Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 July 2021

Type of the Paper (Review)

Small RNAs Asserting Big Roles in Mycobacteria

Fatma S. Coskun[1], Przemysław Płociński[2] and Nicolai S. C. van Oers[1,3,*]

[1] Departments of Immunology, University of Texas Southwestern Medical Center, Dallas, TX, United States; fatma.coskun@utsouthwestern.edu

[2] Department of Immunology and Infectious Biology, Faculty of Biology and Environmental Protection, University of Łódz, Banacha 12/16, 90-237 Łódź, Poland; przemyslaw.plocinski@biol.uni.lodz.pl

[3] Departments of Microbiology, Pediatrics, University of Texas Southwestern Medical Center, Dallas, TX, United States; nicolai.vanoers@utsouthwestern.edu

* Correspondence: nicolai.vanoers@utsouthwestern.edu

Abstract: Tuberculosis (TB) is an infectious disease caused by *Mycobacterium tuberculosis* (Mtb), with 10.4 million new cases per year reported in the human population. Recent studies on the Mtb transcriptome have revealed the abundance of noncoding RNAs expressed at various phases of mycobacteria growth; in culture, in infected mammalian cells and in patients. Among these noncoding RNAs are both small RNAs (sRNAs) between 50-350 nts in length and smaller RNAs (sncRNA) <50 nts. In this review, we provide an up-to-date synopsis of the identification, designation, and function of these Mtb-encoded sRNAs and sncRNAs. The methodological advances including RNA sequencing strategies, small RNA antagonists and locked nucleic acid sequence specific RNA probes advancing the studies on these small RNA are described. Initial insights into the regulation of their expression and putative processing enzymes required for their synthesis and function are discussed. There are many open questions remaining about the biological and pathogenic roles of these small non-coding RNAs, and potential research directions needed to define the role of these mycobacterial noncoding RNAs summarized.

Keywords: Mycobacteria, small RNAs, sncRNAs, RNA processing.

1. Introduction

Mycobacterium tuberculosis (Mtb) remains one of the leading infectious causes of human mortality, supplanted only in 2020 by the COVID-19 pandemic triggered by the SARS-CoV-2 virus. Mtb evolved from an ancestral smooth tubercule bacillus (e.g. *M. canettii*, *M. pseudotuberculosis*), acquiring virulence elements to attain its preferred pathogenicity towards humans[1]. The acquisition of these virulence elements coincided with Mtb undergoing a genomic down-sizing compared to the 100 different smooth tubercule bacilli species characterized [1, 2]. In spite of this downsizing, a core genome is evident among the diverse pathogenic strains of mycobacteria. Several decades of research efforts have been devoted to understanding how the ~4000 protein-coding elements evident in the Mtb genome contribute to growth, survival and pathogenic processes [3-7]. Recent technical advances in deciphering the complex nature of Mtb and related mycobacterial

© 2021 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 July 2021

genomes, including improved large-scale RNA-sequencing strategies, have revealed that ~20% of the Mtb transcriptome comprises small non-coding RNA species (sRNAs) [8]. First described as ranging in size from 50-350 nucleotides (nts), these small RNAs now include some as small as 18 nts [8-13]. These assorted small RNAs are detected in both exponential and/or stationary phase cultures, in infected eukaryotic cells and in patients with tuberculosis (TB), suggesting key roles in all aspects of mycobacterial growth and survival [14]. In the current review, recent discoveries pertaining to these sRNAs are described. Emerging reports detailing their diverse functions along with their transcriptional regulation and processing are discussed. Future directions of research and therapeutic strategies to manipulate such sRNAs is also presented.

2. Identification and Designation of Mycobacterial Small RNAs

The initial screens to identify Mtb-encoded sRNAs were prompted by prior reports on the existence of such small noncoding RNAs in different bacterial species, *Escherichia coli*, *Salmonella enterica,* and *Staphylococcus aureus* [15-17]. As with most pioneering studies, assorted definitions and naming strategies were applied to the Mtb small RNAs [18]. Those sRNAs transcribed from the untranslated region of a coding gene or in the antisense orientation to this sequence are called antisense RNAs and alternatively termed cis-encoded RNAs. Such RNAs often hybridize to the coding RNA with perfect base-pair complementarity, effectively reducing RNA translation of the gene. A second group of sRNAs are called intergenic sRNAs, which are transcribed from intergenic regions. These are sometimes termed trans-encoded RNAs and use imperfect base-pairing with a 6-7 nucleotide seed-sequence targeting complementary sequences on gene targets.

The first sets of Mtb-encoded small RNAs were identified by enriching for low molecular weight RNAs from exponential and stationary phase Mtb cultures, followed by cDNA synthesis and sequence analysis [10]. Four antisense (ASdes, Aspks, AS1726, and AS1890) and five intergenic sRNAs (B11, B55, C8, F6, and G2), ranging in size from 38-238 nts, were identified (Table 1). Later work revealed that ASdes, ASpks, AS1726 and AS1890 were also present in Bacillus Calmette Guerin (BCG) mycobacterial cultures [19]. ASdes was subsequently detected in the plasma of 15 of 27 TB patients and 6 of 24 BCG vaccinated individuals [19].

RNA sequencing approaches have broadened the number of small RNAs detected in Mtb, with 19 and 20 antisense and intergenic sRNAs identified in Mtb cultures, respectively [8]. These sRNAs represent 28% of the exponential phase Mtb transcriptome after excluding ribosomal RNAs [8]. Three of the intergenic small RNAs were analyzed in more detail. Termed Mycobacterium tuberculosis sRNA (MTS), MTS1338 predominates in stationary phase cultures while MTS0997 and MTS2823 are detected in stationary cultures and increase 12-fold and 1.6-fold during exponential phase growth [8]. A more recent screen identified 27 additional antisense sRNAs in dormant TB along with 2 antisense and 8 intergenic sRNAs previously identified [20].

A screen for small RNAs in *Mycobacterium bovis* identified 37 sRNAs, 19 of which were confirmed by Northern blotting (Table 1) [11]. These were termed Mcr1-McR19, for Mycobacterium Cloned RNAs. Not surprisingly as *M.* bovis originated from Mtb, Mcr6, Mcr11, and Mcr14 were the same as the aforementioned C8, MTS0997, and F6 sRNAs, respectively. With an increasing number of RNA sequencing screens performed by various labs, Mcr3, Mcr4, Mcr8, Mcr9, Mcr11, and Mcr14 are consistently detected in *M. bovis* while B11, B55, C8, F6, G2, and ASdes are evident in Mtb [9]. Given that many sRNAs, described in other bacterial species, were formed in more extreme growth conditions, distinct clusters of Mtb-encoded sRNA were consequently also found to be induced following either oxidative stress, pH stress, membrane

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 July 2021

stress, or nutrient or iron starvation conditions [12]. Such studies have expanded the number of sRNAs to >189, with a new nomenclature used to differentiate them [21]. From these 189 sRNAs, most had sizes ranging from 51-300 nts.

