures, bacterial M1 RNAs are grouped into two classes – type A and B – with Bacteroidetes M1 RNAs falling into the type A category. We recently validated the M1 RNA in *B. thetaiotaomicron* as a primary transcript of ~390 nt with evidence of processing at the 3' end, resulting in a mature transcript of ~360 nt (Ryan et al. 2020).

Regulatory elements within 5'UTRs of mRNAs

The 5' UTR of bacterial mRNAs may contain *cis*-regulatory features that control the expression of the downstream coding sequence (CDS). Riboswitches, for example, are *cis*-regulatory elements in front of mRNAs for metabolic enzymes and transporters (McCown et al. 2017). Riboswitches consist of two domains – the aptamer region binds with high specificity to a metabolite and this ligand binding, in turn, induces a conformational change in the so-called expression platform, thereby influencing mRNA expression at the level of transcription elongation or translation initiation (Lotz and Suess 2018). RNA thermometers, too, are regulatory elements within 5' UTRs of mRNAs, but in this case a conformational change is triggered by

Table 1: Partially characterized noncoding RNA elements in *Bacteroidetes*.

RNA class	Name(s)	Description	Length	Species ^a	Reference(s)
Housekeeping RNA	tmRNA (SsrA)	tRNA- and mRNA-like properties; rescues stalled ribosomes	~400 nt	<u><i>Bacteroides thetaiotaomicron</i></u> , widely conserved across <i>Bacteroidetes</i>	Ryan et al. (2020)
Housekeeping RNA	4.5S RNA (Ffs; BTnc259)	Structural component of the SRP complex that directs proteins to the membrane	~110 nt	<u><i>B. thetaiotaomicron</i></u> , widely conserved across <i>Bacteroidetes</i>	Prezza and Ryan et al. (in prep); Ryan et al. (2020)
Housekeeping RNA	M1 RNA (RnpB; BTnc059)	Ribozyme; together with RnpA protein, M1 forms RNase P which is e.g., involved in the maturation of tRNAs	~360 nt	<u><i>B. thetaiotaomicron</i></u> , widely conserved across <i>Bacteroidetes</i>	Ryan et al. (2020)
Riboswitch	TPP riboswitch	Senses the active form of vitamin B ₁ ; acts at the level of transcription elongation or translation initiation	~100 nt	<u><i>B. thetaiotaomicron</i></u> , <u><i>Bacteroides uniformis</i></u> , <u><i>Bacteroides vulgatus</i></u> ; widely conserved across <i>Bacteroidetes</i>	Costliow et al. (2019), Costliow and Degnan (2017), Rodionov et al. (2002)
Riboswitch	AdoCbl riboswitch	Senses the active form of vitamin B ₁₂ ; acts at the level of translation initiation	~200 nt	<u><i>B. thetaiotaomicron</i></u> , <u><i>Porphyromonas gingivalis</i></u> ; widely conserved across <i>Bacteroidetes</i>	Degnan et al. (2014); Hirano et al. (2012); Vitreschak et al. (2003)
mRNA leader	<i>roc</i> leader	Repression of <i>Roc</i> colonization factor in the presence of glucose or fructose via an unknown mechanism	54 nt	<u><i>B. thetaiotaomicron</i></u> , conserved across <i>Bacteroides</i> spp.	Townsend et al. (2019)
<i>Cis</i> -antisense RNA	DonS	Repression of <i>donC</i> (<i>susC</i> homolog of the Don PUL) via transcriptional interference and/or RNA–RNA interaction	128 nt	<u><i>Bacteroides fragilis</i></u> ; conserved across <i>Bacteroides</i> spp.	Cao et al. (2016)
<i>Trans</i> -encoded sRNA (protein antagonist)	6S RNA (SsrS; BTnc017)	Sponges RNA polymerase by molecular mimicry to globally shut down transcription	~190 nt	<u><i>B. thetaiotaomicron</i></u> ; widely conserved across <i>Bacteroidetes</i>	Prezza and Ryan et al. (in prep); Ryan et al. (2020)
<i>Trans</i> -encoded sRNA (base-pairing)	RteR	Inhibition of conjugative transfer of CTnDOT; mechanism likely involves base-pairing to the nascent target transcript, leading to premature transcription termination and dis-coordinate expression of the <i>tra</i> operon	90 nt	<u><i>B. thetaiotaomicron</i></u> ; partially conserved across <i>Bacteroidetes</i>	Jeters et al. (2009); Waters and Salyers (2012)
<i>Trans</i> -encoded sRNA (base-pairing)	GibS (BTnc035)	Intergenic sRNA induced in stationary phase and in the presence of <i>N</i> -acetylglucosamine as the sole carbon source; represses <i>BT_0771</i> and <i>BT_3893</i> by direct binding to their translation initiation regions; activates <i>BT_1871</i> (directly or indirectly)	145 nt	<u><i>B. thetaiotaomicron</i></u> ; partially conserved across <i>Bacteroides</i> spp.	Ryan et al. (2020)

^aUnderlined species are the ones in which the respective RNA element was validated/characterized.

temperature alterations, affecting the access of ribosomes to the downstream CDS (Kortmann and Narberhaus 2012). The length of an mRNA leader can thus hint at the presence of a *cis*-regulatory element. We recently determined the median and average lengths of 5' UTRs in *B. thetaiotaomicron* (32 nt; 52 nt) with a fraction (13.5%) of mRNAs having unusually long (>100 nt) leader sequences (Ryan et al. 2020), arguing that *cis*-regulatory elements may be prevalent in *Bacteroides*.

Indeed, several riboswitches have been inferred from homology searches and characterized in *Bacteroidetes*. The thiamine pyrophosphate (TPP)-sensing riboswitch is the only known class to occur in all three domains of life and is involved in the regulation of metabolism and transport of TPP, the active form of vitamin B₁, that is required for a range of cellular processes (Sudarsan et al. 2003). Comparative genomics predicted TPP riboswitches across the *Bacteroidetes*, typically in front of operons of

genes involved in the biosynthesis or import of thiamine (Costliow and Degnan 2017; Rodionov et al. 2002). TPP riboswitches illustrate how riboswitch aptamers and expression platforms can be mixed and matched, leading to different regulatory consequences. That is, the TPP riboswitches of *Bacteroides vulgatus* and *Bacteroides uniformis*, and a thiamine biosynthetic TPP riboswitch in *B. thetaiotaomicron* locate >50 nt upstream of the CDS and work at the level of transcription (Costliow et al. 2019). In contrast, the two TPP riboswitches governing thiamine import operons in *B. thetaiotaomicron* locate immediately upstream of their cognate start codon and function at the level of translation initiation, which led to the hypothesis that the distance between a riboswitch and the downstream start codon hints at its mode-of-action (Costliow et al. 2019) (Figure 2a). While transcriptional control is considered tighter, translational riboswitches enable faster and reversible responses to metabolite sensing. However, not only the mode-of-action, but also the ligand threshold concentrations differ between *B. thetaiotaomicron* biosynthetic (~10 nM) and transport riboswitches (~100 nM), suggesting hierarchical control of thiamine synthesis and import.

Vitamin B₁₂ is an essential cofactor for several enzymes. The majority of gut-associated bacteria contain importers for vitamin B₁₂ (Degnan et al. 2014) and homologs of the TonB-dependent transporter, BtuB, are widely distributed across the Bacteroidetes. *B. thetaiotaomicron* contains three *btuB* gene copies (*BT_1489*, *BT_1953*, *BT_2094*), and all of them are associated with a riboswitch (Degnan et al. 2014; Vitreschak et al. 2003). The predicted mechanism of regulation involves binding of adenosylcobalamin (AdoCbl) – the biologically active form of vitamin B₁₂ – to the aptamer that in turn refolds the region around the ribosome-binding site (RBS) to block translation initiation (Vitreschak et al. 2003) (Figure 2b). Reminiscent of TPP riboswitches, AdoCbl riboswitches of individual transporter mRNAs respond at different ligand concentrations, again suggesting hierarchical control (Degnan et al. 2014). In our recent transcriptome study, we identified another putative AdoCbl riboswitch in the 5' UTR of *BT_1915*, which encodes a pyruvate carboxylase subunit A protein (Ryan et al. 2020).

Thermo-sensing RNAs have been described mostly in bacteria that experience temperature changes throughout their lifecycles, such as pathogens that face a temperature increase when entering their mammalian host. From this point of view, it was somewhat surprising that – based on sequence similarity to known RNA thermometers (namely ROSE_3 [RF02523] and PrfA thermoregulatory UTR [RF00038]) – thermo-sensing RNA candidates were

predicted in *B. thetaiotaomicron* (Ryan et al. 2020) that, despite many other variable parameters (Figure 1), populates a niche with a fairly constant temperature. The validity of these predictions and their biological relevance, however, need further investigation.

