OPEN ACCESS

Bioinformatic and experimental methods to identify and validate bacterial RNA-human RNA interactions

Kylie I Krohmaly,[1,2] Robert J Freishtat,[1,3,4] Andrea L Hahn 🔟 [1,4,5]

[1]Center for Genetic Medicine Research, Children's National Research Institute, Washington, District of Columbia, USA
[2]Institute for Biomedical Sciences, The George Washington University School of Medicine and Health Sciences, Washington, District of Columbia, USA
[3]Division of Emergency Medicine, Children's National Hospital, Washington, District of Columbia, USA
[4]Department of Pediatrics, The George Washington University School of Medicine and Health Sciences, Washington, District of Columbia, USA
[5]Division of Infectious Diseases, Children's National Hospital, Washington, District of Columbia, USA

Correspondence to
Dr Andrea L Hahn, Children's National Hospital, Washington, District of Columbia, USA; alhahn@childrensnational.org

Accepted 31 August 2022

ABSTRACT

Ample evidence supports the importance of the microbiota on human health and disease. Recent studies suggest that extracellular vesicles are an important means of bacterial-host communication, in part via the transport of small RNAs (sRNAs). Bacterial sRNAs have been shown to co-precipitate with human and mouse RNA-induced silencing complex, hinting that some may regulate gene expression as eukaryotic microRNAs do. Bioinformatic tools, including those that can incorporate an sRNA's secondary structure, can be used to predict interactions between bacterial sRNAs and human messenger RNAs (mRNAs). Validation of these potential interactions using reproducible experimental methods is essential to move the field forward. This review will cover the evidence of interspecies communication via sRNAs, bioinformatic tools currently available to identify potential bacterial sRNA-host (specifically, human) mRNA interactions, and experimental methods to identify and validate those interactions.

SECTION 1: BACKGROUND AND EVIDENCE OF BACTERIA-TO-HUMAN COMMUNICATION VIA SMALL RNAS

The function and dysfunction of microbiota have been implicated in numerous human conditions, including inflammatory bowel diseases, asthma, and even infertility.[1–3] However, the mechanism by which the community of microbiota that reside on and in humans contribute to human disease remains poorly understood. More recent evidence supports that microbiota impact the host in part via molecular mechanisms induced by bacterial extracellular vesicles (EVs) and their contents.[4]

Bacterial EVs[5] are lipid bilayer-bound products deriving from the outer and/or inner membranes of bacteria[6] and can be between 10 and 500 nm depending on the species.[7] EVs can be isolated through various methods including filtration, ultracentrifugation, precipitation kits, or size-exclusion chromatography, and characterized by size, protein content, lipid content, function, and more.[8] There is no gold standard for bacterial EV isolation and characterization, however the minimal information for studies of EV guidelines updated in 2018 provide recommended protocols and reporting criteria to enhance reproducibility of methods.[8]

Bacterial small RNAs are among the RNA cargo of bacterial EVs

Bacterial EVs carry a range of cargo, including RNA, DNA, and proteins.[6 9 10] RNA sizes and types found within EVs include protein coding genes, transfer RNAs (tRNAs), ribosomal RNAs (rRNAs), and short non-coding RNAs.[9 11–13] Short non-coding RNAs are regulatory molecules that are able to modulate gene expression via base pairing with a potential target transcript.[14] Eukaryotic short non-coding RNAs often take the form of microRNAs (miRNA) which are processed from miRNA genes,[15] while in bacteria, short non-coding RNAs are called small RNAs (sRNAs) and are generally derived from intergenic regions.[16 17] tRNA fragments are considered a type of sRNA,[18 19] and are found in bacterial EVs.[9] Additionally, both rRNA and tRNA fragments have been implicated in inducing apoptosis in human cells.[20]

Whereas eukaryotic miRNAs are around 20 nucleotides (nts) long, bacterial sRNAs can be as long as 500+ nt,[21] but are generally between 50 and 400 nt. This extra length allows bacterial sRNAs to form complex secondary structures that impact the stability of the sRNA, how it regulates its target transcript, and how and if it binds its target.[22–28 29 30] Bacteria also produce miRNA-sized RNAs,[18 29–33] which may function similar to human miRNAs. sRNA is a consistent and often large portion of EV cargo.[9 11–13] Interestingly, bacterial EV RNA cargo is distinct from the intracellular RNA population,[9 18 34] and from the non-EV extracellular RNA population,[9] indicating that RNA is strategically loaded into EVs for their target cell.

Bacterial sRNAs circulate in the bloodstream and affect mammalian gene expression

Bacterial sRNAs[35–37] and particularly those encapsulated in bacterial EVs[38 39] have been detected in the plasma of both healthy individuals and those that are affected by various illnesses. Mammalian cells internalize bacterial EVs[11 40] and furthermore, EV cargo—specifically, their RNA cargo—localizes to both the

BMJ

cytoplasm[33][40] and the nucleus.[11][33][40] Exogenous bacterial EVs also traverse the blood-brain barrier.[33][40]

Bacterial sRNAs have been bioinformatically predicted and experimentally validated to affect certain pathways in human and murine cells. Using three periodontal pathogens' EVs, Choi *et al* confirmed the presence of miRNA-sized sRNAs.[32] Using bioinformatic tools, they predicted human immune system-related mRNA targets of several sRNAs.[32] Jurkat T cells exposed to synthetic oligos of the most highly expressed sRNAs showed that these sRNAs downregulated several cytokines (interleukin (IL)-5, IL-13, and IL-15).[32]

Additionally, Han *et al* investigated the potential of EV-derived RNAs from *Aggregatibacter actinomycetemcomitans* (Aa), a periodontal pathogen, to induce an immune response.[29] U937 cells exposed to both intact Aa EVs and lysed Aa EVs (both exposed to RNase) showed that intact EVs induced a greater tumor necrosis factor-alpha response than lysed EVs.[33] Compared with RNase-treated lysed EVs, cells exposed to DNase-treated lysed EVs also upregulated nuclear factor kappa B (NF-κB) activation and toll-like receptor 8 expression (both of which sense single-stranded RNAs); the authors attributed the difference in outcome to the RNAs of the DNase-treated lysed EVs.[35] The authors also found that BV2 (ie, murine microglial) cells activated NF-κB and upregulated IL-6 after exposure to DNase-treated lysed Aa EVs or intact EVs more so than RNase-treated lysed EVs.[41]

Bacterial sRNAs interact with mammalian RISC

RNA-induced silencing complex (RISC) in mammalian cells comprises Argonaute (Ago) and a miRNA, which then can bind the miRNA's target mRNA to result in its degradation.[15] Han *et al* described bacterial sRNAs that interact with host RISC after Aa EV exposure to U397 cells using total RNA sequencing (RNA-Seq) and RNA immunoprecipitation and sequencing (RIP-Seq).[33] While Aa-derived sRNAs only accounted for 0.02% of reads via RNA-Seq, 0.35% of reads originated from Aa when looking at RISC-associated sRNA through RIP-Seq.[33]

In a similar study, Furuse *et al* investigated five pathogenic intracellular bacteria which produced RISC-enriched RNAs (compared with total RNA levels) that were close to the same length as authentic eukaryotic miRNAs.[29] Additionally, one bacterial species, *Mycobacterium marinum* (Mm), produced an RNA, MM-H, the expected size (23 nt) with similar characteristics to canonical miRNAs (ie, stem-loop structure prediction, discrete 5' end consistent with Dicer trimming).[29]

These studies demonstrate that bacterial sRNAs have the potential to operate as human regulatory RNAs via RISC. Bacterial sRNAs—specifically, bacterial miRNA-sized RNAs—interacting with RISC are a likely mechanism by which some sRNAs exert their impact on host gene expression.[29][33] While these findings taken together are very compelling, more studies are needed to confirm this potential mechanism and investigate any off-target effects (perhaps relating to the machinery itself), or the possibility that host miRNA may be targeting the bacterial sRNAs. Nonetheless, discovering that sRNAs interact with miRNA machinery is encouraging evidence that the interactions and

resulting effects may be sequence-specific rather than due to foreign single-stranded RNA.