Most of the RNA screens with Mtb focused on RNAs >50 nts, yet many smaller RNAs <50 nts have been characterized in non-mycobacterial species [22-25]. For example, *Salmonella* expresses a small RNA called Sal-1, which is generated from the 5′ end of a ribosomal RNA by the eukaryotic miRNA processing enzymes [22]. Sal-1 targets the inducible nitric oxide synthase, with the pathogenic role for this sRNA established by the increased killing of Sal-1 deficient *Salmonella* lacking in infected epithelial cells [22]. This resembles eukaryotic miRNAs, which are small noncoding RNAs (20-22 nts) that use 6-7 nucleotide seed sequences to mediate the degradation of mRNAs [26, 27]. The first screen for such miRNA-like sRNAs was undertaken with *Mycobacterium marinum* [28]. In this screen, a single 23 nt RNA was discovered, with features characteristic of a eukaryotic miRNA including the requisite interaction with the Argonaute protein, part of the eukaryotic RNA-induced silencing complex [28]. The *M. marinum* sRNA has no ascribed functions or targets to date. In a more recent screen for miRNA-like sRNAs from TB-infected patients, 6 distinct Mtb encoded miRNA-like sRNAs were discovered in the serum [29]. While all of these had 22 nucleotide lengths consistent with the size of miRNAs, their extremely high GC content was unusual (86%-100%), making it unlikely these are generated by eukaryotic miRNA processing enzymes. In a broader screen for miRNA-like sequences using additional selection criteria that included coverage of plant small RNAs, a set of 35 small RNAs were identified in Mtb-infected macrophages [13]. These sRNAs increased over the 6-day infection period. Termed smaller noncoding RNAs (sncRNAs), their sizes ranged from18-30 nts. All 35 sncRNAs had an average GC content of 50%. In a technical advance to determine the levels of these sncRNAs, a miRNA-based quantitative RT-PCR was developed. This assay incorporates locked-nucleic acid technologies to provide extremely high specificity and selectively for short RNAs. The expression changes of three of the Mtb-encoded sncRNAs, sncRNA-1, sncRNA-6, and sncRNA-8, was verified with this technique [13].

To summarize, diverse mycobacterial species produce small RNA transcripts ranging in size from 18-350 nts. The secondary structure of representative examples of such sRNAs is shown in Figure 1. The mycobacterial sRNAs have diverse sizes and extensive predicted 2° structures, yet lack commonality. Only a handful of these sRNAs have been functionally characterized. Some are more abundant in infected cell lines and in patients, implying roles in pathogenesis. Scientists are beginning to explore their genomic targets, production and processing requirements, and contributions to pathogenesis. We describe next the current state of knowledge of these sRNAs.

3. Functional Roles of Mycobacterial sRNAs and sncRNAs

A key clue to putative biological roles for the sRNAs relates to their expression patterns during various phases of mycobacterial growth, as reviewed elsewhere [30]. Additional insights are suggested by the conditions of growth and survival needed for their induction. These include oxidative stress, nutrient deprivation, DNA damage, and acidic environments such as the phagolysosome in macrophages and dendritic cells. Putative functional roles for the numerous sRNAs need also to consider the stability of the sRNA, affected by both the relative GC content and secondary RNA structures. Examples of several better characterized sRNAs are B11/6C, MTS1338/DrrS, MTS2823/Ms1, MTS0097/Mcr11, ncRv11846/MrsI and sncRNA-1 (Figure 2, Table 1).

sRNA B11 (93 nts), later named 6C due to its similarity to a small RNA found in other bacterial species, forms two stem-loops via 6 conserved cytosines [31, 32]. Target sequence searches have suggested that B11/6C regulates Mtb transcripts coupled to DNA replication and protein secretion. In mechanistic studies in *M. smegmatis*, 6C was found to interact with

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 July 2021

two mRNA targets, *panD* and *dnaB* (Figure 2A, Table 1). Overexpression of 6C inhibited *M. smegmatis* growth. Taking advantage of this, a screen for RNA chaperons that mediate the interaction between 6C and its targets was developed. In experiments where *M. smegmatis* overexpressing 6C were exposed to saturation mutagenesis, few colonies that were recovered were those which had mutations in the overexpression cassette. Otherwise, no colonies in the saturation mutagenesis library was recovered, which indicates that the interaction between B11/6C and its targets does not require an essential chaperone protein [32]. In gram-negative bacteria sRNA/mRNA interactions require the Hfq protein, or a recently identified ProQ chaperone. Yet, Mtb lacks a conserved Hfq or ProQ homolog. Instead, the interaction between the sRNA and mRNA targets in Mtb were proposed to involve direct Watson-Crick base pairing with the GC content of the sRNA and the target [33]. While this may apply to certain sRNAs, it is likely that most Mtb sRNAs target the mRNAs or DNA through as yet uncharacterized accessory proteins.

Several groups have used mycobacterial RNA over-expression vectors to further understand how the various sRNAs function. The over-expression of MTS1338 (117 nts) prevents Mtb replication, suggesting it targets key genes needed for mycobacterial growth (Figure 2B, Table 1) [8]. MTS1338 was later named as DosR Regulated sRNA (DrrS) as it was found to be induced by DosR. Elevated levels of MTS2823 (300 nts) also inhibit Mtb growth, with RNA sequencing data revealing many transcripts involved in metabolism are downregulated (Figure 2C, Table 1) [8]. MTS2823 is also termed Ms1 as it was functionally characterized in *M. smegmatis* and has homology to the 6S sRNA [34]. In *E. coli*, 6S sRNA has a secondary RNA structure that resembles an open promoter. The sigma factor bound RNA polymerase (RNAP) holoenzyme has a high affinity for this RNA structure [35]. The 6S sRNA complexes the RNAP, competitively reducing transcriptional activity [35]. Yet, studies in *M. smegmatis* suggest that MTS2823/Ms1 competes with the sigma factor for binding to RNAP, hence suppressing transcriptional activity. Given the complexity of defining RNA-protein complexes, a revised model is proposed that MTS2823/Ms1 sequesters the RNAP [36].

MTS0997 (131 nts), later named Mcr11, regulates Mtb fatty acid metabolism (Figure 2D, Table 1) [37]. It positively regulates Rv3282, *fadA3*, and *lipB* translation by binding a 7-11 nucleotide region upstream of the start codon. Supplementing fatty acids during cultures deactivates this mechanism, revealing a feedback loop to control metabolic functions in Mtb. The regulatory role for sRNAs in Mtb metabolism was also revealed with the sRNA, ncRv11846 (106 nts). An ortholog of the *E. Coli* sRNA RhyB, NcRV11846 is also termed Mycobacterial regulatory sRNA in Iron (MrsI) (Figure 2E, Table 1) [12]. NcRv11846/MrsI is expressed following iron starvation [16]. This sRNA uses a 6-nucleotide seed sequence to negatively regulate the genes *hypF* and *bfrA*, which encode for nonessential iron-containing proteins. This translational roadblock increases the levels of free iron available. There are multiple additional sRNAs identified that are reduced in expression in response to iron starvation [12]. The roles of these sRNAs remains an open question.