The Roc protein (for regulator of colonization; BT_3172) of *B. thetaiotaomicron* is a HTCS regulator that activates transcription of a defined PUL (BT_3173-3174) which metabolizes host glycans (Townsend et al. 2013). The 5' UTR of *roc* mRNA is relatively long (54 nt), pointing to a *cis*-regulatory RNA element, albeit too short to harbor a riboswitch. The *roc* leader sequence mediates repression of the associated CDS and downstream genes in that operon in the presence of glucose and fructose (Townsend et al. 2019). While the mechanism of Roc repression is elusive, judged from several highly conserved residues within the leader sequence, it was hypothesized that regulation could occur through the binding of a *trans*-acting sRNA or a regulatory protein (Townsend et al. 2019) (Figure 2c). From an evolutionary standpoint, this regulation may guarantee that *B. thetaiotaomicron* finds its proper niche in the host gut: colonization factors shall be expressed only in a microenvironment where this glycan generalist has a selective advantage over competing microbes that catabolize simple sugars, but fail to process complex polysaccharides.

Cis-encoded antisense RNAs

Bacterial RNAs partially overlapping with genes on the opposite strand were first identified on accessory genetic elements, where they fulfill specialized functions including the control of plasmid replication and conjugation or lysis/lysogeny decisions in phages (Wagner et al. 2002). The advent of RNA-seq in bacterial transcriptomics led to the unexpected finding that antisense RNAs are widespread also in the core genome (Georg and Hess, 2018). In *B. thetaiotaomicron*, for example, we detected ~1100 antisense transcription start sites (TSSs), comprising one fourth of all identified initiation sites (Ryan et al. 2020). While absolute numbers should be interpreted with caution – as they may be heavily influenced by technical variation between RNA-seq protocols and analysis pipelines – this value is of a similar magnitude to those determined for proteobacterial model organisms such as *Salmonella enterica* (13%; Kröger et al. 2012), *Escherichia coli* (37%; Thomason et al. 2015), and *Helicobacter pylori* (41%; Sharma et al. 2010). It is currently debated to what extent this plethora of antisense RNAs is functional or merely the result of spurious transcription (Lloréns-Rico et al. 2016). For a handful of antisense candidates in

Figure 2: Working principles of characterized *Bacteroides* regulatory RNA elements.

Proposed mechanisms of the thiamine pyrophosphate (TPP) (a) and adenosylcobalamin (AdoCbl) (b) riboswitches, of the *roc* mRNA leader (c), of the DonS antisense RNA (d), and of the RteR (e) and GibS (f) sRNAs. Coding genes are indicated as grey arrows, mRNAs are in dark blue and regulatory RNA genes/elements in red. See main text for details.

Proteobacteria and Firmicutes, however, functionality has been demonstrated (Wade and Grainger 2014). These RNAs affect expression of their cognate sense-overlapping target by a variety of mechanisms, including transcriptional interference and – post-transcriptionally – base-pairing and masking of the target mRNA's RBS or shielding or generating an RNase cleavage site within the CDS.

An exploratory RNA-seq study recently discovered a specialized class of *cis*-antisense transcripts in *B. fragilis* – the PUL-overlapping antisense RNAs (Cao et al. 2016). These 78–128 nt-long RNAs are enriched within PULs for host-derived glycan processing and divergently encoded to the respective *susC* homolog. Functional characterization of the Don (*degradation of N-glycans*) PUL system revealed the cognate antisense RNA – termed DonS – to repress its PUL. The mode-of-action employed by DonS (and other PUL-associated antisense RNAs) is elusive, but either a transcriptional interference mechanism or – since they overlap the translation initiation region of the cognate *susC* homologs – translational inhibition appear plausible (Figure 2d). In case of the latter, post-transcriptional PUL repression by the antisense RNAs would complement transcriptional control through the corresponding anti-sigma factor, together ensuring tight PUL repression in the presence of a prioritized carbon source. Indeed, a $\Delta donS$ *B. fragilis* mutant was unable to effectively shut down Don when glucose was added to the medium (Cao et al. 2016).

PUL-associated antisense RNAs were also predicted for related species, including *B. thetaiotaomicron*, *B. vulgatus*, and *Bacteroides ovatus* (Cao et al. 2016). Our laboratory further expanded on this class of PUL-overlapping antisense RNAs with the identification of five additional candidates in *B. thetaiotaomicron*, three of which (BTnc055, BTnc136, BTnc252) are antisense to the translation initiation region, while the remaining two (BTnc011, BTnc111) are antisense to the CDS of the cognate *susC* homolog. (Ryan et al. 2020)

Trans-encoded small RNAs

sRNAs are short RNA molecules – typically between 50 and 300 nt in length – that are encoded by independent genes or arise from the UTRs of mRNAs. They can regulate target gene expression via two distinct, yet not mutually exclusive mechanisms: directly, by mediating imperfect base-pair interactions with specific *trans*-encoded target mRNAs, or indirectly, by titrating regulatory RNA-binding proteins (RBPs). Amongst the latter, are the sRNA antagonists of the translational regulator CsrA/RsmA (Romeo and Babitzke

2018), the Ro60-interacting Y RNAs with versatile cellular functions (Sim and Wolin 2018), and the 6S RNA (encoded by the *ssrS* gene), which sponges the RNA polymerase holoenzyme to globally tune transcriptional activity (Wassarman 2018). In Bacteroidetes, no CsrA system is known and Y RNAs have not been identified either; however, a 6S RNA homolog was predicted in this phylum (Wehner et al. 2014) and recently validated in *B. thetaiotaomicron* as a ~190 nt-long transcript (Ryan et al. 2020). Generally, 6S RNA adopts the characteristic structure of a long hairpin with a central asymmetric bulge and sequesters RNA polymerase by molecular mimicry of the transcription bubble in genomic DNA. To reactivate transcription, RNA polymerase uses 6S RNA as a template, resulting in the synthesis of 14–20 nt-long product RNAs (Wassarman and Saecker, 2006) that, too, were detected in *B. thetaiotaomicron* (Ryan et al. 2020). This argues that this type of transcriptional control is an ultra-conserved mechanism across the bacterial kingdom (Barrick et al. 2005).

In contrast to protein antagonists, base-pairing sRNAs anneal through short “seed” sequences with partially complementary target sites in mRNAs and repress or activate their targets through a variety of mechanisms (Wagner and Romby 2015). In Gram-negative species, sRNAs often depend on assisting RNA chaperones such as the Sm-like Hfq or the FinO domain-containing proteins (Holmqvist and Vogel 2018). In the best-understood scenario of sRNA-mediated target control from γ -Proteobacteria, Hfq binding stabilizes the sRNA in the cytosol and facilitates its annealing to the target sequence – classically within the 5' region of an mRNA, overlapping with the Shine-Dalgarno sequence and/or start codon to block translation initiation (Hör et al. 2020). With respect to the Bacteroidetes, this raises several interesting questions: are *trans*-acting sRNAs prevalent in this phylum? If so, and given that obvious homologs of Hfq and FinO are missing, do Bacteroidetes sRNAs work in a protein-independent manner or are there other global RBPs that chaperone the sRNAs? Since Bacteroidetes further lack the classical Shine-Dalgarno sequence of Proteobacteria (Nakagawa et al. 2010), would sRNAs still preferentially bind to the 5' region of mRNAs and inhibit translation, or are other modes of target control more common in Bacteroidetes?

In our recent work, we addressed some of these questions: application of differential RNA-seq (Sharma and Vogel 2014) to refine the transcriptome annotation of *B. thetaiotaomicron* type strain VPI-5482, led to the discovery of 269 noncoding RNA elements, including 151 putative sRNAs (Ryan et al. 2020). This number is comparable with the sRNA complement in proteobacterial species (Dugar et al. 2013; Kröger et al. 2013, 2018; Sharma et al. 2010; Vogel et al. 2003). Of 14 selected sRNA candidates, nine were validated by

Northern blot, including a 3' UTR-derived sRNA. 3'-derived sRNAs had previously only been described in Proteobacteria (Miyakoshi et al. 2015); this finding from *Bacteroides* now implies that the 3' end of mRNAs may constitute a reservoir for regulatory RNAs throughout the bacterial phylogenetic tree. We further identified rarer cases of 5' UTR-derived and intra-operonic sRNA candidates, suggesting an expanded sequence space for the origin of sRNAs (Adams and Storz 2020; Jose et al. 2019). Sequence conservation often is a predictor of functional importance. We therefore selected a core set of 49 intergenic sRNAs – harboring both, a cognate TSS and a predicted Rho-independent transcription terminator – for conservation analyses. Sequence alignment across the Bacteroidetes phylum revealed 22 of them to be conserved in two or more species, while the remaining 27 were *B. thetaiotaomicron*-specific (Prezza and Ryan et al. in prep; Ryan et al. 2020).