Although Furuse *et al* investigated if MM-H interacted with a perfectly complementary target using a reporter assay,[29] no human target prediction was performed for the sRNA. Future studies will need to investigate the potential specific targets of the discovered sRNAs more deeply. There is a need for identification and validation to determine sRNA candidates for microbial-host communication in conditions for which host microbiota have been implicated (eg, inflammatory bowel disease or asthma).[1][2] In the following two sections, we will discuss bioinformatic methods (table 1) and experimental methods (table 2) that could be employed to discover these interactions and identify and validate potential human targets of bacterial sRNAs.

SECTION 2: BIOINFORMATIC APPROACHES FOR PREDICTING BACTERIAL SRNA-HUMAN MRNA INTERACTIONS

Homology and alignment tools for sRNA target prediction in silico

As RNA-RNA interactions are mediated via base pairing, sequence complementarity for at least the seed region (ie, the perfectly complementary region that 'seeds' the interaction) is required. This region of the sRNA may be conserved across species, and a novel sRNA's function in one species may be similar to that of a homologous or identical sRNA (ie, a perfect match after alignment) in another species.

Koeppen *et al* identified human targets for 10 *Pseudomonas aeruginosa* sRNAs in EVs via *BLAST* (and other tools discussed below), which they then used to guide their functional validation.[18] Jorth and Whitely, after identifying and characterizing a lysine riboswitch controlling a lysine transporter in Aa, identified a homologous transporter gene in *Haemophilus influenzae* (Hi) through BLASTp.[42] They confirmed the riboswitch's effect on the transporter in Hi, demonstrating conservation of the system across the two species.[42] In other words, they aligned one species' sRNA target in BLAST in order to determine if the latter produced a similar target. This type of approach can be ued in identifying sRNA targets in different genomes; after identifying the transcript and homologous transcripts in other genomes, and one can then go backwards from there to identify sRNAs that target the transcript. BLAST (and other alignment tools combined with appropriate databases such as bowtie[41]) can also be useful in identifying sRNAs themselves through alignment to sRNA databases—with whole sequences, and with portions of sequences that may be conserved. For more tools on sRNA prediction and detection, refer to table 3.

The use of alignment and homology to annotate sRNAs with targets in different species (and even within a species) is currently limited in part because there may be species-specific sRNAs, and because there are still many sRNAs that need functional validation and many species left to investigate. For example, one group assigned functions to detected bacterial sRNAs using *Rfam*[43] and other databases by aligning the detected sRNAs to the annotated sRNA sequences within the databases.[44] However, only 19 of 422 sRNAs were assigned a function.[44] These databases can contain predicted and experimentally validated sRNAs.

Table 1 Summary of bioinformatic tools that can be used to predict bacterial RNA-human RNA targets

Tool	Uses energy*	Uses accessibility	Homology/ Conservation	Reference	GitHub or website link
Rfam			✓	43	https://rfam.xfam.org
BLAST			✓	89	https://blast.ncbi.nlm.nih.gov/Blast.cgi
RNAhybrid	✓			46	https://bibiserv.cebitec.uni-bielefeld.de/rnahybrid?id=rnahybrid_view_submission
RNAnue	✓	✓		53	https://github.com/lbvt/RNAnue
RNAup	✓	✓		23	https://www.tbi.univie.ac.at/RNA/RNAup.1.html
RNAplex	✓	✓		24	https://www.tbi.univie.ac.at/RNA/RNAplex.1.html
IntaRNA	✓	✓		22 27	http://rna.informatik.uni-freiburg.de/IntaRNA/Input.jsp;jsessionid=41FFC5A7C2ED917ED3C204B6E3D9CB6B
RIsearch2	✓	✓		47	http://rth.dk/resources/risearch
DGE and Correlation Analysis				55 56	—
IPA (TargetScan)			✓	59 62	https://www.qiagen.com/us/products/discovery-and-translational-research/next-generation-sequencing/informatics-and-data/interpretation-content-databases/ingenuity-pathway-analysis/
miRanda	✓		✓	64	http://cbio.mskcc.org/miRNA2003/miranda.html
DIANA-microT	✓		✓	66	https://diana.e-ce.uth.gr/home

*Negative energy is associated with a higher likelihood of the interaction taking place.
DGE, differential gene expression; IPA, ingenuity pathway analysis.

sRNA alignment and homology may be more widely used after more sRNA targets have been identified and fully characterized.

Importantly, alignment to directly determine sRNA targets (ie, a reverse-complement match to an sRNA) may be less reliable, however, as the secondary structure of the sRNA and mRNA are not taken into account in alignment tools. This is important to consider because although a sequence match may be found between a pair, the sequences may not be able to interact—for example, the seed region may be occluded.

Energy-based and accessibility-based models consider secondary structure

Besides homology, energy-based models are RNA-RNA interaction prediction algorithms and pipelines that calculate interaction or hybridization energy; that is, the energy

Table 2 Summary of experimental approaches that can be used to identify and validate bacterial RNA-human RNA interactions

Method	Requires knowledge of sRNA or mRNA beforehand	Uses anchor to pull or enrich RNAs	Crosslinking/Ligation	Reference
EMSA	✓ (both)			49 67–69
Reporter systems	✓ (both)			29 49
Expression alteration	✓ (both)			70 71
RNA Pull-Down	✓ (both)			72
MAPS	✓ (one)	✓ (MS2-RNA)		73–75
RIP-Seq		✓ (Ago)		29 33
qCLASH		✓ (Ago, Hfq)	✓	84 85
RIL-Seq		✓ (Hfq)	✓	86
RAP-RNA		✓ (biotinylated probe)	✓	80
MARIO		✓ (biotinylated cysteine residues)	✓	88
Modified CLASH			✓	76
LIGR-Seq			✓	77
PARIS			✓	78
SPLASH		✓ (biotinylated psoralin)	✓	79
RIC-Seq		✓ (biotinylated cytidine (bis) phosphate-3′ RNA ends)	✓	87

EMSA, electrophoretic mobility shift assay; LIGR-Seq, ligation of interacting RNA followed by high throughput sequencing; MAPS, MS2-affinity purification coupled with RNA sequencing; MARIO, mapping RNA interactome in vivo; mRNA, messenger RNA; PARIS, psoralen analysis of RNA interactions and structures; qCLASH, quick cross-linking and sequencing of hybrids; RAP-RNA, RNA antisense purification to systemically map RNA-RNA interactions; RIC-Seq, RNA in situ conformation; RIP-Seq, RNA immunoprecipitation and sequencing; SPLASH, sequencing of psoralen crosslinked, ligated, and selected hybrids; sRNA, small RNA.