Among the diverse sncRNAs, sncRNA-1 remains the best characterized (Figure 2F) [13]. This non-coding RNA is present in the RD1 pathogenicity locus, in-between *esxA* and *espI*. Over-expressing sncRNA-1 alters the Mtb transcriptome, with multiple genes required for fatty acid biogenesis increased in expression. Screening putative targets of sncRNA-1 by seed-seed complementarity searches suggested two targets of this sRNA, Rv0242c and Rv1094. These are two key genes involved in the oleic acid biogenesis pathway. Both genes had putative sncRNA-1 binding sites within their 5′ UTRs. By substituting selected nucleotides involved in Watson-Crick base pairing, either within the 5′ UTR or in the sncRNA seed sequence eliminated the positive regulation. To identify putative processing requirements needed for the generation of sncRNA-1, nucleotide substitutions were created within the hairpin loop and antisense complementarity strand of the precursor form of this snRNA. This revealed that the processing of the longer RNA transcript was

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 July 2021

dependent on both the formation of the hairpin loop and the nucleotides at a putative cleavage site needed to form sncRNA-1. Notably, the expression of the longer RNA sequence that still contained sncRNA-1 no longer regulated gene expression.

One novel approach to study microRNA functions is the use of locked nucleic acid power inhibitors (LNA-PIs). These are not cleaved by RNA processing enzymes and have been modified to transduce into cells without liposomes [38]. They hybridize with the target miRNA conferring them unfunctional. These LNA-PIs were tested in mycobacteria, which are inherently difficult to transduce. Notably, such LNAs are easily incorporated into mycobacteria and can antagonize sncRNAs in Mtb [13, 39]. Incubation of Mtb with an LNA-PI selectively targeting sncRNA-1 abolished the upregulation of the Rv0242c [13]. This reduced the survival of Mtb in macrophages, revealing a key pathogenic contribution of this sncRNA. While the functions of sncRNA-6 and sncRNA-8 remain unexplored, sncRNA-6 also undergoes a sequence specific processing from a longer RNA transcript [13]. Like sncRNA-1, mutations that disrupt the hairpin loop in which sncRNA-6 resides or the mutations at the cleavage site of sncRNA-6 prevents its processing. All told, the findings reveal a critical role for smaller noncoding RNAs in Mtb pathogenicity.

4. Regulation of Mycobacterial sRNAs/sncRNAs Expression

As more sRNAs/sncRNAs are being discovered in diverse mycobacterial species, the regulatory elements controlling their expression and processing are slowly being characterized. This includes the identification of different trans- and cis-regulatory factors. In mycobacteria, sigA is the primary transcription factor, which is a member of the sigma70 family [40]. SigA recognizes the consensus cis regulatory sequence, the TTGCGA–N_{18}–TANNNT hexamer that is present at -35 and -10 region upstream of the transcription start site (Figure 3A) [41, 42]. SigA binding enables RNA Polymerase to transcribe at promoter sites responsible for the expression of housekeeping regulons and for mycobacterial growth [40, 43]. Miotto *et. al.* developed computational predictions to identify sigA-regulated sRNAs [44]. Of the sRNAs identified in the screen, 46.9% had the consensus SigA promoter sequence in the 5′ UTR, with 8.5% containing an intrinsic or factor dependent terminator sequence in the 3′ UTR. While 13.6% of the sRNAs had both 5′ and 3′ elements, the presence of these and their impact on transcription requires further study. The remaining 31.0% of the sRNA had neither site, suggesting the involvement of other regulatory factors. For example, MTS2823/Ms1 contains a -10 element, starting 5-nucleotide upstream of +1 position along with a distinct -35 element, suggesting that a distinct sigma factor regulates its expression. MTS2823/Ms1 contains different regulatory elements (−491/+9 region) that contribute its expression.

Coupled with the cis-regulatory elements are novel trans-regulatory elements being identified that control sRNAs/sncRNA expression. Among these are alternate transcription factors or sigma factors. For instance, sRNA ncRv11846/MrsI has an IdeR binding site in its promoter region. IdeR is an iron-responsive master regulator of genes coupled to iron metabolism, including the sRNA ncRv11846/MrsI [45] (Figure 3C) [12]. sRNA, MTS0997/Mcr11, resides between two protein-coding genes Rv1264 and Rv1265, with the protein products of these 2 genes involved in the metabolism of cAMP [46]. Rv1264 encodes an adenylyl cyclase, which catalyzes ATP to cAMP. Rv1265 is a transcription factor that binds to both ATP and DNA. DNA binding studies have shown that Rv1265 induces MTS0997/Mcr11 expression (Figure 3D). Rv1265 is now termed AbmR for ATP binding Mcr11 regulator. Mapping studies of the 5′ end of MTS0997/Mcr11 revealed that its -35 element coincides with the transcriptional start site of AmbR, which is oriented in the opposite direction [8]. This suggests that MTS0997/Mcr11 and AbmR share a transcriptional start site that is positively regulated by AbmR itself (Figure 3D). This strategy of shared start sites is also reported for the sRNA MTS1338/DrrS, transcribed in the opposite direction as Rv1733c. Rv1733c encodes a protein involved in cell wall

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 July 2021

biogenesis, and is a component of the DosR regulon [8]. The DosR regulon, induced by nitric oxide (NO), is the primary mediator of the hypoxic stress response [47]. MTS1338/DrrS is also upregulated in response to NO, and the MTS1338/DrrS promoter is activated by DosR, established with b-galactosidase reporter assays (Figure 3E) [48]. Overall, this might present a mechanism for a shared transcription start between a sRNA and the neighboring gene oriented in the opposite direction.

In summary, identification of the cis- and trans-acting factors is revealing many diverse types or regulatory elements involved in the sRNA expression. Little is known about the regulation of the sncRNAs.