Since sRNAs are prevalent in the Bacteroidetes, how do they function in the absence of homologs of known global RBPs and classical Shine-Dalgarno sequences? Up to now, two *trans*-encoded *B. thetaiotaomicron* sRNAs have been functionally characterized. RteR (regulation of tetracycline resistance elements RNA) is a sRNA of 90 nt, encoded downstream of the *exc* gene (the *rteR* promoter overlaps the *exc* stop codon) within the excision region of the integrative conjugative transposon CTnDOT, and is widely conserved within *Bacteroides* spp. (Jeters et al. 2009; Waters and Salysers 2012). In *B. thetaiotaomicron*, RteR promotes dis-coordinate expression of the *tra* operon, whose products are required to assemble the mating apparatus for the transfer of CTnDOT. That is, while the mRNA levels of *traA* – the first gene in the operon, encoding a conjugative transposon protein – are not affected by the sRNA, the downstream *traB–Q* genes are repressed (Waters and Salysers 2012). However, RteR shows no obvious effect on the half-life of *tra* mRNA, speaking against a post-transcriptional effect. Rather, a co-transcriptional mechanism was proposed (Figure 2e): since a sequence resembling an intrinsic transcription terminator as well as stretches with partial sequence complementarity to RteR were identified within the *traB* CDS, RteR may interact with the nascent *tra* transcript, inducing a conformational change that results in premature termination of *tra* transcription. Experimental evidence for this mechanism is still required and it is currently also unknown whether the prematurely terminated transcript is still a substrate for TraA synthesis. Irrespectively, however, RteR inhibits conjugative transfer of CTnDOT and thereby influences the spread of antibiotic resistance genes (Waters and Salysers 2012).

That *Bacteroides* sRNAs can also mediate post-transcriptional control of gene expression was recently

supported by our findings on GibS (GlcNAc-induced *Bacteroides* sRNA) (Ryan et al. 2020). In *B. thetaiotaomicron*, GibS is transcribed from an intergenic region in between a putative para-aminobenzoate synthase cluster (*BT_0763–68*) and a glycogen biosynthesis operon (*BT_0769–71*). GibS is associated with a *Bacteroides* promoter motif that is enriched in front of stationary phase-induced genes and, accordingly, GibS levels increase over growth in rich medium. GibS steady-state levels, however, increase even further when the bacteria grow in minimal medium with N-acetyl-D-glucosamine (GlcNAc) – a monosaccharide constituent of host-derived glycosaminoglycans – as the sole carbon source. Structural analysis by chemical and enzymatic probing revealed the conformation of this 145 nt-long sRNA, composed of a single-stranded 5' region (~40 nt), followed by two meta-stable hairpins and a Rho-independent terminator. Genome-wide differential expression analysis guided the identification of GibS target operons, which are related to metabolic processes. In particular, GibS activates an operon comprising a galactosidase and a periplasmic glucosidase gene (*BT_1871–BT_1872*) and represses the *BT_0769–BT_0771* operon harboring genes for glucan-branching enzymes, as well as *BT_3893*, which codes for a hypothetical protein. *In silico* prediction and *in vitro* validation experiments revealed the unstructured 5' region of GibS, and two distinct seed regions therein, to be at the heart of this regulation. That is, GibS employs one or both of its seed regions, respectively, to anneal with sequence stretches spanning the start codons of *BT_0771* or *BT_3893* (Figure 2f). The physiological role of GibS needs further investigation; a Δ *gibS* *B. thetaiotaomicron* mutant shows no strong growth phenotype in rich medium, but grows slightly faster than an isogenic wild-type when feeding on GlcNAc (Ryan et al. 2020).

RBP candidates and ribonucleases in Bacteroidetes

Global RNA binding proteins such as Hfq and FinO-like proteins are mediators of sRNA-target interactions in many Gram-negative bacteria (Holmqvist and Vogel 2018; Woodson et al. 2018). The role of Hfq has been extensively studied in Proteobacteria, particularly in *E. coli* and *S. enterica*, where it forms a homohexameric ring with three RNA-interacting surfaces to stabilize the bound sRNA and facilitate its annealing to target mRNAs (Kavita et al. 2018; Santiago-Frangos and Woodson 2018; Vogel and Luisi 2011). FinO-like proteins have only recently emerged as global RNA binders in α -, β -, and γ -Proteobacteria, but

Table 2: Ribonucleases predicted for *Bacteroidetes*.

Ribonuclease	Identifier in <i>B. thetaiotaomicron</i>	Ribonuclease activity	Substrate class(es)
RNase E/G family	BT_1500	Endonuclease	tRNA, rRNA, mRNA, sRNA
RNase III	BT_3357	Endonuclease	rRNA, mRNA, sRNA
RNase P	BT_3227, BTnc059 (M1 RNA)	Endonuclease	tRNA, mRNA, rRNA
RNase BN	BT_1315	3' → 5' exonuclease, endonuclease	mRNA, tRNA
RNase H-like	BT_3391	Endonuclease	RNA-DNA hybrids
RNase HII	BT_3412	Endonuclease	RNA-DNA hybrids
RNase R	BT_3077	3' → 5' exonuclease	rRNA, sRNA
RNase Z	BT_4346	3' → 5' exonuclease, endonuclease	mRNA, tRNA
RNase YbeY	BT_3283	Endonuclease	rRNA

Genes in *B. thetaiotaomicron* that are annotated as putative ribonucleases and highly conserved across the *Bacteroidetes* phylum. The list is derived from <https://biocyc.org/>. The substrate classes were inferred from Georg and Hess (2018) and Mohanty and Kushner (2018).

resolved mechanisms of sRNAs associated with these proteins remain sparse (Olejniczak and Storz 2017). *Bacteroidetes* species lack homologs of both, Hfq and FinO. In contrast, proteins containing cold-shock or K homology domains – which also have the ability to bind RNA – are prevalent in the *Bacteroidetes* (Prezza and Ryan et al. in prep) and it remains to be investigated whether any of them functionally substitutes for Hfq or FinO-like proteins in this phylum.

Central to sRNA regulatory pathways in Gram-negative bacteria is the activity of RNases (Mohanty and Kushner 2018). In *Proteobacteria*, RNase E is a central player in sRNA regulatory processes: it is involved in the processing of several RNA species, including 3'-derived sRNAs (Chao et al. 2017), and can be recruited by sRNA-Hfq complexes to induce endonucleolytic cleavages within target mRNAs, initiating their rapid decay (Bandyra and Luisi 2018). In the *Bacteroidetes* phylum, several putative RNases (Table 2) have been annotated through automated homology searches (<https://biocyc.org/>) (Karp et al. 2019). This includes a member of the family of RNase E/G-like endonucleases (BT_1500 in *B. thetaiotaomicron*);

whether or not it is involved in sRNA-guided target degradation might be a subject of future studies. The general paucity of information on these vital cellular enzymes in *Bacteroidetes* offers exciting avenues of future research.

Bacteroidetes CRISPR systems

The large intestine harbors a vast consortium of bacteriophages that shape the microbiota composition and impose a high selection pressure on gut-resident bacteria (Mirzaei and Maurice 2017) (Figure 1). CRISPR-Cas systems are present in many prokaryotic species and provide adaptive immunity against phage infections (Marraffini 2015). Computational predictions indicate that about half of the sequenced *Bacteroidetes* species possess at least one CRISPR locus (Makarova et al. 2020; Pourcel et al. 2020) (Figure 3). While the classical type II-C system is the most frequent in *Bacteroidetes*, almost all of the known occurrences of the type VI system are restricted to this phylum (Makarova et al. 2020). Type VI CRISPR systems involve Cas13, which – unlike most other Cas nucleases – targets RNA rather than DNA. However, once activated through the recognition of transcribed phage RNA, Cas13 degrades nearby bacterial RNAs through its nonspecific RNase activity, thereby inducing bacterial dormancy upon phage infection (Meeske et al. 2019). These findings stem from *Listeria*; whether Cas13 plays a similar role for the persistence of *Bacteroidetes* populations exposed to the gut phageome is not yet known. From an evolutionary standpoint, as of now we can only speculate as to why certain intestinal *Bacteroidetes* species contain, while others lack, CRISPR systems, despite occupying similar host niches. The presumed burden associated with a type VI CRISPR system, for example, that might attack random RNA could outweigh its benefits in relation to other anti-phage systems (Hampton et al. 2020). This would particularly hold true for species such as *B. thetaiotaomicron* that can switch between multiple surface CPSs, which already provides a certain degree of protection against phage adsorption (Porter et al. 2020).

Apart from global, species-level predictions, CRISPR systems remain mostly unexplored in specific *Bacteroidetes* members. A notable exception is *B. fragilis*. Detailed inspection of strains with available genome sequence revealed that most (100 of 109) of the analyzed *B. fragilis* genomes carry at least one CRISPR system of type III-B, II-C or I-B (Tajkarimi and Wexler 2017). Seventy one strains also harbor a CRISPR array that lacks any associated Cas protein and is consistently found directly

Figure 3: Prevalence of CRISPR-Cas systems across the Bacteroidetes. Phylogenetic tree of 485 Bacteroidetes genome assemblies generated with ETE 3 (Huerta-Cepas et al. 2016), color-coded based on presence/absence of CRISPR systems as per (Makarova et al. 2020) and with “landmark” species highlighted. The background colors group tree branches that belong to the same phylogenetic class. While type VI CRISPR systems (purple squares) are quite common in Bacteroidetes, they have been barely observed outside this phylum.

upstream of the *hipAB* operon, with a putative role in persister cell formation and antibiotic resistance. This genomic co-localization led the authors to hypothesize that the “orphan” CRISPR could affect *hipAB* expression (Tajkarimi and Wexler 2017), but experimental validation of this hypothesis is still needed.