Table 3 Summary of bioinformatic tools that detect bacterial small RNAs and their genes

Tool name	Input is bacterial genome	Input is sequencing data	Homology/Conservation based	Reference	GitHub or website link
Transterm	✓			[90]	https://transterm.cbcb.umd.edu
ARNold	✓			[91]	http://rssf.i2bc.paris-saclay.fr/toolbox/arnold/
RNIE	✓			[92]	https://github.com/ppgardne/RNIEs
sRNADetect		✓		[44]	https://github.com/BioinformaticsLabAtMUN/sRNA-Detect
Rockhopper		✓		[93]	https://cs.wellesley.edu/~btjaden/Rockhopper/
APERO		✓		[94]	https://github.com/Simon-Leonard/APERO/blob/master/example/script.R
DET'RPROK		✓		[95]	http://rssf.i2bc.paris-saclay.fr/Software/detrprok.php
BLAST		✓	✓	[89]	https://blast.ncbi.nlm.nih.gov/Blast.cgi
GLASSgo		✓	✓	[96]	https://github.com/lotts/GLASSgo

required for two RNAs to interact via base pairing.[22–24] An interaction is favorable when the energy is negative, which means that it is a spontaneous interaction,[45] and is used as a proxy to interaction stability.[27] One such model is *RNAhybrid*.[46] However, RNAhybrid does not consider the secondary structure of the two RNAs before they interact.[46] Importantly, the binding site that is predicted in RNAhybrid may be concealed and unable to interact in its natural secondary structure, generating a false positive 'hit'. Other tools take secondary structure into account by considering the 'accessible energy' of the interaction, which is the energy needed to make the binding site accessible or open,[22 23 27 47] or the probability that the binding site is unpaired in the secondary structure.[24] In these models, the reported energy is typically the sum of accessible energy and hybridization or interaction energy,[22 23 27] with the lowest energy being a more favorable or plausible interaction.

RNAup,[23] *RNAplex*,[24] and *IntaRNA*[22 27] are designed for short non-coding RNA interaction prediction and consider both hybridization and accessible energy. Generally, at least one sRNA and putative target are required. These targets may be a few mRNAs to a whole transcriptome or genome, as in IntaRNA.[22 27] RNAplex and IntaRNA are able to be applied to both bacterial sRNA and miRNA target prediction.[24 27] Umu and Gardner performed a benchmark study of 15 RNA-RNA interaction prediction tools comparing datasets of eukaryotic, bacterial, and archaeal short RNAs along with their RNA targets.[48] Overall, energy-based and accessibility-based tools performed the best in terms of true positive rates (sensitivity) and positive predictive values (precision).[48] Specifically, RNAup overall performed the best, followed by IntaRNA, and RNAplex.[48]

From the developers of GLASSgo, IntaRNA is an energy-based and accessibility-based RNA-RNA interaction prediction tool that predicts both prokaryotic[49] and eukaryotic interactions.[50] Recently, IntaRNA has been incorporated into ViRBase V.3.0, a virus-host interaction database, to aid in interspecies RNA-RNA interaction binding site prediction.[51] Uniquely, in addition to outputting the potential binding sites and energy scores as RNAplex and RNAup do, IntaRNA provides minimal energy profiles that allow visual exploration of interaction patterns.[27]

RIsearch2[47] facilitates exploration of predicted sRNA targets by integrating experimental transcriptomic data. This pipeline assumes that there are target (intended) interactions and off-target interactions, and calculates targeting probabilities for each potential target of an sRNA using interaction probabilities derived from both energies.[47] As output, RIsearch2 reports the predicted interaction target, the start and stop positions with a visualized binding site, and the energy of the interaction.[52]

Similarly, *RNAnue*[53] is a bioinformatic pipeline that analyzes experimental RNA-RNA interaction data derived from crosslinking and ligation methods (discussed in section 3). RNAnue considers hybridization and accessible energy, and complementarity of RNA partners.[53] The main outputs of RNAnue are alignments of the RNA partners (chimeras, discussed further in section 3) to the genome and complementarity and energy scores.[53] Overall, when benchmarked using experimental data produced by the methods, the RNAnue pipeline outperformed the original bioinformatic pipelines.[53]

Bioinformatic tools are able to predict interkingdom RNA-RNA interactions

Dual RNA sequencing (dRNA-Seq) is simply the simultaneous sequencing of two or more organisms without enriching or depleting any RNA type.[54] After sequencing, reads are aligned to human and bacterial genomes in parallel, and subsequent differential expression, pathway, or network analysis, and more can be performed to detect sRNAs in the dataset and determine their targets.[55 56] Westermann *et al* used dRNA-Seq and *differential gene expression (DGE) and correlation analyses* to identify a novel *Salmonella enterica* sRNA, PinT, after infecting HeLa cells.[55 56] Through this strategy, they also determined that PinT regulates bacterial virulence factors which impact host signaling.[55 56] Although this sRNA was not shown to have a direct human RNA target, this technique is still applicable to identifying direct interspecies RNA-RNA interactions.

Ingenuity pathway analysis (*IPA*)[57] is a bioinformatic tool that allows transcriptome integration and pathway visualization facilitating sRNA target exploration and has been employed to predict bacterial sRNA-induced human gene expression changes.[18] IPA was developed for eukaryotic transcriptome data (such as that generated from dRNA-Seq[58]), as well as metabolome and proteome

data analysis. IPA integrates experimentally validated interactions from miRNA-mRNA interaction databases, and TargetScan, a miRNA-mRNA interaction prediction tool.[59] Using IPA, Mika-Gospodorz et al identified differentially expressed genes and affected pathways in host cells infected with an intracellular bacterium,[60] while Koeppen et al analyzed which human proteins were impacted by the addition of their synthetic sRNA of interest in human bronchial epithelial cells.[18] Using a platform such as IPA to analyze dual RNA sequencing data[54 58 61] would facilitate target prediction and contextualization within the biological system. *TargetScan* (currently at release V.8.0) predicts miRNA targets by scanning 6–8 nt long sites on mRNAs complementary to the miRNA.[62 63] It requires the selection of a species supported by TargetScan in addition to either a gene or transcript ID and/or select a conserved miRNA family, miRBase annotation, or that the user enter the name of a known miRNA.[62 63]