5. Processing of Mycobacterial sRNAs and sncRNAs

Many sRNAs are generated as full-length mature transcripts with no obvious processing steps. Yet, several of the smaller species do undergo some form of processing [36, 37, 48]. Included among these are MTS2823/Ms1, MTS1338/DrrS, sncRNA-1 and sncRNA-6. MTS2823/Ms1 is a 300 nt transcript detected in both exponential and stationary phase cultures. Notably, it also exists as a 250 nt transcript in stationary phase, suggesting some form of processing [8]. MTS1338/DrrS is transcribed as a precursor transcript of >400 nts (referred to as DrrS+) that is cleaved at the 3′ end to yield the mature 108 nts form [48]. MTS0997/Mcr11 has a 3′ end that varies in size by 3-14 nts, implying a 3′ RNA processing occurs like that for MTS133/DrrS [37]. Both sncRNA-1 and sncRNA-6, which have final sizes of 25 nts and 21 nts, require processing enzymes for their generation (Figure 3) [13]. These sncRNAs were predicted to exists as precursor transcripts >115 nts that have defined hairpin loops and 2° RNA structures. Nucleotide substitutions within the putative cleave sites caused in the formation of multiple intermediate size-transcripts (40-115 nts), detected by Northern blotting [13]. Taken together, multiple experiments establish the existence of a small RNA processing system in mycobacteria. Several candidate RNA processing enzymes have been reported to date. Among these are ribonuclease E (RNase E), polynucleotide phosphorylase (PNPase or GpsI), ribonuclease J (RNase J) and the ATP-dependent RNA helicase RhlE (Figure 4) [49]. All are components of the RNA degradasome. Except RhlE, all are essential for *in vitro* growth, determined by identifying key genes through a transposon mutagenesis screen (Himar1 transposon libraries) [7]. Mechanistically, RNase E recognizes the 5′ phosphate of the transcript and then cuts at an A/U rich sequence of the ssRNA [50]. PNPase and RNase J are 3′ and 5′ specific exonucleases, respectively that stop upon the presence of a dsRNA sequence [51]. Many research teams made use of CRISPR interference mediated knock down of the RNA processing enzymes to study their role in the generation of specific sRNAs [36, 49, 52]. Sikova *et. al.* have investigated the contribution of the core RNase enzymes in the processing of MTS2823/Ms1 [36]. Knockdown of PNPase increased the levels of MTS2823/Ms1 ~30%, while the targeting of RNase E and RNase J had no effect on this sRNA, revealing some target specificity. These findings further suggest that the processing of MTS2823/Ms1 likely involves additional RNA processing enzymes. Another possibility is that residual protein levels of PNPase were still resulting in some processing of the longer RNA transcript. Taken together, the limited number of studies on the RNA processing enzymes leave open many questions about how Mtb produces sRNAs from longer transcripts.

6. tRNA processing enzymes as potential players for sRNA maturation

Transfer RNAs (tRNAs) share some common features with small RNAs, being relatively short and highly structured non-coding RNA molecules. tRNA maturation involves several steps, with both 3′ and 5′ ends being extensively processed in an orchestrated, sequential order. Besides the core RNA degradosome components, tRNA processing enzymes are also involved in maturation and turnover of certain sRNA species. In many organisms, transcripts encoding tRNAs are also a source of regulatory small RNAs namely tRNA-derived small RNAs [53]. The mechanisms of tRNA maturation

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 July 2021

in Mtb are not well characterized and require future studies. Compared to *E. coli* and *B. subtilis* used as model bacteria for RNA processing, Mtb encodes for RNase P, which is involved in the initial processing of the 5′ end of tRNA molecules [54]. The suite of 3′ end processing enzymes includes RNase PH, RNase Z, the oligoribonuclease [55], RNase D (Rv2681) and a divergent functional and structural ortholog of RNase T (Rv2179c) [56]. Since a large proportion of mycobacterial tRNAs require an enzymatic addition of the CCA sequence at their 3′ end, they are likely additionally processed by the Poly(A) polymerase and/or PNPase. All the ribonucleases described above have the potential to be involved in processing of small non-coding RNAs other than tRNA.

tRNA cleavage is seen in all kingdoms of life as a regulatory mechanism, adding another layer to the complexity of gene regulation mechanisms. This has been observed in *Streptomyces coelicolor*, another actinomycetes related to Mycobacteria [57]. tRNA cleavage has recently been reported in Mtb [58]. Tuberculosis encodes numerous toxin-antitoxin systems, with many requiring a ribonuclease component. The VapC11 ribonuclease of the virulence associated TA system, VapBC, specifically cleaves two tRNA species, $tRNA^{Gln32\text{-}CUG}$ and $tRNA^{Leu3\text{-}CAG}$ [58]. Mtb encodes for about 50 VapC ribonuclease toxins, with these having the potential to directly target noncoding RNAs. It is also likely that these would cleave tRNAs to yield tRNA derived functional sRNAs in Mtb, as they seem to have in higher eukaryotes [59]. Direct or indirect interplay between VapC toxins and sRNA is inevitable. In fact, overexpression of the MTS2823 restricts expression of at least five VapC homologues [8], but the exact mechanism remains unexplored.

7. The hunt for the mycobacterial Hfq equivalent

In the majority of bacteria species, trans-encoded sRNAs require RNA chaperones, either Hfq or ProQ, to ensure appropriate sRNA:mRNA base pairing [60]. CsrA, a conserved small RNA binding protein was recently shown to assist complex formation between sRNA and its mRNA targets in *Bacillus subtilis* [61]. The RNA chaperones closely cooperate and interact with the core RNA degradosome to ensure efficient regulation of gene expression [62]. As previously mentioned, the Mtb lacks any identifiable orthologues of Hfq, ProQ or CsrA chaperones and must exploit alternative proteins or mechanisms for efficient sRNA-mediated mRNA decay. Recent studies from *E. coli* and *S. aureus* have revealed that Cold shock domain containing proteins (CSPs), involved in binding and melting RNA species [63, 64], also interact with several sRNAs. CSP proteins are typically respond to stress and could aid in a coordinated response to external stimuli that is not limited to cold sensing [65]. Hence, these proteins may likely be involved in sRNA mediated regulation of gene expression in Mtb. To corroborate this notion, CspA and CspB both associate with the core RNA degradosome in Mtb [49]. Future studies will likely reveal their relevance to the functionalities of the RNA degrading machinery. Interestingly, the mycobacterial CspA gene itself is co-expressed with the sRNA molecule, ncRv3648c. Exploiting active RNA structure unwinding, with the help of ATP-dependent RNA helicases, could theoretically support sRNA folding in the absence of passive unwinding mechanisms provided by Hfq-like chaperones. A previous study from *E. coli* have reported the requirement of the CsdA DEAD-box helicase for low temperature riboregulation of rpoS mRNA via sRNA-mediated mechanism, where the activity of Hfq was not sufficient for translational activation of rpoS expression [66].

Intriguingly, *M. smegmatis*, *M. dioxanotrophicus* and *M. goodie* encode a eukaryotic-like protein with a full length TROVE domain (KEGG database search www.genome.jp), sharing over 35% identity with the human 60 kDa SS-A/Ro ribonucleoprotein ortholog (SIM analysis results [67]). The 60 kDa SS-A/Ro ribonucleoprotein binds to misfolded small RNAs and pre-5S rRNA in eukaryotes [68]. It is thought to function as an RNA chaperone that stabilizes small RNAs of the Y family and protects them from enzymatic degradation [69]. In mycobacteria, the protein coding element was likely acquired by horizontal gene transfer from a mycobacteriophage, with a similar gene identified in the mycobacteriophage

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 July 2021

Sparky (KEGG database ortholog search, www.genome.jp). It remains unknown whether the TROVE protein has acquired some functions related to sRNA metabolism in fast growing mycobacterial species or it is simply a useless remnant of a previous bacteriophage infection.