Commonalities and differences between regulatory RNA activities in Bacteroidetes and Proteobacteria

With just two *trans*-encoded sRNAs (RteR [Waters and Salyers 2012] and GibS [Ryan et al. 2020]) and one class of *cis*-antisense RNAs (DonS-like RNAs; Cao et al. 2016) functionally characterized, Bacteroidetes RNA research is still in its infancy. However, recent global transcriptomic approaches to *B. fragilis* (Cao et al. 2016), *B. thetaiotaomicron* (Ryan et al. 2020), and even an extra-intestinal Bacteroidetes member (Hirano et al. 2012; Høvik et al. 2012) have revealed a plethora of regulatory RNA candidates and suggest a bright future for this field. With respect to the nomenclature of Bacteroidetes noncoding RNA candidates, we recently introduced an analogous concept to that used for Proteobacteria, namely to name RNA genes “*BTncXXX*” where “*BT*” designates the species (here: *B. thetaiotaomicron*), “*nc*” refers to *noncoding*, followed by a three-digits number according to the ranked position on the chromosome. Only upon functional characterization, this operational identifier may be replaced by a trivial (four-letter) name. If adopted by the community and used consistently, this nomenclature should facilitate cross-comparison between independent studies.

The major housekeeping RNAs (tmRNA, 4.5S RNA, M1 RNA) are present in *Bacteroides* spp. (Prezza and Ryan et al. in prep; Ryan et al. 2020). High-resolution annotation of the *B. thetaiotaomicron* transcriptome (Ryan et al. 2020) further indicates that the numbers of representatives of individual regulatory RNA classes – e.g., 78 *cis*-antisense RNAs and 124 intergenic sRNAs – are similar to that reported in well-characterized bacteria from other phyla. Conservation of these newly identified *Bacteroides* RNAs, however, barely extends beyond the genus level, except for specialized sRNAs such as 6S RNA. Limited conservation seems to be a general feature of noncoding RNAs and promises many new RNA biological aspects to be learned from the study of

Bacteroidetes. Already now, careful inspection of the identified sRNA candidates in *B. thetaiotaomicron* enables some speculations regarding the commonalities and differences of Bacteroidetes sRNAs and their proteobacterial counterparts. For example, we found the average length and “structuredness” (i.e., the minimal free energy of *in silico*-predicted sRNA structures normalized by genomic GC content) of intergenic sRNAs in *B. thetaiotaomicron* to be similar to that of Proteobacteria (Prezza and Ryan et al. in prep).

Conversely, an obvious difference is the absence of Hfq homologs from Bacteroidetes. Whereas Gram-positive bacteria exemplify that sRNA regulation can occur without an assisting chaperone (Brantl and Brückner, 2014), this is relatively uncommon for Gram-negatives. Rather, there could be unrelated Bacteroidetes RBPs that functionally substitute Hfq. A second difference relates to the absence of a classical Shine-Dalgarno sequence from Bacteroidetes translation initiation regions (Nakagawa et al. 2010). Instead, there is a characteristic enrichment of adenine residues in the 5' UTR of mRNAs at positions –3, –6, and from –11 to –15 (relative to the start codon) that enhance translation efficiency (Baez et al. 2019). This region is targeted by GibS (in two out of three targets; (Ryan et al. 2020)) and is necessary and sufficient to repress *roc* mRNA, a mechanism that could be mediated by an unknown sRNA (Townsend et al. 2019). It is tempting to speculate that *trans*-acting sRNAs in the Bacteroidetes would compensate the absence of Hfq by evolving extended seed regions or utilizing multiple seeds to cover this entire region and mediate efficient regulation. If so, GibS-mediated repression of *BT_3893* – which involves two seed regions – would be a case in point, although the generalizability of these findings is unclear. In fact, for repression of a second *bona fide* GibS target (*BT_0771*) a single seed seems to be sufficient. Both of these GibS targets were repressed at the mRNA level; however, if mRNA decay is a secondary effect of the interference with translation initiation, or whether an RNase is actively recruited for target degradation has to be seen.

In one of the densest microbial ecosystems, fitness depends on efficient competition with companion microbes for nutrients. From a physiological perspective, regulatory RNAs are often implicated in the regulation of metabolic processes in Proteobacteria (Bobrovskyy and Vanderpool 2013), and this is likely even more so the case in specialized glycan degraders. For example, *Bacteroides* spp. evolved the unusual ability to metabolize more than a dozen plant- and host-derived polysaccharides (McNeil

1984; Salyers et al. 1977). However, expression of the large protein machineries to bind, process, and import complex carbohydrates is energetically costly. Therefore, *Bacteroides* would pay a high price if not able to tightly control their metabolic capacities. Transcriptional control on its own can only induce or shut off *de novo* synthesis of the corresponding mRNAs; however, clearance of the pool of pre-existing mRNAs from the cytosol upon sensing a preferred carbon source requires post-transcriptional mechanisms. In addition, RNA-mediated expression control allows fast adaptation, which should provide another advantage in the face of rapid nutrient fluctuations and fierce competition. We therefore expect that many more regulatory RNAs – in addition to the TPP and AdoCbl riboswitch, the DonS-like antisense RNAs, the *roc* leader, and GibS sRNA – modulate metabolism in Bacteroidetes.

Open questions and how to address them

This review provides an overview of the *status quo* of our collective knowledge of RNA biology in gut-associated Bacteroidetes. However, the field is just beginning to develop and many open questions remain to be addressed. For example, some of the best-studied *trans*-acting sRNAs in Proteobacteria regulate several dozens of target mRNAs, generating post-transcriptional regulatory networks with comparable complexity to regulons governed by transcription factors (Papenfort and Vogel 2009). Whether this applies also for Bacteroidetes sRNAs is currently unknown. The targets of the two *Bacteroides* sRNAs that have been functionally characterized were either identified by educated guesses (effect of RteR on the *tra* operon) or by differential expression analysis (Δ *gibS* mutant vs. wild-type vs. overexpression strain). The former bears the risk that additional targets are missed, whereas differential expression cannot distinguish direct from secondary effects. In contrast, technologies now routinely used for sRNA target screens in Proteobacteria, such as sRNA pulse-expression (Massé et al. 2005; Papenfort et al. 2006) (Figure 4a) or sRNA affinity purification followed by RNA-seq (Lalaouna et al. 2017), could be transferred to Bacteroidetes to identify direct sRNA target candidates in an unbiased, genome-wide manner. To validate *bona fide* targets, reporter systems such as *lacZ* (Huntzinger et al. 2005) or fluorescent target fusions (Urban and Vogel 2007) – provided the samples are given enough time under normoxic conditions for protein maturation to occur or use of oxygen-independent alternatives (Chia

et al. 2020) – could be harnessed (Figure 4b). Alternatively, luciferase-based reporter constructs – as used to dissect the mechanisms of TPP riboswitches in *Bacteroides* spp. (Costliow et al. 2019) – could be adopted for sRNA target verification.

As of now, we do not know whether the function of Bacteroidetes sRNAs depends on assisting chaperones or if regulatory RNAs work in a protein-independent manner in this phylum. Recent technological breakthroughs (Gerovac et al. 2020; Queiroz et al. 2019; Shchepachev et al. 2019; Smirnov et al. 2017; Urdaneta et al. 2019) led to the identification of novel RBPs even in species that have served us as RNA research models for decades (Attaiech et al. 2016; Pagliuso et al. 2019; Smirnov et al. 2016) and should likewise foster RBP discovery in Bacteroidetes (Figure 4c).

Which RNAs are employed by Bacteroidetes to efficiently colonize host niches? Dual RNA-seq (Westermann et al. 2017) of hypoxic cell culture models colonized with *Bacteroides* spp. (Figure 4d), or hybrid selection RNA-seq (Donaldson et al. 2020) of *Bacteroides* colonizing *in vivo* tissues have great potential to discover regulatory RNAs induced during host interaction. Enhanced expression often indicates functional relevance under the given condition and this could subsequently be tested by fitness screening of the respective knockout mutants. Alternatively, genome-wide perturbation screens such as transposon insertion sequencing (Cain et al. 2020) have been applied to uncover genetic factors contributing to *Bacteroides* fitness in the host (Goodman et al. 2009; Wu et al. 2015). Random mutagenesis, however, is inherently biased toward the disruption of longer genes, typically resulting in an underrepresentation of sRNA mutants in the library. Targeted approaches such as CRISPR interference, whose applicability was already demonstrated for *B. thetaiotaomicron* (Mimee et al. 2015) and that could simultaneously knock down hundreds of sRNAs, appear to be promising alternatives (Figure 4e).