Koeppen et al also predicted interkingdom RNA-RNA interactions between bacterial sRNAs and human mRNAs using *miRanda*.[18] miRanda computes miRNA targets by assessing complementarity of the two input sequences and then calculating free energy of the interaction; the lowest (most favorable) energy of the resulting linked sequence structure is computed[64] miRanda also considers conservation of the miRNA by scanning it against 3' UTR datasets.[64] Additionally, Choi et al used *DIANA-microT* for the same purpose.[32] Using alignment of the driver sequence (ie, the first 9 nt of the miRNA, to databases of targets' 3' UTRs[65] and coding sequences[66]), the algorithm identifies miRNA recognition elements, which are 7–9 consecutive and complementary nucleotides to the driver sequence.[65] DIANA-microT uses RNAhybrid to compute hybridization energy for sites with fewer than seven consecutive complementary nucleotides.[65] Pairings are ranked based on conservation of the miRNA recognition elements in several species.[65] Additionally, DIANA-microT contains integrated pipelines that can, for example, analyze expression data or perform enrichment analysis of miRNA-target pairings.[66] While all three tools were designed to predict human miRNA targets, Koeppen et al and Choi et al have established their use in predicting interspecies—even more, interkingdom—RNA-RNA interactions (ie, human targets of bacterial sRNAs).[18 32]

SECTION 3: EXPERIMENTAL APPROACHES FOR IDENTIFYING AND VALIDATING BACTERIAL SRNA-HUMAN MRNA INTERACTIONS
Use of two known interacting RNA partners to validate an interaction
Originally used to investigate nucleic acid-protein interactions, an electrophoretic mobility shift assay (*EMSA*) can validate the interaction of two suspected RNA partners. This first involves in vitro transcription of the two RNAs and labeling of one of them for detection, such as radiolabel[49 67 68] or digoxigenin.[69] The two RNAs are then incubated and shifted on a gel together; if the two RNAs interact and form a duplex, a second radiolabeled band will form and be visible.[49 67–69]

Reporter systems are also helpful in determining the interaction of two known RNAs. Furuse et al used a Renilla

luciferase (Rluc) reporter assay to determine if MM-H would bind and repress two perfectly complementary sequences.[29] One modification to this approach is to clone a potential human target transcript into the Rluc vector, providing a clearer picture of if MM-H can repress a human transcript. Also, in *lacZ* reporter system by Mai et al, derepression of the sRNA will decrease beta-galactosidase activity if it interacts and binds with the candidate transcript's 5' UTR.[49] Mai et al observed this decrease in beta-galactosidase activity after 6C sRNA was derepressed, indicating its binding to its target, the *lacZ*-fused *panD* transcript.[49]

Alteration of sequence and expression can similarly determine if two known RNAs interact. Zhang et al, interested in cotton-fungal interactions, identified several cotton miRNAs in infecting fungal hyphae through deep sequencing of sRNA.[70] Zhang et al mutated fungal transcripts to eliminate base pairing with the identified miRNAs.[70] After transformation into the fungus, expression and cotton infection, the resistant transcripts accumulated while the wild-type transcripts were silenced, demonstrating direct targeting by the cotton miRNAs.[70] Similarly, Jiao and Peng overexpressed and knocked-down a particular wheat miRNA (using a viral-transformed bacterial infiltration method); miRNA overexpression resulted in downregulation of target fungal transcripts, while knocked-down miRNA allowed an accumulation of target transcripts.[71]

RNA pull-down assays can demonstrate direct RNA-RNA interaction between known RNAs.[72] Here, both interacting partners must be known, as transfection of one known RNA occurs and the end point is reverse transcription-quantitative PCR (RT qPCR), which requires primers of both sequences for detection. In this technique, a biotin-labelled known RNA of interest is transfected into cells and after lysis, the cells are added to magnetic streptavidin beads.[72] The known biotinylated RNA base pairing to its suspected partner are pulled down using the beads, extracted, then amplified with RT qPCR using primers of both sequences.[72] For Yang et al, base-pairing without crosslinking was enough to pull-down the pair and allow detection of the target RNA via RT-qPCR.[72] This method is helpful in validating a few RNA-RNA interactions at a time, but again the interacting pairs need to be known beforehand.

Targeted identification to enable identification of one or both RNA partners
'MS2-affinity purification coupled with RNA sequencing' (*MAPS*) requires knowledge of one of the two sequences.[73–75] First, an MS2 (a coat protein from a bacteriophage) aptamer is fused to a sequence of interest and expressed in vivo.[73–75] Lysates undergo affinity chromatography using the MS2 maltose binding protein, which binds the MS2 aptamer with high affinity and specificity,[75] to enrich RNAs fused to the MS2 aptamer and those base pairing with fused RNAs.[73–75] RNA is then extracted and sequenced.[73–75] This method involves manipulation of the native RNA of interest, which could potentially impact the secondary structure of the RNA or its binding. In contrast to RNA pull-down assays, only one interacting partner needs to be known as MAPS uses sequencing of the associated RNAs instead of RT qPCR.

Furuse et al and Han et al used RIP-Seq to identify microRNA-sized bacterial sRNAs associated with
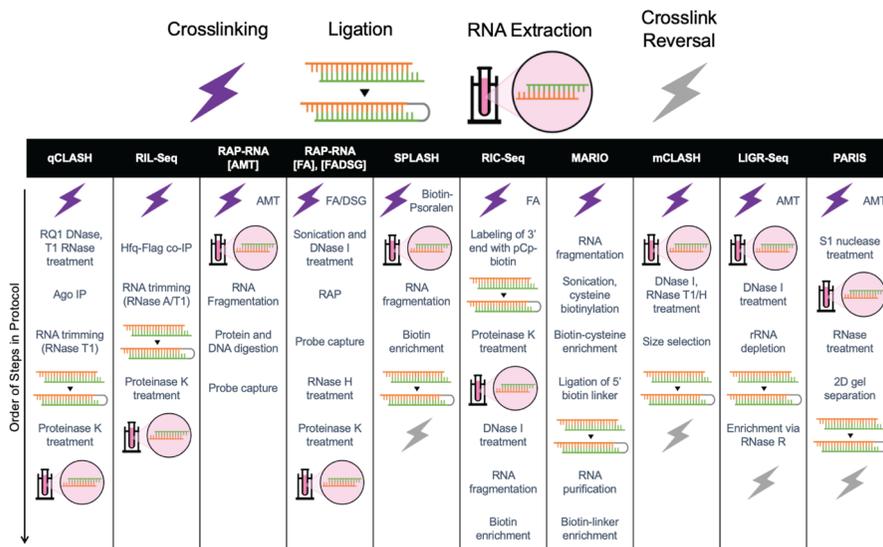
Figure 1 Summary and comparison of crosslinking and ligation methods. Crosslinking, RNA extraction, ligation, crosslink reversal, and RNA purification and sequencing represent a core set of steps shared by many discussed methods and are indicated with images. For each method, the protocol is listed stepwise in descending order, starting at crosslinking. Differences between the methods are indicated with text as they appear in the protocol. RNA purification and sequencing (not depicted) is always the final step in the protocol. 2D, two-dimensional; AMT, 4'-aminomethyltrioxsalen; DSG, disuccinimidyl glutarate; FA, formaldehyde; LIGR-Seq, ligation of interacting RNA followed by high throughput sequencing; PARIS, psoralen analysis of RNA interactions and structures; RAP-RNA, RNA antisense purification to systemically map RNA-RNA interactions; RIC-Seq, RNA in situ conformation; RIL-Seq, RNA interaction by ligation and sequencing; rRNA, ribosomal RNA.

mammalian RISC.[29][33] After infection with intracellular bacteria[29] or dosage with EVs,[33] cells are lysed and incubated with beads coated with anti-Ago antibodies. RNA is then extracted and subjected to deep sRNA-Seq.[29][33] This technique allows identification of mammalian RISC-associated bacterial sRNAs, and previously discussed bioinformatic tools may be employed to help with this.