8. Concluding Remarks

Mtb-encoded small RNAs are emerging as new regulators of mycobacterial growth, survival, and pathogenesis. To date, the functions for only a handful of these sRNAs have been described. Given their distinct sizes, their function contributions likely differ when comparing those <50 nt and those between 50-450 nt. The recently developed CRISPR interference-based assays hold great potential for studying the function of sRNAs and sncRNAs [12, 34, 52]. In addition, LNA power inhibitors recently validated on sncRNA-1 are a very tractable system for blocking sRNA functions [13]. Another remaining question is the type and sequence specificity of the regulatory elements controlling sRNA transcription. The lack of a comprehensive study to define such regulatory elements has limited the identification of additional sRNAs. Lastly, many studies suggest that the assorted sRNAs are processed after transcription, and this adds another complication due to the lack of elaborate techniques to define exact processing events. The processing may be growth phase or stress-dependent, altering the function of the RNA in selected physiological conditions. Moreover, the mycobacterial proteome involved in RNA editing is still poorly annotated. Using Mycobrowser, candidate genes involved in RNA processing in Mtb are evident (Table 2). However, more studies are needed to investigate the role of these putative RNA binding/processing proteins since most have been identified via computational predictions. As summarized, the last decade has identified many distinct Mtb-encoded RNAs. The next decade will likely address the open questions mentioned throughout this review. Identification of new sRNAs/sncRNAs involved in pathogenesis and their regulatory mechanisms will enhance our understanding of tools that Mtb utilizes to escape macrophage killing, which will eventually help eradicate the TB.

9. Figure Legends

Figure 1: Predicted secondary structure of B11/6C, MTS1338/DrrS, ncRv11846/MrsI, MTS2823/Ms1, MTS0997/Mcr11, sncRNA-1, sncRNA-6 and sncRNA-8. The verified seed sequences are highlighted in yellow. The structures were obtained using the RNAFold web server and may be different than previous reports.

Figure 2: The function of B11/6C (A), MTS1338/DrrS (B), MTS2823/Ms1 (C), MTS0997/Mcr11 (D), ncRv11846/MrsI (E), and sncRNA-1. Indirect interactions are shown with a dashed line. (A) B11/6C gets induced by unknown factors and positively regulates the expression genes coupled to growth. (B) DrrS+ is induced by the DosR regulon upon nitric oxide stress. Then, it undergoes post transcriptional processing to yield MTS1338/DrrS. MTS1338/DrrS promotes the expression of three operons which cause defects in Mtb growth and promotes persistence. The mechanism of this MTS1338/DrrS mediated regulation has not been characterized. (C) MTS2823/Ms1 sequesters RNA Polymerase (RNAP) at the stationary phase. Upon entrance to the outgrowth phase, MTS2823/Ms1 is degraded by PNPase and some other RNases not yet identified which releases RNAP to promote global transcription. (D) The expression of MTS0997/Mcr11 is regulated by AbmR, an ATP-bound transcription factor. After transcription, MTS0997/Mcr11 undergoes processing at the 3' end and then regulate the expression of genes involved in fatty acid production in a site-specific manner. This is negatively regulated by fatty acids. (E) In iron restricted environments, the iron responsive transcription factor IdeR induces the expression of ncRv11846/MrsI which in return hinders the translation of nonessential iron storing proteins. This increases the level of free iron that can be used for essential functions. (F) sncRNA-1 is induced in infected macrophages and gets processed to yield 25 nts RNA. The processed sncRNA-1 enhances the expression of Rv0242c

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 July 2021

and Rv1094, two genes involved in oleic acid production. This regulatory network then promotes Mtb growth and survival inside macrophages.

Figure 3: Cis and trans regulatory elements involved in sRNA expression is shown. (A) sigA recognizes a consensus sequence to induce the expression of a set of sRNAs identified by Miotto *et. al.* (B) -35 and -10 elements upstream of MTS2823/Ms1 is shown. Distal regulatory elements not shown here also contribute to the expression of MTS2823/Ms1. (C) IdeR potentially regulates the expression of ncRv11846/MrsI through the IdeR-box found in its promoter region. (D) The expression of MTS0997/Mcr11 is regulated by AmbR located in the upstream of MTS0997/Mcr11 and expressed in the opposite orientation. (E) The expression of MTS1338/DrrS is regulated by the DosR transcription factor.

Figure 4: The genes coupled to the Mtb degradosome are shown. RNase J is an endoribonuclease and a 5′-3′ exoribonuclease. RNase E is an endoribonuclease that cleaves ssRNA at A/U rich sites after recognizing the 5′ phosphate in proximity. PNPase is a 3′-5′ exoribonuclease, which is the only RNase implicated in the processing of a sRNA, MTS2823/Ms1.

Author Contributions: Conceptualization, F.S.C. and N.S.C.O, writing—review and editing, F.S.C. P.P. and N.S.C.O. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the UT Southwestern Medical Center Beecherl funding to N.S.C.v.O. P.P. was supported by National Science Centre, Poland 2015/19/D/NZ1/02842, SONATA 10.

Institutional Review Board Statement: "Not applicable."

Informed Consent Statement: Not applicable."

Acknowledgments: We appreciate the help of Dr. Didem Agac in critically reviewing the review article

Conflicts of Interest: The authors declare no conflict of interest.

10. References

1. Supply, P.; Marceau, M.; Mangenot, S.; Roche, D.; Rouanet, C.; Khanna, V.; Majlessi, L.; Criscuolo, A.; Tap, J.; Pawlik, A.; Fiette, L.; Orgeur, M.; Fabre, M.; Parmentier, C.; Frigui, W.; Simeone, R.; Boritsch, E. C.; Debrie, A.-S.; Willery, E.; Walker, D.; Quail, M. A.; Ma, L.; Bouchier, C.; Salvignol, G.; Sayes, F.; Cascioferro, A.; Seemann, T.; Barbe, V.; Locht, C.; Gutierrez, M.-C.; Leclerc, C.; Bentley, S. D.; Stinear, T. P.; Brisse, S.; Medigue, C.; Parkhill, J.; Cruveiller, S.; Brosch, R., Genomic analysis of smooth tubercle bacilli provides insights into ancestry and pathoadaptation of Mycobacterium tuberculosis. *Nat Genet* **2013,** *45* (2), 172-179.

2. Cole, S. T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S. V.; Eiglmeier, K.; Gas, S.; Barry, C. E., 3rd; Tekaia, F.; Badcock, K.; Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.; Murphy, L.; Oliver, K.; Osborne, J.; Quail, M. A.; Rajandream, M. A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, J.; Squares, R.; Squares, S.; Sulston, J. E.; Taylor, K.; Whitehead, S.; Barrell, B. G., Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature* **1998,** *393* (6685), 537-44.