Bacteroides spp. release outer membrane vesicles (OMVs) to share polysaccharide processing machineries with neighboring bacteria (Rakoff-Nahoum et al. 2014) and to deliver anti-inflammatory compounds to mammalian host cells (Shen et al. 2012). As bacterial vesicles can also contain ribonucleoprotein complexes, extracellular RNA molecules of pathogenic species shuttled via OMVs were suggested to mediate the cross-talk with co-colonizing bacteria or mammalian host cells (Koeppen et al. 2016; Lécivain and Beckmann 2020). RNA delivery may also occur in the opposite direction, i.e., from the host epithelium to bacterial microbiota members (Liu et al. 2016). *Bacteroides* spp. – at the interface of host-microbe and microbe-microbe encounters – are therefore promising

Figure 4: Technologies to foster functional studies of *Bacteroidetes* sRNAs.

a, sRNA target identification via pulse-expression. An sRNA is induced for a short time period (~10 min; symbolized by the timer) in the bacterial cell, followed by global RNA-seq to pinpoint altered mRNA levels in response to the sRNA pulse. b, Translational fusions of the predicted target region to a colorimetric, fluorescent or luminescent reporter gene can be employed to validate sRNA target transcripts and – through the introduction of point mutations – target sites. c, RBP identification with gradient sequencing (Grad-seq). Sedimentation of cellular RNA-protein complexes by density centrifugation and detection of RNA and protein molecules in the resulting fractions from low (LMW) to high molecular weight (HMW) by RNA-seq and mass-spectrometry (MS) can be used to screen for sRNA-binding proteins. d, Dual RNA-seq profiles bacterial and host gene expression during their interaction. It allows for the identification of *in vivo*-induced sRNAs and interspecies expression correlation may pinpoint host target processes of individual sRNAs. e, CRISPR interference (CRISPRi) screening for sRNA mutant fitness. Guide RNAs (gRNAs) designed against the sRNA complement of a bacterium and catalytically inactive Cas9 nuclease (dCas9) are introduced in a bacterial population (input), which is subsequently grown under a defined selection pressure (e.g., host colonization) and the remaining bacteria (output) are compared to the input pool to identify functionally important sRNAs under the screened condition.

model organisms to further explore the concept of functional extracellular RNA.

The *Bacteroides* community has made a strong commitment to make their global datasets easily accessible and easily usable by peers. It is thus worth emphasizing that multiple open-source databases exist that may be consulted for information on the *Bacteroides* genome and transcriptome, such as our Theta-Base (featuring transcriptome annotation and gene expression data for *B. thetaiotaomicron*; www.helmholtz-hiri.de/en/datasets/bacteroides; Ryan et al. 2020), the

Fitness Browser (Tn-seq data for *B. thetaiotaomicron* under a variety of conditions; <http://fit.genomics.lbl.gov>; Liu et al. 2019), and PULDB (an overview of predicted and published PUL systems across *Bacteroidetes*; www.cazy.org/PULDB_new/; Terrapon et al. 2018). These platforms provide excellent entry points for any functional study.

Bacteroidetes RNA research is beginning to prosper. With this review, we highlight recent progress made in this new field and hope to boost future studies. The case examples, arising concepts, and open questions reported

herein shall serve the community as a common starting ground for the years to come.

Acknowledgments: The authors thank Lars Barquist for insightful comments on this review.

Author contribution: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

Research funding: Work in the Westermann laboratory is funded by the German Research Foundation (DFG; project number: WE 6689/1-1).

Conflict of interest statement: The authors declare no conflicts of interest regarding this article.

References

- Adams, P., and Storz, G. (2020). Prevalence of small base-pairing RNAs derived from diverse genomic loci. *Biochim. Biophys. Acta Gene Regul. Mech.* <https://doi.org/10.1016/j.bbagr.2020.194524>. 32147527.
- Altman, S. (2011). Ribonuclease P. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 366: 2936–2941.
- Attaiech, L., Boughammoura, A., Brochier-Armanet, C., Allatif, O., Peillard-Fiorente, F., Edwards, R.A., Omar, A.R., MacMillan, A.M., Glover, M., and Charpentier, X. (2016). Silencing of natural transformation by an RNA chaperone and a multitarget small RNA. *Proc. Natl. Acad. Sci. U. S. A.* 113: 8813–8818.
- Bacic, M.K., and Smith, C.J. (2008). Laboratory maintenance and cultivation of *Bacteroides* species. *Curr. Protoc. Microbiol.* Chapter 13, Unit 13 C.11, <https://doi.org/10.1002/9780471729259.mc13c01s9>.
- Baez, W.D., Roy, B., McNutt, Z.A., Shatoff, E.A., Chen, S., Bundschuh, R., and Fredrick, K. (2019). Global analysis of protein synthesis in *Flavobacterium johnsoniae* reveals the use of Kozak-like sequences in diverse bacteria. *Nucleic Acids Res.* 47: 10477–10488.
- Bandyra, K.J., and Luisi, B.F. (2018). RNase E and the high-fidelity orchestration of RNA metabolism. *Microbiol. Spectr.* 6, <https://doi.org/10.1128/microbiolspec.RWR-0008-2017>.
- Barrick, J.E., Sudarsan, N., Weinberg, Z., Ruzzo, W.L., and Breaker, R.R. (2005). 6S RNA is a widespread regulator of eubacterial RNA polymerase that resembles an open promoter. *RNA* 11: 774–784.
- Bloom, S.M., Bijanki, V.N., Nava, G.M., Sun, L., Malvin, N.P., Donermeyer, D.L., Dunne, W.M. Jr., Allen, P.M., and Stappenbeck, T.S. (2011). Commensal *Bacteroides* species induce colitis in host-genotype-specific fashion in a mouse model of inflammatory bowel disease. *Cell Host Microbe* 9: 390–403.
- Bobrovskyy, M., and Vanderpool, C.K. (2013). Regulation of bacterial metabolism by small RNAs using diverse mechanisms. *Annu. Rev. Genet.* 47: 209–232. [10.1146/annurev-genet-111212-133445](https://doi.org/10.1146/annurev-genet-111212-133445).
- Brantl, S., and Brückner, R. (2014). Small regulatory RNAs from low-GC Gram-positive bacteria. *RNA Biol.* 11: 443–456.
- Breaker, R.R. (2012). Riboswitches and the RNA world. *Cold Spring Harb. Perspect. Biol.* 4, <https://doi.org/10.1101/cshperspect.a003566>.
- Cain, A.K., Barquist, L., Goodman, A.L., Paulsen, I.T., Parkhill, J., and van Opijnen, T. (2020). A decade of advances in transposon-insertion sequencing. *Nat. Rev. Genet.* 21: 1–15.
- Cao, Y., Förstner, K.U., Vogel, J., and Smith, C.J. (2016). *cis*-encoded small RNAs, a conserved mechanism for repression of polysaccharide utilization in *Bacteroides*. *J. Bacteriol.* 198: 2410–2418.
- Chao, Y., Li, L., Girodat, D., Förstner, K.U., Said, N., Corcoran, C., Šmiga, M., Papenfort, K., Reinhardt, R., and Wieden, H.J., et al. (2017). *In vivo* cleavage map illuminates the central role of RNase E in coding and non-coding RNA pathways. *Mol. Cell.* 65: 39–51.
- Chatzidaki-Livanis, M., Coyne, M.J., and Comstock, L.E. (2009). A family of transcriptional antitermination factors necessary for synthesis of the capsular polysaccharides of *Bacteroides fragilis*. *J. Bacteriol.* 191: 7288–7295.
- Chatzidaki-Livanis, M., Weinacht, K.G., and Comstock, L.E. (2010). Trans locus inhibitors limit concomitant polysaccharide synthesis in the human gut symbiont *Bacteroides fragilis*. *Proc. Natl. Acad. Sci. U. S. A.* 107: 11976–11980.
- Chia, H.E., Zuo, T., Koropatkin, N.M., Marsh, E.N.G., and Biteen, J.S. (2020). Imaging living obligate anaerobic bacteria with bilin-binding fluorescent proteins. *Curr. Res. Microb. Sci.* 1: 1–6.
- Costliow, Z.A., and Degnan, P.H. (2017). Thiamine acquisition strategies impact metabolism and competition in the gut microbe *Bacteroides thetaiotaomicron*. *mSystems* 2: e00116–17.
- Costliow, Z.A., Degnan, P.H., and Vanderpool, C.K. (2019). Thiamine pyrophosphate riboswitches in *Bacteroides* species regulate transcription or translation of thiamine transport and biosynthesis genes. *bioRxiv*, <https://www.biorxiv.org/content/10.1101/867226v1>.
- Coyne, M.J., Weinacht, K.G., Krinos, C.M., and Comstock, L.E. (2003). Mpi recombinase globally modulates the surface architecture of a human commensal bacterium. *Proc. Natl. Acad. Sci. U. S. A.* 100: 10446–10451.
- Degnan, P.H., Barry, N.A., Mok, K.C., Taga, M.E., and Goodman, A.L. (2014). Human gut microbes use multiple transporters to distinguish vitamin B₁₂ analogs and compete in the gut. *Cell Host Microbe* 15: 47–57.
- D’Elia, J.N., and Salysers, A.A. (1996). Effect of regulatory protein levels on utilization of starch by *Bacteroides thetaiotaomicron*. *J. Bacteriol.* 178: 7180–7186.
- Donaldson, G.P., Lee, S.M., and Mazmanian, S.K. (2016). Gut biogeography of the bacterial microbiota. *Nat. Rev. Microbiol.* 14: 20–32.
- Donaldson, G.P., Chou, W.C., Manson, A.L., Rogov, P., Abeel, T., Bochicchio, J., Ciulla, D., Melnikov, A., Ernst, P.B., and Chu, H., et al. (2020). Spatially distinct physiology of *Bacteroides fragilis* within the proximal colon of gnotobiotic mice. *Nat. Microbiol.* 5: 746–756.
- Dugar, G., Herbig, A., Förstner, K.U., Heidrich, N., Reinhardt, R., Nieselt, K., and Sharma, C.M. (2013). High-resolution transcriptome maps reveal strain-specific regulatory features of multiple *Campylobacter jejuni* isolates. *PLoS Genet.* 9: e1003495.