Crosslinking and/or ligation methods to enable global or near-global identification of interacting RNA partners

As some RNA-RNA interactions are facilitated by certain proteins, crosslinking RNA to proteins and/or ligation to the linked RNA with subsequent pull-down of specific anchor proteins known to facilitate or be involved with RNA-RNA interactions can aid in identifying targets of sRNAs that use that protein. Generally, techniques that involve both crosslinking and ligation follow these steps: (1) cells are grown and nearby RNAs or proteins are crosslinked; (2) cells are lysed and RNA is extracted; (3) crosslinked RNAs are ligated at one or both ends, forming a chimera; (4) crosslinks are reversed; (5) RNA is purified, prepared for sequencing and sequenced (figure 1). Timing of RNA pull-down depends on the technique. Crosslinking can be achieved by several molecules, but psoralen-derivative 4'-aminomethyltrioxsalen (AMT) is commonly used.[76–80] AMT inserts itself in nucleic acids and on irradiation with 365 nm UV light, AMT will covalently bind two nearby RNAs together (intermolecularly and intramolecularly), with a preference for thymine or uridine,[81][82] which may introduce bias for regions containing more uridine.[83] Formaldehyde crosslinks RNA and protein to protein, while disuccinimidyl glutarate crosslinks protein to protein.[80] Crosslinking RNA with protein,

and protein to protein, enables inclusion of RNA whose direct interactions are facilitated by proteins, and also RNA indirectly involved through peripheral proteins.[80] Conveniently, these crosslinks are reversed on irradiation with 254 nm UV light. To our knowledge, the methods discussed in this section have not been applied to identify direct RNA-RNA interactions between bacteria and humans.

Gay *et al* debuted a 'quick cross-linking and sequencing of hybrids' (*qCLASH*) that enables the pull-down of Ago-crosslinked RNAs.[84] In this technique, only UV irradiation crosslinks RNAs to proteins. Specifically, Ago proteins, along with the crosslinked RNAs, are then immunoprecipitated on beads.[84] RNA pairs are ligated while on the beads and Ago is digested, while RNA is purified and sequenced.[84] In a similar CLASH, Hfq, a bacterial chaperone that interacts with sRNAs, and associated RNAs are pulled down.[85]

'RNA interaction by ligation and sequencing' (*RIL-Seq*) similarly targets Hfq and has very similar general steps to qCLASH and the Hfq-CLASH.[86] RIL-Seq, however, requires Hfq to express a Flag tag to facilitate its precipitation.[86] The computational portion of RIL-Seq (and CLASH) uses the ends of sequence fragments to map to the genome, as some of these reads are chimeras made of two RNA molecules that could be from different loci in the genome (or different genomes altogether, if looking at interspecies RNA-RNA interactions), and identifies statistically significant over-represented chimeras.[86]

'RNA antisense purification to systemically map RNA-RNA interactions' (*RAP-RNA*) uses biotinylated DNA probes after utilizing only crosslinking.[80] These probes are antisense to an RNA of interest. There are three versions of RAP-RNA that differ mainly in type of crosslinking

molecule: RAP-RNA (FA), (FA+DSG) or (AMT).[80] In RAP-RNA (AMT), RNA-RNA complexes are crosslinked with AMT, and RNA antisense to the biotinylated probes are sequenced.[80] In this technique, the crosslinks are not reversed, and reverse transcription is halted at the crosslinked RNA which enables mapping of RNA-RNA interaction sites, as opposed to just the identification of interaction partners.

To positively select crosslinked interaction pairs, 'sequencing of psoralen crosslinked, ligated, and selected hybrids' (SPLASH) uses biotinylated psoralen to crosslink interacting RNAs, enabling pull-down of the RNA complexes by streptavidin-coated beads before ligation.[79] Similarly, 'RNA in situ conformation' (RIC-Seq) labels interaction pairs with biotinylated cytidine (bis) phosphate (pCp) at the 3' ends of transcripts before ligation of both ends of the pair, which are then positively selected for using biotin-pCp.[87]

For clear delineation between and enrichment for ligated interaction pairs after sequencing, 'mapping RNA interactome in vivo' (MARIO) ligates unambiguous biotin-tagged RNA linkers between the RNA pairs.[88] First, all RNA-protein and protein-protein complexes are crosslinked via UV irradiation for capture of RNA-RNA interactions facilitated by any protein, and the proteins are denatured while the RNA is fragmented.[88] Biotinylation occurs at cysteine residues and is used to immobilize the RNA-protein complexes on streptavidin beads. Next, a biotin-tagged RNA linker is ligated to the 5' ends of RNAs, and then the interacting RNA pairs are ligated to form the chimeric product 'RNA1-linker-RNA2'.[88] Nguyen et al developed 'MARIO tools' (available at: https://mariotools.ucsd.edu), which is a bioinformatic pipeline to aid in the analysis of MARIO output and determine enriched interacting pairs from different genomic locations.[88]

Some techniques omit protein crosslinking and opt for only RNA-RNA crosslinking. Some techniques use the same general crosslinking and ligation procedure summarized above without pull-down of any kind and some minor additions. In *modified CLASH* by Liu *et al*, extracted RNAs are size-selected before ligation.[76] In 'ligation of interacting RNA followed by high throughput sequencing' (LIGR-Seq), RNase R digests uncrosslinked (likely non-interacting) RNA and thus enriches crosslinked duplexes after ligation.[77] Additionally, in 'psoralen analysis of RNA interactions and structures' (PARIS) proteins are totally digested, RNA is partially digested, and crosslinked RNA is purified through gel electrophoresis.[78] RNAnue, mentioned in the previous section, outperformed the original analyzation pipelines for SPLASH-generated, LIGR-Seq-generated, and PARIS-generated datasets.[53] This or a program like it would be ideal for analyzation of data from crosslinking/ligation methods that generate chimeras.

Various techniques use both photoactive crosslinking and subsequent ligation of RNA partners to determine direct RNA-RNA interactions without the use of anchors such as RNA-binding proteins or complementary oligonucleotide baits. Crosslinking RNAs to each other, instead of anchoring proteins, and ligation of those partners enables investigation into a global picture of RNA interaction, including protein-independent interaction.[76] In contrast to some methods mentioned above, most of these methods

do not require the use of genetically modified proteins or systems in order to investigate RNA-RNA interactions. Furthermore, these techniques do not require an RNA of interest, as they are able to map global or near-global RNA-RNA interactions. However, as the RNA ligation step is inefficient and ligation of non-interacting RNA occurs,[83] the number of chimeric reads is typically low. Schönberger *et al*[83] offer several improvements to these methods, including introducing non-complementary overhangs to improve ligation efficiency and a combinatorial approach to enriching crosslinked RNA partners.