3. Ohol, Y. M.; Goetz, D. H.; Chan, K.; Shiloh, M. U.; Craik, C. S.; Cox, J. S., Mycobacterium tuberculosis MycP1 protease plays a dual role in regulation of ESX-1 secretion and virulence. *Cell Host Microbe* **2010,** *7* (3), 210-20.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 July 2021

4. Garces, A.; Atmakuri, K.; Chase, M. R.; Woodworth, J. S.; Krastins, B.; Rothchild, A. C.; Ramsdell, T. L.; Lopez, M. F.; Behar, S. M.; Sarracino, D. A.; Fortune, S. M., EspA acts as a critical mediator of ESX1-dependent virulence in Mycobacterium tuberculosis by affecting bacterial cell wall integrity. *PLoS Pathog* **2010,** *6* (6), e1000957.

5. Singh, R.; Singh, A.; Tyagi, A. K., Deciphering the genes involved in pathogenesis of Mycobacterium tuberculosis. *Tuberculosis (Edinb)* **2005,** *85* (5-6), 325-35.

6. Sassetti, C. M.; Rubin, E. J., Genetic requirements for mycobacterial survival during infection. *Proc Natl Acad Sci U S A* **2003,** *100* (22), 12989-94.

7. DeJesus, M. A.; Gerrick, E. R.; Xu, W.; Park, S. W.; Long, J. E.; Boutte, C. C.; Rubin, E. J.; Schnappinger, D.; Ehrt, S.; Fortune, S. M.; Sassetti, C. M.; Ioerger, T. R., Comprehensive Essentiality Analysis of the Mycobacterium tuberculosis Genome via Saturating Transposon Mutagenesis. *mBio* **2017,** *8* (1).

8. Arnvig, K. B.; Comas, I.; Thomson, N. R.; Houghton, J.; Boshoff, H. I.; Croucher, N. J.; Rose, G.; Perkins, T. T.; Parkhill, J.; Dougan, G.; Young, D. B., Sequence-based analysis uncovers an abundance of non-coding RNA in the total transcriptome of Mycobacterium tuberculosis. *PLoS Pathog* **2011,** *7* (11), e1002342.

9. Pellin, D.; Miotto, P.; Ambrosi, A.; Cirillo, D. M.; Di Serio, C., A genome-wide identification analysis of small regulatory RNAs in Mycobacterium tuberculosis by RNA-Seq and conservation analysis. *PLoS One* **2012,** *7* (3), e32723.

10. Arnvig, K. B.; Young, D. B., Identification of small RNAs in Mycobacterium tuberculosis. *Mol Microbiol* **2009,** *73* (3), 397-408.

11. DiChiara, J. M.; Contreras-Martinez, L. M.; Livny, J.; Smith, D.; McDonough, K. A.; Belfort, M., Multiple small RNAs identified in Mycobacterium bovis BCG are also expressed in Mycobacterium tuberculosis and Mycobacterium smegmatis. *Nucleic Acids Res* **2010,** *38* (12), 4067-78.

12. Gerrick, E. R.; Barbier, T.; Chase, M. R.; Xu, R.; Francois, J.; Lin, V. H.; Szucs, M. J.; Rock, J. M.; Ahmad, R.; Tjaden, B.; Livny, J.; Fortune, S. M., Small RNA profiling in Mycobacterium tuberculosis identifies MrsI as necessary for an anticipatory iron sparing response. *Proc Natl Acad Sci U S A* **2018,** *115* (25), 6464-6469.

13. Coskun, F. S.; Srivastava, S.; Raj, P.; Dozmorov, I.; Belkaya, S.; Mehra, S.; Golden, N. A.; Bucsan, A. N.; Chapagain, M. L.; Wakeland, E. K.; Kaushal, D.; Gumbo, T.; van Oers, N. S. C., sncRNA-1 Is a Small Noncoding RNA Produced by Mycobacterium tuberculosis in Infected Cells That Positively Regulates Genes Coupled to Oleic Acid Biosynthesis. *Front Microbiol* **2020,** *11*, 1631.

14. Carrier, M. C.; Lalaouna, D.; Masse, E., Broadening the Definition of Bacterial Small RNAs: Characteristics and Mechanisms of Action. *Annu Rev Microbiol* **2018,** *72*, 141-161.

15. Wassarman, K. M.; Zhang, A.; Storz, G., Small RNAs in Escherichia coli. *Trends Microbiol* **1999,** *7* (1), 37-45.

16. Gottesman, S., The small RNA regulators of Escherichia coli: roles and mechanisms*. *Annu Rev Microbiol* **2004,** *58*, 303-28.

17. Waters, L. S.; Storz, G., Regulatory RNAs in bacteria. *Cell* **2009,** *136* (4), 615-28.

18. Gottesman, S.; Storz, G., Bacterial small RNA regulators: versatile roles and rapidly evolving variations. *Cold Spring Harb Perspect Biol* **2011,** *3* (12).

19. Fu, Y.; Li, W.; Wu, Z.; Tao, Y.; Wang, X.; Wei, J.; Jiang, P.; Wu, J.; Zhang, Z.; Zhang, W.; Zhao, J.; Zhang, F., Detection of mycobacterial small RNA in the bacterial culture supernatant and plasma of patients with active tuberculosis. *Biochem Biophys Res Commun* **2018,** *503* (2), 490-494.

20. Ignatov, D. V.; Salina, E. G.; Fursov, M. V.; Skvortsov, T. A.; Azhikina, T. L.; Kaprelyants, A. S., Dormant non-culturable Mycobacterium tuberculosis retains stable low-abundant mRNA. *BMC Genomics* **2015,** *16*, 954.

21. Lamichhane, G.; Arnvig, K. B.; McDonough, K. A., Definition and annotation of (myco)bacterial non-coding RNA. *Tuberculosis (Edinb)* **2013,** *93* (1), 26-9.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 July 2021

22. Gu, H.; Zhao, C.; Zhang, T.; Liang, H.; Wang, X. M.; Pan, Y.; Chen, X.; Zhao, Q.; Li, D.; Liu, F.; Zhang, C. Y.; Zen, K., Salmonella produce microRNA-like RNA fragment Sal-1 in the infected cells to facilitate intracellular survival. *Sci Rep* **2017,** *7* (1), 2392.

23. Kang, S. M.; Choi, J. W.; Lee, Y.; Hong, S. H.; Lee, H. J., Identification of microRNA-size, small RNAs in Escherichia coli. *Curr Microbiol* **2013,** *67* (5), 609-13.

24. Dang, T. H. Y.; Tyagi, S.; D'Cunha, G.; Bhave, M.; Crawford, R.; Ivanova, E. P., Computational prediction of microRNAs in marine bacteria of the genus Thalassospira. *PLoS One* **2019,** *14* (3), e0212996.

25. Shmaryahu, A.; Carrasco, M.; Valenzuela, P. D., Prediction of bacterial microRNAs and possible targets in human cell transcriptome. *J Microbiol* **2014,** *52* (6), 482-9.