- Georg, J., and Hess, W.R. (2018). Widespread antisense transcription in prokaryotes. *Microbiol. Spectr.* 6: 191–210.
- Gerovac, M., El Mouali, Y., Kuper, J., Kisker, C., Barquist, L., and Vogel, J. (2020). Global discovery of bacterial RNA-binding proteins by RNase-sensitive gradient profiles reports a new FinO domain protein. *RNA*. in press. <https://doi.org/10.1261/rna.076992.120>.
- Goldstein, E.J. (1996). Anaerobic bacteremia. *Clin. Infect. Dis.* 23(Suppl. 1): S97–101.
- Goodman, A.L., McNulty, N.P., Zhao, Y., Leip, D., Mitra, R.D., Lozupone, C.A., Knight, R., and Gordon, J.I. (2009). Identifying genetic determinants needed to establish a human gut symbiont in its habitat. *Cell Host Microbe* 6: 279–289.
- Hampton, H.G., Watson, B.N.J., and Fineran, P.C. (2020). The arms race between bacteria and their phage foes. *Nature* 577: 327–336.
- Hirano, T., Beck, D.A., Demuth, D.R., Hackett, M., and Lamont, R.J. (2012). Deep sequencing of *Porphyromonas gingivalis* and comparative transcriptome analysis of a LuxS mutant. *Front. Cell Infect. Microbiol.* 2: 79.
- Hör, J., Matera, G., Vogel, J., Gottesman, S., and Storz, G. (2020). Trans-acting small RNAs and their effects on gene expression in *Escherichia coli* and *Salmonella enterica*. *EcoSal Plus* 9, <https://doi.org/10.1128/ecosalplus.esp-0030-2019>.
- Høvik, H., Yu, W.H., Olsen, I., and Chen, T. (2012). Comprehensive transcriptome analysis of the periodontopathogenic bacterium *Porphyromonas gingivalis* W83. *J. Bacteriol.* 194: 100–114.
- Holmqvist, E., and Vogel, J. (2018). RNA-binding proteins in bacteria. *Nat. Rev. Microbiol.* 16: 601–615.
- Holmqvist, E., and Wagner, E.G.H. (2017). Impact of bacterial sRNAs in stress responses. *Biochem. Soc. Trans.* 45: 1203–1212.
- Huerta-Cepas, J., Serra, F., and Bork, P. (2016). ETE 3: reconstruction, analysis, and visualization of phylogenomic data. *Mol. Biol. Evol.* 33: 1635–1638.
- Huntzinger, E., Boisset, S., Saveanu, C., Benito, Y., Geissmann, T., Namane, A., Lina, G., Etienne, J., Ehresmann, B., and Ehresmann, C., et al. (2005). *Staphylococcus aureus* RNAIII and the endoribonuclease III coordinately regulate spa gene expression. *EMBO J.* 24: 824–835.
- Huttenhower, C., Gevers, D., Knight, R., Abubucker, S., Badger, J.H., Chinwalla, A.T., Creasy, H.H., Earl, A.M., FitzGerald, M.G., and Fulton, R.S., et al. (2012). Structure, function and diversity of the healthy human microbiome. *Nature* 486: 207–214.
- Jeters, R.T., Wang, G.R., Moon, K., Shoemaker, N.B., and Salys, A.A. (2009). Tetracycline-associated transcriptional regulation of transfer genes of the *Bacteroides* conjugative transposon CTnDOT. *J. Bacteriol.* 191: 6374–6382.
- Jose, B.R., Gardner, P.P., and Barquist, L. (2019). Transcriptional noise and exaptation as sources for bacterial sRNAs. *Biochem. Soc. Trans.* 47: 527–539.
- Kalvari, I., Nawrocki, E.P., Argasinska, J., Quinones-Olvera, N., Finn, R.D., Bateman, A., and Petrov, A.I. (2018). Non-coding RNA analysis using the rfam database. *Curr. Protoc. Bioinformatics* 62: e51.
- Karp, P.D., Billington, R., Caspi, R., Fulcher, C.A., Latendresse, M., Kothari, A., Keseler, I.M., Krummenacker, M., Midford, P.E., and Ong, Q., et al. (2019). The BioCyc collection of microbial genomes and metabolic pathways. *Brief Bioinform.* 20: 1085–1093.
- Kavita, K., de Mets, F., and Gottesman, S. (2018). New aspects of RNA-based regulation by Hfq and its partner sRNAs. *Curr. Opin. Microbiol.* <https://doi.org/10.1016/j.mib.2017.10.014>. 29125938.
- Knight, D.J., and Girling, K.J. (2003). Gut flora in health and disease. *Lancet* 361: 1831.
- Koeppen, K., Hampton, T.H., Jarek, M., Scharfe, M., Gerber, S.A., Mielcarz, D.W., Demers, E.G., Dolben, E.L., Hammond, J.H., and Hogan, D.A., et al. (2016). A novel mechanism of host-pathogen interaction through sRNA in bacterial outer membrane vesicles. *PLoS Pathog.* 12: e1005672.
- Kortmann, J., and Narberhaus, F. (2012). Bacterial RNA thermometers: molecular zippers and switches. *Nat. Rev. Microbiol.* 10: 255–265.
- Kröger, C., Dillon, S.C., Cameron, A.D., Papenfort, K., Sivasankaran, S.K., Hokamp, K., Chao, Y., Sittka, A., Hébrard, M., and Händler, K., et al. (2012). The transcriptional landscape and small RNAs of *Salmonella enterica* serovar Typhimurium. *Proc. Natl. Acad. Sci. U. S. A.* 109: E1277–1286.
- Kröger, C., Colgan, A., Srikumar, S., Händler, K., Sivasankaran, S.K., Hammarlöf, D.L., Canals, R., Grissom, J.E., Conway, T., and Hokamp, K., et al. (2013). An infection-relevant transcriptomic compendium for *Salmonella enterica* serovar Typhimurium. *Cell Host. Microbe* 14: 683–695.
- Kröger, C., MacKenzie, K.D., Alshabib, E.Y., Kirzinger, M.W.B., Suchan, D.M., Chao, T.C., Akulova, V., Miranda-CasoLuengo, A.A., Monzon, V.A., and Conway, T., et al. (2018). The primary transcriptome, small RNAs and regulation of antimicrobial resistance in *Acinetobacter baumannii* ATCC 17978. *Nucleic Acids Res.* 46: 9684–9698.
- Lalaouna, D., Prévost, K., Eyraud, A., and Massé, E. (2017). Identification of unknown RNA partners using MAPS. *Methods* 117: 28–34.
- Lécroivain, A.L., and Beckmann, B.M. (2020). Bacterial RNA in extracellular vesicles: a new regulator of host-pathogen interactions? *Biochim. Biophys. Acta Gene Regul. Mech.* 1863: 194519.
- Lee, S.M., Donaldson, G.P., Mikulski, Z., Boyajian, S., Ley, K., and Mazmanian, S.K. (2013). Bacterial colonization factors control specificity and stability of the gut microbiota. *Nature* 501: 426–429.
- Lee, Y.K., Mehrabian, P., Boyajian, S., Wu, W.L., Selicha, J., Vonderfecht, S., and Mazmanian, S.K. (2018). The protective role of *Bacteroides fragilis* in a murine model of colitis-associated colorectal cancer. *mSphere* 3, <https://doi.org/10.1128/msphere.00587-18>.
- Ley, R.E., Turnbaugh, P.J., Klein, S., and Gordon, J.I. (2006). Microbial ecology: human gut microbes associated with obesity. *Nature* 444: 1022–1023.
- Liu, C.H., Lee, S.M., Vanlare, J.M., Kasper, D.L., and Mazmanian, S.K. (2008). Regulation of surface architecture by symbiotic bacteria mediates host colonization. *Proc. Natl. Acad. Sci. U. S. A.* 105: 3951–3956.
- Liu, S., da Cunha, A.P., Rezende, R.M., Cialic, R., Wei, Z., Bry, L., Comstock, L.E., Gandhi, R., and Weiner, H.L. (2016). The host shapes the gut microbiota via fecal microRNA. *Cell Host Microbe* 19: 32–43.
- Liu, H., Price, M.N., Carlson, H.K., Chen, Y., Ray, J., Shiver, A.L., Petzold, C.J., Huang, K.C., Arkin, A.P., and Deutschbauer, A.M. (2019). Large-scale chemical-genetics of the human gut bacterium *Bacteroides thetaiotaomicron*. *bioRxiv*, <https://www.biorxiv.org/content/10.1101/573055v1>.
- Lloréns-Rico, V., Cano, J., Kamminga, T., Gil, R., Latorre, A., Chen, W.H., Bork, P., Glass, J.I., Serrano, L., and Lluch-Senar, M. (2016). Bacterial antisense RNAs are mainly the product of transcriptional noise. *Sci. Adv.* 2: e1501363.
- Lotz, T.S., and Suess, B. (2018). Small-molecule-binding riboswitches. *Microbiol. Spectr.* 6. <https://doi.org/10.1128/microbiolspec.RWR-0025-2018>.