These crosslinking and ligation techniques could be used to identify interspecies or interkingdom RNA-RNA interactions—in particular, bacterial sRNA-human mRNA interactions—if employed either after exposure of a singular sRNA of interest or a mixed population of sRNAs (such as that inside bacterial EVs), as these techniques identify RNA-RNA interactions globally. While the main advantage of these techniques is the global or near-global identification of RNA pairing partners, one disadvantage is the possibility of a high number of pairing partners that need to be evaluated. However, excluding interactions that do not involve interspecies pairings lowers that number considerably, especially since a small proportion of such interactions are expected.[29 33]

CONCLUSION

Bacterial sRNAs are frequently found within EVs.[9 11–13] These EVs can affect immune system-related gene expression,[32 33 40] and bacterial sRNAs associate with eukaryotic miRNA machinery.[29 33] While not a complete list of bioinformatic and experimental approaches that exist, this review summarizes a wide range of techniques that can be employed or modified for use to identify and validate the human targets of newly identified bacterial sRNAs (and more broadly, interspecies RNA-RNA interactions).

We discussed bioinformatic tools that allow annotation of sRNAs via aligning to known sRNAs (alignment/homology tools), and those that consider energy and accessibility to assess interaction favorability (summarized in table 1). Although the latter are more thorough and consider secondary structure, using a combination of bioinformatic techniques to identify sRNA targets is the most robust approach. Even when a particular target prediction tool is designed for a certain biological system, it may be applicable elsewhere—such as IntaRNA, which was originally intended for sRNA target prediction,[22] but has been used for eukaryotic target prediction as well.[50] Similarly, miRNA target prediction tools, such as TargetScan (IPA), DIANA microT, and miRanda, have been successfully used to predict human targets of bacterial sRNAs.[18 32]

Experimental techniques to identify RNA-RNA interactions that require the use of an anchor (eg, a biotinylated probe) and those that employ crosslinking and ligation of interacting RNA pairs were also discussed (summarized in table 2). EMSA, reporter systems, altering expression, and RNA pull-down assays are useful for validation of interactions, whereas crosslinking and ligation methods are more global and will be useful for exploration and identification of RNA-RNA interaction networks. While experimental methods to verify RNA-RNA interactions have often been

described for use in one cell type or species, they can be adapted in experimental design to allow identification and validation of interkingdom RNA-RNA interactions. For example, the aforementioned crosslinking and ligation methods could be employed after cell exposure to bacterial sRNAs or EVs.

To our knowledge, the vast majority of the bioinformatic and experimental methods discussed here have not been used to analyze interspecies interactions, particularly bacteria-human. Employing these methods, particularly the experimental methods, to determine bacterial-human RNA-RNA interactions will facilitate the important research that is needed to determine the potential impact microbiota exert on the host through these interspecies RNA-RNA interactions. It is only with improved understanding of these interactions that we can begin to effectively alter or adapt microbiota and positively impact human disease.

ORCID iD
Andrea L Hahn http://orcid.org/0000-0002-5117-0080

REFERENCES
1 Sultan S, El-Mowafy M, Elgaml A, *et al*. Metabolic influences of gut microbiota dysbiosis on inflammatory bowel disease. *Front Physiol* 2021;12:715506.
2 Hufnagl K, Pali-Schöll I, Roth-Walter F, *et al*. Dysbiosis of the gut and lung microbiome has a role in asthma. *Semin Immunopathol* 2020;42:75–93.
3 Saraf VS, Sheikh SA, Ahmad A, *et al*. Vaginal microbiome: normalcy vs dysbiosis. *Arch Microbiol* 2021;203:3793–802.
4 Macia L, Nanan R, Hosseini-Beheshti E, *et al*. Host- and Microbiota-Derived extracellular vesicles, immune function, and disease development. *Int J Mol Sci* 2019;21:201.
5 Ñahui Palomino RA, Vanpouille C, Costantini PE, *et al*. Microbiota-host communications: bacterial extracellular vesicles as a common language. *PLoS Pathog* 2021;17:e1009508.
6 Pérez-Cruz C, Delgado L, López-Iglesias C, *et al*. Outer-inner membrane vesicles naturally secreted by gram-negative pathogenic bacteria. *PLoS One* 2015;10:e0116896.
7 Bose S, Aggarwal S, Singh DV, *et al*. Extracellular vesicles: an emerging platform in Gram-positive bacteria. *Microb Cell* 2020;7:312–22.
8 Théry C, Witwer KW, Aikawa E, *et al*. Minimal information for studies of extracellular vesicles 2018 (MISEV2018): a position statement of the International Society for extracellular vesicles and update of the MISEV2014 guidelines. *J Extracell Vesicles* 2018;7:1535750.
9 Ghosal A, Upadhyaya BB, Fritz JV, *et al*. The extracellular RNA complement of Escherichia coli. *Microbiologyopen* 2015;4:252–66.
10 Chamberlain CA, Hatch M, Garrett TJ. Extracellular vesicle analysis by paper spray ionization mass spectrometry. *Metabolites* 2021;11:308.
11 Blenkiron C, Simonov D, Muthukaruppan A, *et al*. Uropathogenic Escherichia coli releases extracellular vesicles that are associated with RNA. *PLoS One* 2016;11:e0160440.