26. Bartel, D. P., MicroRNAs: target recognition and regulatory functions. *Cell* **2009,** *136* (2), 215-33.

27. Guo, H.; Ingolia, N. T.; Weissman, J. S.; Bartel, D. P., Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature* **2010,** *466* (7308), 835-40.

28. Furuse, Y.; Finethy, R.; Saka, H. A.; Xet-Mull, A. M.; Sisk, D. M.; Smith, K. L.; Lee, S.; Coers, J.; Valdivia, R. H.; Tobin, D. M.; Cullen, B. R., Search for microRNAs expressed by intracellular bacterial pathogens in infected mammalian cells. *PLoS One* **2014,** *9* (9), e106434.

29. Chakrabarty, S.; Kumar, A.; Raviprasad, K.; Mallya, S.; Satyamoorthy, K.; Chawla, K., Host and MTB genome encoded miRNA markers for diagnosis of tuberculosis. *Tuberculosis (Edinb)* **2019,** *116*, 37-43.

30. Ostrik, A. A.; Azhikina, T. L.; Salina, E. G., Small Noncoding RNAs and Their Role in the Pathogenesis of Mycobacterium tuberculosis Infection. *Biochemistry (Mosc)* **2021,** *86* (Suppl 1), S109-S119.

31. Weinberg, Z.; Barrick, J. E.; Yao, Z.; Roth, A.; Kim, J. N.; Gore, J.; Wang, J. X.; Lee, E. R.; Block, K. F.; Sudarsan, N.; Neph, S.; Tompa, M.; Ruzzo, W. L.; Breaker, R. R., Identification of 22 candidate structured RNAs in bacteria using the CMfinder comparative genomics pipeline. *Nucleic Acids Res* **2007,** *35* (14), 4809-19.

32. Mai, J.; Rao, C.; Watt, J.; Sun, X.; Lin, C.; Zhang, L.; Liu, J., Mycobacterium tuberculosis 6C sRNA binds multiple mRNA targets via C-rich loops independent of RNA chaperones. *Nucleic Acids Res* **2019,** *47* (8), 4292-4307.

33. Bloch, S.; Wegrzyn, A.; Wegrzyn, G.; Nejman-Falenczyk, B., Small and Smaller-sRNAs and MicroRNAs in the Regulation of Toxin Gene Expression in Prokaryotic Cells: A Mini-Review. *Toxins (Basel)* **2017,** *9* (6).

34. Hnilicova, J.; Jirat Matejckova, J.; Sikova, M.; Pospisil, J.; Halada, P.; Panek, J.; Krasny, L., Ms1, a novel sRNA interacting with the RNA polymerase core in mycobacteria. *Nucleic Acids Res* **2014,** *42* (18), 11763-76.

35. Cavanagh, A. T.; Wassarman, K. M., 6S RNA, a global regulator of transcription in Escherichia coli, Bacillus subtilis, and beyond. *Annu Rev Microbiol* **2014,** *68*, 45-60.

36. Sikova, M.; Janouskova, M.; Ramaniuk, O.; Palenikova, P.; Pospisil, J.; Bartl, P.; Suder, A.; Pajer, P.; Kubickova, P.; Pavlis, O.; Hradilova, M.; Vitovska, D.; Sanderova, H.; Prevorovsky, M.; Hnilicova, J.; Krasny, L., Ms1 RNA increases the amount of RNA polymerase in Mycobacterium smegmatis. *Mol Microbiol* **2019,** *111* (2), 354-372.

37. Girardin, R. C.; McDonough, K. A., Small RNA Mcr11 requires the transcription factor AbmR for stable expression and regulates genes involved in the central metabolism of Mycobacterium tuberculosis. *Mol Microbiol* **2020,** *113* (2), 504-520.

38. Grunweller, A.; Hartmann, R. K., Locked nucleic acid oligonucleotides: the next generation of antisense agents? *BioDrugs* **2007,** *21* (4), 235-43.

39. Manier, S.; Powers, J. T.; Sacco, A.; Glavey, S. V.; Huynh, D.; Reagan, M. R.; Salem, K. Z.; Moschetta, M.; Shi, J.; Mishima, Y.; Roche-Lestienne, C.; Leleu, X.; Roccaro, A. M.; Daley, G. Q.; Ghobrial, I. M., The LIN28B/let-7 axis is a novel therapeutic pathway in multiple myeloma. *Leukemia* **2017,** *31* (4), 853-860.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 July 2021

40. Gomez, M.; Doukhan, L.; Nair, G.; Smith, I., sigA is an essential gene in Mycobacterium smegmatis. *Mol Microbiol* **1998,** *29* (2), 617-28.

41. Hu, Y.; Coates, A. R., Transcription of two sigma 70 homologue genes, sigA and sigB, in stationary-phase Mycobacterium tuberculosis. *J Bacteriol* **1999,** *181* (2), 469-76.

42. Sachdeva, P.; Misra, R.; Tyagi, A. K.; Singh, Y., The sigma factors of Mycobacterium tuberculosis: regulation of the regulators. *FEBS J* **2010,** *277* (3), 605-26.

43. Wu, S.; Howard, S. T.; Lakey, D. L.; Kipnis, A.; Samten, B.; Safi, H.; Gruppo, V.; Wizel, B.; Shams, H.; Basaraba, R. J.; Orme, I. M.; Barnes, P. F., The principal sigma factor sigA mediates enhanced growth of Mycobacterium tuberculosis in vivo. *Mol Microbiol* **2004,** *51* (6), 1551-62.

44. Miotto, P.; Forti, F.; Ambrosi, A.; Pellin, D.; Veiga, D. F.; Balazsi, G.; Gennaro, M. L.; Di Serio, C.; Ghisotti, D.; Cirillo, D. M., Genome-wide discovery of small RNAs in Mycobacterium tuberculosis. *PLoS One* **2012,** *7* (12), e51950.

45. Schmitt, M. P.; Predich, M.; Doukhan, L.; Smith, I.; Holmes, R. K., Characterization of an iron-dependent regulatory protein (IdeR) of Mycobacterium tuberculosis as a functional homolog of the diphtheria toxin repressor (DtxR) from Corynebacterium diphtheriae. *Infect Immun* **1995,** *63* (11), 4284-9.

46. Girardin, R. C.; Bai, G.; He, J.; Sui, H.; McDonough, K. A., AbmR (Rv1265) is a novel transcription factor of Mycobacterium tuberculosis that regulates host cell association and expression of the non-coding small RNA Mcr11. *Mol Microbiol* **2018,** *110* (5), 811-830.

47. Park, H. D.; Guinn, K. M.; Harrell, M. I.; Liao, R.; Voskuil, M. I.; Tompa, M.; Schoolnik, G. K.; Sherman, D. R., Rv3133c/dosR is a transcription factor that mediates the hypoxic response of Mycobacterium tuberculosis. *Mol Microbiol* **2003,** *48* (3), 833-43.