- Makarova, K.S., Wolf, Y.I., Iranzo, J., Shmakov, S.A., Alkhnbashi, O.S., Brouns, S.J.J., Charpentier, E., Cheng, D., Haft, D.H., and Horvath, P., et al. (2020). Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. *Nat. Rev. Microbiol.* 18: 67–83.
- Mao, C., Bhardwaj, K., Sharkady, S.M., Fish, R.I., Driscoll, T., Wower, J., Zwieb, C., Sobral, B.W., and Williams, K.P. (2009). Variations on the tmRNA gene. *RNA Biol.* 6: 355–361.
- Marraffini, L.A. (2015). CRISPR-Cas immunity in prokaryotes. *Nature* 526: 55–61.
- Martens, E.C., Chiang, H.C., and Gordon, J.I. (2008). Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. *Cell Host Microbe* 4: 447–457.
- Massé, E., Vanderpool, C.K., and Gottesman, S. (2005). Effect of RyhB small RNA on global iron use in *Escherichia coli*. *J. Bacteriol.* 187: 6962–6971.
- McCown, P.J., Corbino, K.A., Stav, S., Sherlock, M.E., and Breaker, R.R. (2017). Riboswitch diversity and distribution. *RNA* 23: 995–1011.
- McNeil, N.I. (1984). The contribution of the large intestine to energy supplies in man. *Am. J. Clin. Nutr.* 39: 338–342.
- Meeske, A.J., Nakandakari-Higa, S., and Marraffini, L.A. (2019). Cas13-induced cellular dormancy prevents the rise of CRISPR-resistant bacteriophage. *Nature* 570: 241–245.
- Mimee, M., Tucker, A.C., Voigt, C.A., and Lu, T.K. (2015). Programming a human commensal bacterium, *Bacteroides thetaiotaomicron*, to sense and respond to stimuli in the murine gut microbiota. *Cell Syst.* 1: 62–71.
- Mirzaei, M.K., and Maurice, C.F. (2017). Ménage à trois in the human gut: interactions between host, bacteria and phages. *Nat. Rev. Microbiol.* 15: 397–408.
- Miyakoshi, M., Chao, Y., and Vogel, J. (2015). Regulatory small RNAs from the 3' regions of bacterial mRNAs. *Curr. Opin. Microbiol.* 24: 132–139.
- Mohanty, B.K., and Kushner, S.R. (2018). Enzymes involved in posttranscriptional RNA metabolism in Gram-negative bacteria. *Microbiol. Spectr.* 6, <https://doi.org/10.1128/microbiolspec.RWR-0011-2017>.
- Nakagawa, S., Niimura, Y., Miura, K., and Gojobori, T. (2010). Dynamic evolution of translation initiation mechanisms in prokaryotes. *Proc. Natl. Acad. Sci. U. S. A.* 107: 6382–6387.
- Olejniczak, M., and Storz, G. (2017). ProQ/FinO-domain proteins: another ubiquitous family of RNA matchmakers? *Mol. Microbiol.* 104: 905–915.
- Pagliuso, A., Tham, T.N., Allemand, E., Robertin, S., Dupuy, B., Bertrand, Q., Bécavin, C., Koutero, M., Najburg, V., and Nahori, M.A., et al. (2019). An RNA-binding protein secreted by a bacterial pathogen modulates RIG-I signaling. *Cell Host Microbe* 26: 823–835, e811.
- Papenfort, K., and Vogel, J. (2009). Multiple target regulation by small noncoding RNAs rewires gene expression at the post-transcriptional level. *Res. Microbiol.* 160: 278–287.
- Papenfort, K., Pfeiffer, V., Mika, F., Lucchini, S., Hinton, J.C., and Vogel, J. (2006). SigmaE-dependent small RNAs of *Salmonella* respond to membrane stress by accelerating global omp mRNA decay. *Mol. Microbiol.* 62: 1674–1688.
- Peluso, P., Herschlag, D., Nock, S., Freymann, D.M., Johnson, A.E., and Walter, P. (2000). Role of 4.5S RNA in assembly of the bacterial signal recognition particle with its receptor. *Science* 288: 1640–1643.
- Porter, N.T., Hryckowian, A.J., Merrill, B.D., Fuentes, J.J., Gardner, J.O., Glowacki, R.W.P., Singh, S., Crawford, R.D., Snitkin, E.S., and Sonnenburg, J.L., et al. (2020). Multiple phase-variable mechanisms, including capsular polysaccharides, modify bacteriophage susceptibility in *Bacteroides thetaiotaomicron*. *bioRxiv*, <https://www.biorxiv.org/content/10.1101/521070v1>.
- Pourcel, C., Touchon, M., Villeriot, N., Vernadet, J.P., Couvin, D., Toffano-Nioche, C., and Vergnaud, G. (2020). CRISPRCasdb a successor of CRISPRdb containing CRISPR arrays and cas genes from complete genome sequences, and tools to download and query lists of repeats and spacers. *Nucleic Acids Res.* 48: D535–d544.
- Prezza, G., Ryan, D., Mädler, G., Reichardt, S., Barquist, L., and Westermann, A.J. (in prep). *Comparative genomics provides structural and functional insights into Bacteroides small RNA families*. In prep.
- Queiroz, R., Smith, T., Villanueva, E., Marti-Solano, M., Monti, M., Pizzinga, M., Mirea, D.-M., Ramakrishna, M., Harvey, R., Dezi, V., Thomas, G., Willis, A., and Lilley, K. (2019). Comprehensive identification of RNA-protein interactions in any organism using orthogonal organic phase separation (OOPS). *Nat. Biotechnol.*, <https://doi.org/10.1038/s41587-018-0001-2>. 30607034.
- Rakoff-Nahoum, S., Coyne, M.J., and Comstock, L.E. (2014). An ecological network of polysaccharide utilization among human intestinal symbionts. *Curr. Biol.* 24: 40–49.
- Reeves, A.R., Wang, G.R., and Salyers, A.A. (1997). Characterization of four outer membrane proteins that play a role in utilization of starch by *Bacteroides thetaiotaomicron*. *J. Bacteriol.* 179: 643–649.
- Rodionov, D.A., Vitreschak, A.G., Mironov, A.A., and Gelfand, M.S. (2002). Comparative genomics of thiamin biosynthesis in procaryotes. New genes and regulatory mechanisms. *J. Biol. Chem.* 277: 48949–48959.
- Romeo, T., and Babitzke, P. (2018). Global regulation by CsrA and its RNA antagonists. *Microbiol. Spectr.* 6, <https://doi.org/10.1128/microbiolspec.RWR-0009-2017>.
- Round, J.L., Lee, S.M., Li, J., Tran, G., Jabri, B., Chatila, T.A., and Mazmanian, S.K. (2011). The Toll-like receptor 2 pathway establishes colonization by a commensal of the human microbiota. *Science* 332: 974–977.
- Ryan, D., Jenniches, L., Reichardt, S., Barquist, L., and Westermann, A.J. (2020). A high-resolution transcriptome map identifies small RNA regulation of metabolism in the gut microbe *Bacteroides thetaiotaomicron*. *Nat. Commun.* 11: 3557.
- Santiago-Frangos, A., and Woodson, S.A. (2018). Hfq chaperone brings speed dating to bacterial sRNA. *Wiley Interdiscip. Rev. RNA* 9: e1475.
- Salyers, A.A., Vercellotti, J.R., West, S.E., and Wilkins, T.D. (1977). Fermentation of mucin and plant polysaccharides by strains of *Bacteroides* from the human colon. *Appl. Environ. Microbiol.* 33: 319–322.
- Sharma, C.M., and Vogel, J. (2014). Differential RNA-seq: the approach behind and the biological insight gained. *Curr. Opin. Microbiol.* 19: 97–105.
- Sharma, C.M., Hoffmann, S., Darfeuille, F., Reignier, J., Findeiss, S., Sittka, A., Chabas, S., Reiche, K., Hackermüller, J., and Reinhardt, R., et al. (2010). The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature* 464: 250–255.
- Shchepachev, V., Bresson, S., Spanos, C., Petfalski, E., Fischer, L., Rappsilber, J., and Tollervey, D. (2019). Defining the RNA interactome by total RNA-associated protein purification. *Mol. Syst. Biol.*, <https://doi.org/10.15252/msb.20188689>. 30962360.