12 Joshi B, Singh B, Nadeem A, *et al*. Transcriptome Profiling of *Staphylococcus aureus* Associated Extracellular Vesicles Reveals Presence of Small RNA-Cargo. *Front Mol Biosci* 2020;7:566207.
13 Tsatsaronis JA, Franch-Arroyo S, Resch U, *et al*. Extracellular vesicle RNA: a universal mediator of microbial communication? *Trends Microbiol* 2018;26:401–10.
14 Beisel CL, Storz G. Base pairing small RNAs and their roles in global regulatory networks. *FEMS Microbiol Rev* 2010;34:866–82.
15 Nakanishi K. Anatomy of RISC: how do small RNAs and chaperones activate Argonaute proteins? *Wiley Interdiscip Rev RNA* 2016;7:637–60.
16 Fritsch TE, Siqueira FM, Schrank IS. Global analysis of sRNA target genes in Mycoplasma hyopneumoniae. *BMC Genomics* 2018;19:767.
17 Miravet-Verde S, Lloréns-Rico V, Serrano L. Alternative transcriptional regulation in genome-reduced bacteria. *Curr Opin Microbiol* 2017;39:89–95.
18 Koeppen K, Hampton TH, Jarek M, *et al*. A novel mechanism of host-pathogen interaction through sRNA in bacterial outer membrane vesicles. *PLoS Pathog* 2016;12:e1005672.
19 Carrier M-C, Lalaouna D, Massé E. Broadening the definition of bacterial small RNAs: characteristics and mechanisms of action. *Annu Rev Microbiol* 2018;72:141–61.
20 Obregón-Henao A, Duque-Correa MA, Rojas M, *et al*. Stable extracellular RNA fragments of Mycobacterium tuberculosis induce early apoptosis in human monocytes via a caspase-8 dependent mechanism. *PLoS One* 2012;7:e29970.
21 RKr G, Luong TT, Lee CY. RNAIII of the *Staphylococcus aureus agr* system activates global regulator MgrA by stabilizing mRNA. *Proceedings of the National Academy of Sciences* 2015;112:14036–41.
22 Busch A, Richter AS, Backofen R. IntaRNA: efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. *Bioinformatics* 2008;24:2849–56.
23 Mückstein U, Tafer H, Hackermüller J, *et al*. Thermodynamics of RNA-RNA binding. *Bioinformatics* 2006;22:1177–82.
24 Tafer H, Amman F, Eggenhofer F, *et al*. Fast accessibility-based prediction of RNA-RNA interactions. *Bioinformatics* 2011;27:1934–40.
25 Wenzel A, Akbasli E, Gorodkin J. RIsearch: fast RNA-RNA interaction search using a simplified nearest-neighbor energy model. *Bioinformatics* 2012;28:2738–46.
26 Fukunaga T, Hamada M. RIblast: an ultrafast RNA-RNA interaction prediction system based on a seed-and-extension approach. *Bioinformatics* 2017;33:2666–74.
27 Mann M, Wright PR, Backofen R. IntaRNA 2.0: enhanced and customizable prediction of RNA-RNA interactions. *Nucleic Acids Res* 2017;45:W435–9.
28 Andrade JM, Pobre V, Arraiano CM. Small RNA modules confer different stabilities and interact differently with multiple targets. *PLoS One* 2013;8:e52866.
29 Furuse Y, Finethy R, Saka HA, *et al*. Search for microRNAs expressed by intracellular bacterial pathogens in infected mammalian cells. *PLoS One* 2014;9:e106434.
30 Choi J-W, Kwon T-Y, Hong S-H, *et al*. Isolation and characterization of a microRNA-size secretable small RNA in Streptococcus sanguinis. *Cell Biochem Biophys* 2018;76:293–301.
31 Kang S-M, Choi J-W, Lee Y, *et al*. Identification of microRNA-size, small RNAs in Escherichia coli. *Curr Microbiol* 2013;67:609–13.
32 Choi J-W, Kim S-C, Hong S-H, *et al*. Secretable small RNAs via outer membrane vesicles in periodontal pathogens. *J Dent Res* 2017;96:458–66.
33 Han E-C, Choi S-Y, Lee Y, *et al*. Extracellular RNAs in periodontopathogenic outer membrane vesicles promote TNF-α production in human macrophages and cross the blood-brain barrier in mice. *Faseb J* 2019;33:13412–22.
34 Malabirade A, Habier J, Heintz-Buschart A, *et al*. The RNA Complement of Outer Membrane Vesicles From *Salmonella enterica* Serovar Typhimurium Under Distinct Culture Conditions. *Front Microbiol* 2018;9:2015.
35 Ormseth MJ, Wu Q, Zhao S, *et al*. Circulating microbial small RNAs are altered in patients with rheumatoid arthritis. *Ann Rheum Dis* 2020;79:1557–64.
36 Beatty M, Guduric-Fuchs J, Brown E, *et al*. Small RNAs from plants, bacteria and fungi within the order Hypocreales are ubiquitous in human plasma. *BMC Genomics* 2014;15:933.
37 Wang K, Li H, Yuan Y, *et al*. The complex exogenous RNA spectra in human plasma: an interface with human gut biota? *PLoS One* 2012;7:e51009.
38 Tulkens J, De Wever O, Hendrix A. Analyzing bacterial extracellular vesicles in human body fluids by orthogonal biophysical separation and biochemical characterization. *Nat Protoc* 2020;15:40–67.
39 Krohmaly K, Hoptay C, Hahn A. *Detection of bacterial extracellular vesicles in blood from healthy volunteers*. Poster presented at: ISEV2020 Annual Meeting, 2020.

40 Ha JY, Choi S-Y, Lee JH, *et al*. Delivery of periodontopathogenic extracellular vesicles to brain monocytes and microglial IL-6 promotion by RNA cargo. *Front Mol Biosci* 2020;7:596366.

41 Langmead B, Trapnell C, Pop M, *et al*. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 2009;10:R25.

42 Jorth P, Whiteley M. Characterization of a novel riboswitch-regulated lysine transporter in *Aggregatibacter actinomycetemcomitans*. *J Bacteriol* 2010;192:6240–50.

43 Kalvari I, Nawrocki EP, Ontiveros-Palacios N, *et al*. Rfam 14: expanded coverage of metagenomic, viral and microRNA families. *Nucleic Acids Res* 2021;49:D192–200.

44 Grüll MP, Peña-Castillo L, Mulligan ME, *et al*. Genome-wide identification and characterization of small RNAs in *Rhodobacter capsulatus* and identification of small RNAs affected by loss of the response regulator CtrA. *RNA Biol* 2017;14:914–25.

45 Ragan C, Zuker M, Ragan MA. Quantitative prediction of miRNA-mRNA interaction based on equilibrium concentrations. *PLoS Comput Biol* 2011;7:e1001090.

46 Rehmsmeier M, Steffen P, Höchsmann M, *et al*. Fast and effective prediction of microRNA/target duplexes. *RNA* 2004;10:1507–17.

47 Alkan F, Wenzel A, Palasca O, *et al*. RIsearch2: suffix array-based large-scale prediction of RNA-RNA interactions and siRNA off-targets. *Nucleic Acids Res* 2017;45:e60.

48 Umu SU, Gardner PP. A comprehensive benchmark of RNA–RNA interaction prediction tools for all domains of life. *Bioinformatics* 2016;7:btw728–96.

49 Mai J, Rao C, Watt J, *et al*. Mycobacterium tuberculosis 6C sRNA binds multiple mRNA targets via C-rich loops independent of RNA chaperones. *Nucleic Acids Res* 2019;47:4292–307.

50 Zhao P, Li T, Wang Y, *et al*. Lncrna MYCNOS promotes glioblastoma cell proliferation by regulating miR-216b/FOXM1 axis. *Metab Brain Dis* 2021;36:1185–9.

51 Cheng J, Lin Y, Xu L, *et al*. ViRBase v3.0: a virus and host ncRNA-associated interaction Repository with increased coverage and annotation. *Nucleic Acids Res* 2022;50:D928–33.

52 Aklan F, Wenzel A, Palasca O, *et al*. RIsearch2: user manual, 2018. Available: https://rth.dk/resources/risearch/download/Manual_RIsearch2.pdf [Accessed 3 May 2022].

53 Schäfer RA, Voß B. RNAnue: efficient data analysis for RNA-RNA interactomics. *Nucleic Acids Res* 2021;49:5493–501.

54 Westermann AJ, Gorski SA, Vogel J. Dual RNA-seq of pathogen and host. *Nat Rev Microbiol* 2012;10:618–30.

55 Westermann AJ, Förstner KU, Amman F, *et al*. Dual RNA-seq unveils noncoding RNA functions in host-pathogen interactions. *Nature* 2016;529:496–501.

56 Westermann AJ, Barquist L, Vogel J. Resolving host-pathogen interactions by dual RNA-seq. *PLoS Pathog* 2017;13:e1006033.

57 QIAGEN Ingenuity Pathway Analysis (QIAGEN IPA). Qiagen.com. Available: https://digitalinsights.qiagen.com/products-overview/discovery-insights-portfolio/analysis-and-visualization/qiagen-ipa/ [Accessed 4 May 2022].

58 Chung M, Bruno VM, Rasko DA, *et al*. Best practices on the differential expression analysis of multi-species RNA-seq. *Genome Biol* 2021;22:121.