- Shen, Y., Giardino Torchia, M.L., Lawson, G.W., Karp, C.L., Ashwell, J.D., and Mazmanian, S.K. (2012). Outer membrane vesicles of a human commensal mediate immune regulation and disease protection. *Cell Host Microbe* 12: 509–520.
- Sim, S., and Wolin, S.L. (2018). Bacterial Y RNAs: gates, tethers, and tRNA mimics. *Microbiol. Spectr.* 6, <https://doi.org/10.1128/microbiolspec.RWR-0023-2018>.
- Smirnov, A., Förstner, K.U., Holmqvist, E., Otto, A., Günster, R., Becher, D., Reinhardt, R., and Vogel, J. (2016). Grad-seq guides the discovery of ProQ as a major small RNA-binding protein. *Proc. Natl. Acad. Sci. U. S. A.* 113: 11591–11596.
- Smirnov, A., Schneider, C., Hör, J., and Vogel, J. (2017). Discovery of new RNA classes and global RNA-binding proteins. *Curr. Opin. Microbiol.* 39: 152–160.
- Sonnenburg, E.D., Sonnenburg, J.L., Manchester, J.K., Hansen, E.E., Chiang, H.C., and Gordon, J.I. (2006). A hybrid two-component system protein of a prominent human gut symbiont couples glycan sensing in vivo to carbohydrate metabolism. *Proc. Natl. Acad. Sci. U. S. A.* 103: 8834–8839.
- Sonnenburg, E.D., Zheng, H., Joglekar, P., Higginbottom, S.K., Firbank, S.J., Bolam, D.N., and Sonnenburg, J.L. (2010). Specificity of polysaccharide use in intestinal bacteroides species determines diet-induced microbiota alterations. *Cell* 141: 1241–1252.
- Storz, G., Vogel, J., and Wassarman, K.M. (2011). Regulation by small RNAs in bacteria: expanding frontiers. *Mol. Cell.* 43: 880–891.
- Sudarsan, N., Barrick, J.E., and Breaker, R.R. (2003). Metabolite-binding RNA domains are present in the genes of eukaryotes. *RNA* 9: 644–647.
- Svensson, S.L., and Sharma, C.M. (2016). Small RNAs in bacterial virulence and communication. *Microbiol. Spectr.* 4: 169–212.
- Tajkarimi, M., and Wexler, H.M. (2017). CRISPR-cas systems in *Bacteroides fragilis*, an important pathobiont in the human gut microbiome. *Front. Microbiol.* 8: 2234.
- Terrapon, N., Lombard, V., Drula, É., Lapébie, P., Al-Masaudi, S., Gilbert, H.J., and Henrissat, B. (2018). PULDB: the expanded database of polysaccharide utilization loci. *Nucleic Acids Res.* 46: D677–d683.
- Thomason, M.K., Bischler, T., Eisenbart, S.K., Förstner, K.U., Zhang, A., Herbig, A., Nieselt, K., Sharma, C.M., and Storz, G. (2015). Global transcriptional start site mapping using differential RNA sequencing reveals novel antisense RNAs in *Escherichia coli*. *J. Bacteriol.* 197: 18–28.
- Townsend, G.E. 2nd, Raghavan, V., Zwir, I., and Groisman, E.A. (2013). Intramolecular arrangement of sensor and regulator overcomes relaxed specificity in hybrid two-component systems. *Proc. Natl. Acad. Sci. U. S. A.* 110: E161–169.
- Townsend, G.E. 2nd, Han, W., Schwalm, N.D. 3rd, Raghavan, V., Barry, N.A., Goodman, A.L., and Groisman, E.A. (2019). Dietary sugar silences a colonization factor in a mammalian gut symbiont. *Proc. Natl. Acad. Sci. U. S. A.* 116: 233–238.
- Urban, J.H., and Vogel, J. (2007). Translational control and target recognition by *Escherichia coli* small RNAs *in vivo*. *Nucleic Acids Res.* 35: 1018–1037.
- Urdaneta, E.C., Vieira-Vieira, C.H., Hick, T., Wessels, H.H., Figini, D., Moschall, R., Medenbach, J., Ohler, U., Granneman, S., and Selbach, M., et al. (2019). Purification of cross-linked RNA-protein complexes by phenol-toluol extraction. *Nat. Commun.* 10: 990.
- Vitreschak, A.G., Rodionov, D.A., Mironov, A.A., and Gelfand, M.S. (2003). Regulation of the vitamin B12 metabolism and transport in bacteria by a conserved RNA structural element. *RNA* 9: 1084–1097.
- Vogel, J., and Luisi, B.F. (2011). Hfq and its constellation of RNA. *Nat. Rev. Microbiol.* 9: 578–589.
- Vogel, J., Bartels, V., Tang, T.H., Churakov, G., Slagter-Jäger, J.G., Hüttenhofer, A., and Wagner, E.G. (2003). RNomics in *Escherichia coli* detects new sRNA species and indicates parallel transcriptional output in bacteria. *Nucleic Acids Res.* 31: 6435–6443.
- Wade, J.T., and Grainger, D.C. (2014). Pervasive transcription: illuminating the dark matter of bacterial transcriptomes. *Nat. Rev. Microbiol.* 12: 647–653.
- Wagner, E.G.H., and Romby, P. (2015). Small RNAs in bacteria and archaea: who they are, what they do, and how they do it. *Adv. Genet.* 90: 133–208.
- Wagner, E.G., Altuvia, S., and Romby, P. (2002). Antisense RNAs in bacteria and their genetic elements. *Adv. Genet.* 46: 361–398.
- Wassarman, K.M., and Saecker, R.M. (2006). Synthesis-mediated release of a small RNA inhibitor of RNA polymerase. *Science* 314: 1601–1603.
- Wassarman, K.M. (2018). 6S RNA, a global regulator of transcription. *Microbiol. Spectr.* 6, <https://doi.org/10.1128/microbiolspec.RWR-0002-2017>.
- Waters, J.L., and Salyers, A.A. (2012). The small RNA RteR inhibits transfer of the conjugative transposon CTnDOT. *J. Bacteriol.* 194: 5228–5236.
- Wehner, S., Damm, K., Hartmann, R.K., and Marz, M. (2014). Dissemination of 6S RNA among bacteria. *RNA Biol.* 11: 1467–1478.
- Westermann, A.J., Barquist, L., and Vogel, J. (2017). Resolving host-pathogen interactions by dual RNA-seq. *PLoS Pathog.* 13: e1006033.
- Westermann, A.J. (2018). Regulatory RNAs in virulence and host-microbe interactions. *Microbiol. Spectr.* 6.
- Withey, J.H., and Friedman, D.I. (2003). A salvage pathway for protein structures: tmRNA and trans-translation. *Annu. Rev. Microbiol.* 57: 101–123.
- Woodson, S.A., Panja, S., and Santiago-Frangos, A. (2018). Proteins that chaperone RNA regulation. *Microbiol. Spectr.* 6, <https://doi.org/10.1128/microbiolspec.RWR-0026-2018>.
- Wu, M., McNulty, N.P., Rodionov, D.A., Khoroshkin, M.S., Griffin, N.W., Cheng, J., Latreille, P., Kerstetter, R.A., Terrapon, N., and Henrissat, B., et al. (2015). Genetic determinants of *in vivo* fitness and diet responsiveness in multiple human gut *Bacteroides*. *Science* 350: aac5992.

Bionotes

Daniel Ryan

Helmholtz Institute for RNA-based Infection Research (HIRI), Helmholtz Centre for Infection Research (HZI), Josef-Schneider-Str. 2/D15, D-97080, Würzburg, Germany
daniel.ryan@helmholtz-hiri.de
<https://orcid.org/0000-0003-4261-2702>

Daniel Ryan completed his undergraduate studies and his PhD in Biotechnology at KIIT University, Bhubaneswar, India in 2017. Since May 2018, he is a postdoctoral researcher in the Westermann laboratory.

Gianluca Prezza

Helmholtz Institute for RNA-based Infection Research (HIRI), Helmholtz Centre for Infection Research (HZI), Josef-Schneider-Str. 2/D15, D-97080, Würzburg, Germany
gianluca.prezza@helmholtz-hiri.de
<https://orcid.org/0000-0002-0032-4369>

Gianluca Prezza completed his MSc in Life Sciences at the University of Würzburg in 2017. He joined the Westermann group as a PhD student in October 2018.

Alexander J. Westermann

Helmholtz Institute for RNA-based Infection Research (HIRI), Helmholtz Centre for Infection Research (HZI), Josef-Schneider-Str. 2/D15, D-97080, Würzburg, Germany
 Institute of Molecular Infection Biology (IMIB), University of Würzburg, Josef-Schneider-Str. 2/D15, D-97080, Würzburg, Germany
alexander.westermann@uni-wuerzburg.de
<https://orcid.org/0000-0003-3236-0169>

Alexander J. Westermann studied Molecular Biosciences at the University of Heidelberg (Germany) and worked as a visiting scholar in 2009 at UC Berkeley (California, USA). He obtained his PhD and worked as a PostDoc in the lab of Prof. Jörg Vogel at the Institute of Molecular Infection Biology (IMIB) in Würzburg. In 2017 and 2018, he was a visiting researcher in the labs of Prof. Andreas Bäuml (UC Davis, USA) and Prof. David Holden (Imperial College London, UK). Since March 2018, he is a Junior Professor at the IMIB and independent group leader at the HIRI in Würzburg.

Our websites:

- <http://www.imib-wuerzburg.de/research/westermann/research/>
- https://www.helmholtz-hzi.de/en/research/research_topics/bacterial_and_viral_pathogens/host_pathogen_microbiota_interactions/our_research/