59 Qiagen.com. Microrna target filter. Available: https://digitalinsights.qiagen.com/products-overview/discovery-insights-portfolio/analysis-and-visualization/qiagen-ipa/features/microrna-target-filter/ [Accessed 3 May 2022].

60 Mika-Gospodorz B, Giengkam S, Westermann AJ, *et al*. Dual RNA-seq of Orientia tsutsugamushi informs on host-pathogen interactions for this neglected intracellular human pathogen. *Nat Commun* 2020;11:3363.

61 Schulze S, Schleicher J, Guthke R, *et al*. How to predict molecular interactions between species? *Front Microbiol* 2016;7:442.

62 Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* 2005;120:15–20.

63 TargetScan. Available: https://www.targetscan.org/vert_80/ [Accessed 4 May 2022].

64 Enright AJ, John B, Gaul U, *et al*. MicroRNA targets in Drosophila. *Genome Biol* 2003;5:R1.

65 Maragkakis M, Alexiou P, Papadopoulos GL, *et al*. Accurate microRNA target prediction correlates with protein repression levels. *BMC Bioinformatics* 2009;10:295.

66 Paraskevopoulou MD, Georgakilas G, Kostoulas N, *et al*. DIANA-microT web server v5.0: service integration into miRNA functional analysis workflows. *Nucleic Acids Res* 2013;41:W169–73.

67 Henderson CA, Vincent HA, Stone CM, *et al*. Characterization of MICA interactions suggests a potential novel means of gene regulation by small non-coding RNAs. *Nucleic Acids Res* 2013;41:3386–97.

68 Silva IJ, Ortega AD, Viegas SC, *et al*. An rpoS-dependent sRNA regulates the expression of a chaperone involved in protein folding. *RNA* 2013;19:1253–65.

69 Ota C, Morisaki H, Nakata M, *et al*. Streptococcus sanguinis Noncoding *cia*-Dependent Small RNAs Negatively Regulate Expression of Type IV Pilus Retraction ATPase PilT and Biofilm Formation. *Infect Immun* 2018;86:e00894–17.

70 Zhang T, Zhao Y-L, Zhao J-H, *et al*. Cotton plants export microRNAs to inhibit virulence gene expression in a fungal pathogen. *Nat Plants* 2016;2:16153.

71 Jiao J, Peng D. Wheat microRNA1023 suppresses invasion of *Fusarium graminearum* via targeting and silencing *FGSG_03101*. *Journal of Plant Interactions* 2018;13:514–21.

72 Yang M, Chen W, Liu H, *et al*. Long non-coding RNA CBR3 antisense RNA 1 is downregulated in colorectal cancer and inhibits miR-29a-Mediated cell migration and invasion. *Mol Biotechnol* 2022;64:773–9.

73 Carrier M-C, Lalaouna D, Massé E. A game of tag: MAPS catches up on RNA interactomes. *RNA Biol* 2016;13:473–6.

74 Mercier N, Prévost K, Massé E, *et al*. MS2-Affinity purification coupled with RNA sequencing in Gram-positive bacteria. *J Vis Exp* 2021. doi:10.3791/61731. [Epub ahead of print: 23 Feb 2021].

75 Lalaouna D, Desgranges E, Caldelari I. MS2-Affinity purification coupled with RNA sequencing approach in the human pathogen Staphylococcus aureus 2018:393–411.

76 Liu T, Zhang K, Xu S, *et al*. Detecting RNA-RNA interactions in E. coli using a modified clash method. *BMC Genomics* 2017;18:343.

77 Sharma E, Sterne-Weiler T, O'Hanlon D, *et al*. Global mapping of human RNA-RNA interactions. *Mol Cell* 2016;62:618–26.

78 Lu Z, Zhang QC, Lee B, *et al*. RNA duplex map in living cells reveals higher-order transcriptome structure. *Cell* 2016;165:1267–79.

79 Aw JGA, Shen Y, Wilm A, *et al*. In Vivo Mapping of Eukaryotic RNA Interactomes Reveals Principles of Higher-Order Organization and Regulation. *Mol Cell* 2016;62:603–17.

80 Engreitz JM, Sirokman K, McDonel P, *et al*. RNA-RNA interactions enable specific targeting of noncoding RNAs to nascent pre-mRNAs and chromatin sites. *Cell* 2014;159:188–99.

81 Zhen WP, Buchardt O, Nielsen H, *et al*. Site specificity of psoralen-DNA interstrand cross-linking determined by nuclease Bal31 digestion. *Biochemistry* 1986;25:6598–603.

82 Cimino GD, Gamper HB, Isaacs ST, *et al*. Psoralens as photoactive probes of nucleic acid structure and function: organic chemistry, photochemistry, and biochemistry. *Annu Rev Biochem* 1985;54:1151–93.

83 Schönberger B, Schaal C, Schäfer R, *et al*. Rna interactomics: recent advances and remaining challenges. *F1000Res* 2018;7:1824.

84 Gay LA, Turner PC, Renne R. Modified cross-linking, ligation, and sequencing of hybrids (qCLASH) to identify microRNA targets. *Curr Protoc* 2021;1:e257.

85 Iosub IA, van Nues RW, McKellar SW, *et al*. Hfq clash uncovers sRNA-target interaction networks linked to nutrient availability adaptation. *Elife* 2020;9:e54655.

86 Melamed S, Faigenbaum-Romm R, Peer A, *et al*. Mapping the small RNA interactome in bacteria using RIL-seq. *Nat Protoc* 2018;13:1–33.

87 Cai Z, Cao C, Ji L, *et al*. RIC-seq for global in situ profiling of RNA-RNA spatial interactions. *Nature* 2020;582:432–7.

88 Nguyen TC, Cao X, Yu P, *et al*. Mapping RNA-RNA interactome and RNA structure in vivo by Mario. *Nat Commun* 2016;7:12023.

89 Johnson M, Zaretskaya I, Raytselis Y, *et al*. NCBI BLAST: a better web interface. *Nucleic Acids Res* 2008;36:W5–9.

90 Ermolaeva MD, Khalak HG, White O, *et al*. Prediction of transcription terminators in bacterial genomes. *J Mol Biol* 2000;301:27–33.

91 Gautheret D, Lambert A. Direct RNA motif definition and identification from multiple sequence alignments using secondary structure profiles. *J Mol Biol* 2001;313:1003–11.

92 Gardner PP, Barquist L, Bateman A, *et al*. RNIE: genome-wide prediction of bacterial intrinsic terminators. *Nucleic Acids Res* 2011;39:5845–52.

93 McClure R, Balasubramanian D, Sun Y, *et al*. Computational analysis of bacterial RNA-Seq data. *Nucleic Acids Res* 2013;41:e140.

94 Leonard S, Meyer S, Lacour S, *et al*. APERO: a genome-wide approach for identifying bacterial small RNAs from RNA-Seq data. *Nucleic Acids Res* 2019;47:e88.

95 Toffano-Nioche C, Luo Y, Kuchly C, *et al*. Detection of non-coding RNA in bacteria and archaea using the DETR'PROK Galaxy pipeline. *Methods* 2013;63:60–5.

96 Lott SC, Schäfer RA, Mann M, *et al*. GLASSgo - Automated and Reliable Detection of sRNA Homologs From a Single Input Sequence. *Front Genet* 2018;9:124.