48. Moores, A.; Riesco, A. B.; Schwenk, S.; Arnvig, K. B., Expression, maturation and turnover of DrrS, an unusually stable, DosR regulated small RNA in Mycobacterium tuberculosis. *PLoS One* **2017,** *12* (3), e0174079.

49. Plocinski, P.; Macios, M.; Houghton, J.; Niemiec, E.; Plocinska, R.; Brzostek, A.; Slomka, M.; Dziadek, J.; Young, D.; Dziembowski, A., Proteomic and transcriptomic experiments reveal an essential role of RNA degradosome complexes in shaping the transcriptome of Mycobacterium tuberculosis. *Nucleic Acids Res* **2019,** *47* (11), 5892-5905.

50. Baek, Y. M.; Jang, K. J.; Lee, H.; Yoon, S.; Baek, A.; Lee, K.; Kim, D. E., The bacterial endoribonuclease RNase E can cleave RNA in the absence of the RNA chaperone Hfq. *J Biol Chem* **2019,** *294* (44), 16465-16478.

51. Jester, B. C.; Romby, P.; Lioliou, E., When ribonucleases come into play in pathogens: a survey of gram-positive bacteria. *Int J Microbiol* **2012,** *2012*, 592196.

52. Rock, J. M.; Hopkins, F. F.; Chavez, A.; Diallo, M.; Chase, M. R.; Gerrick, E. R.; Pritchard, J. R.; Church, G. M.; Rubin, E. J.; Sassetti, C. M.; Schnappinger, D.; Fortune, S. M., Programmable transcriptional repression in mycobacteria using an orthogonal CRISPR interference platform. *Nat Microbiol* **2017,** *2*, 16274.

53. Oberbauer, V.; Schaefer, M. R., tRNA-Derived Small RNAs: Biogenesis, Modification, Function and Potential Impact on Human Disease Development. *Genes (Basel)* **2018,** *9* (12).

54. Singh, A.; Batra, J. K., Insight into the functional role of unique determinants in RNA component of RNase P of Mycobacterium tuberculosis. *Int J Biol Macromol* **2018,** *119*, 937-944.

55. Taverniti, V.; Forti, F.; Ghisotti, D.; Putzer, H., Mycobacterium smegmatis RNase J is a 5'-3' exo-/endoribonuclease and both RNase J and RNase E are involved in ribosomal RNA maturation. *Mol Microbiol* **2011,** *82* (5), 1260-76.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 July 2021

56. Abendroth, J.; Ollodart, A.; Andrews, E. S.; Myler, P. J.; Staker, B. L.; Edwards, T. E.; Arcus, V. L.; Grundner, C., Mycobacterium tuberculosis Rv2179c protein establishes a new exoribonuclease family with broad phylogenetic distribution. *J Biol Chem* **2014,** *289* (4), 2139-47.
57. Haiser, H. J.; Karginov, F. V.; Hannon, G. J.; Elliot, M. A., Developmentally regulated cleavage of tRNAs in the bacterium Streptomyces coelicolor. *Nucleic Acids Res* **2008,** *36* (3), 732-41.
58. Cintron, M.; Zeng, J. M.; Barth, V. C.; Cruz, J. W.; Husson, R. N.; Woychik, N. A., Accurate target identification for Mycobacterium tuberculosis endoribonuclease toxins requires expression in their native host. *Sci Rep* **2019,** *9* (1), 5949.
59. Stanger, S. J.; Bernstein, I. R.; Anderson, A. L.; Hutcheon, K.; Dun, M. D.; Eamens, A. L.; Nixon, B., The abundance of a transfer RNA-derived RNA fragment small RNA subpopulation is enriched in cauda spermatozoa. *ExRNA* **2020,** *2* (1).
60. Quendera, A. P.; Seixas, A. F.; Dos Santos, R. F.; Santos, I.; Silva, J. P. N.; Arraiano, C. M.; Andrade, J. M., RNA-Binding Proteins Driving the Regulatory Activity of Small Non-coding RNAs in Bacteria. *Front Mol Biosci* **2020,** *7*, 78.
61. Muller, P.; Gimpel, M.; Wildenhain, T.; Brantl, S., A new role for CsrA: promotion of complex formation between an sRNA and its mRNA target in Bacillus subtilis. *RNA Biol* **2019,** *16* (7), 972-987.
62. Bruce, H. A.; Du, D.; Matak-Vinkovic, D.; Bandyra, K. J.; Broadhurst, R. W.; Martin, E.; Sobott, F.; Shkumatov, A. V.; Luisi, B. F., Analysis of the natively unstructured RNA/protein-recognition core in the Escherichia coli RNA degradosome and its interactions with regulatory RNA/Hfq complexes. *Nucleic Acids Res* **2018,** *46* (1), 387-402.
63. Caballero, C. J.; Menendez-Gil, P.; Catalan-Moreno, A.; Vergara-Irigaray, M.; Garcia, B.; Segura, V.; Irurzun, N.; Villanueva, M.; Ruiz de Los Mozos, I.; Solano, C.; Lasa, I.; Toledo-Arana, A., The regulon of the RNA chaperone CspA and its auto-regulation in Staphylococcus aureus. *Nucleic Acids Res* **2018,** *46* (3), 1345-1361.
64. Rennella, E.; Sara, T.; Juen, M.; Wunderlich, C.; Imbert, L.; Solyom, Z.; Favier, A.; Ayala, I.; Weinhaupl, K.; Schanda, P.; Konrat, R.; Kreutz, C.; Brutscher, B., RNA binding and chaperone activity of the E. coli cold-shock protein CspA. *Nucleic Acids Res* **2017,** *45* (7), 4255-4268.
65. Mazzon, R. R.; Lang, E. A.; Silva, C. A.; Marques, M. V., Cold shock genes cspA and cspB from Caulobacter crescentus are posttranscriptionally regulated and important for cold adaptation. *J Bacteriol* **2012,** *194* (23), 6507-17.
66. Resch, A.; Vecerek, B.; Palavra, K.; Blasi, U., Requirement of the CsdA DEAD-box helicase for low temperature riboregulation of rpoS mRNA. *RNA Biol* **2010,** *7* (6), 796-802.
67. Huang, X.; Miller, W., A Time-Efficient, Linear-Space Local Similarity Algorithm. *Advances in Applied Mathematics* **1991,** *12*, 337-357.
68. Hogg, J. R.; Collins, K., Human Y5 RNA specializes a Ro ribonucleoprotein for 5S ribosomal RNA quality control. *Genes Dev* **2007,** *21* (23), 3067-72.
69. Chen, X.; Smith, J. D.; Shi, H.; Yang, D. D.; Flavell, R. A.; Wolin, S. L., The Ro autoantigen binds misfolded U2 small nuclear RNAs and assists mammalian cell survival after UV irradiation. *Curr Biol* **2003,** *13* (24), 2206-11.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 July 2021

Table 1: List of sRNA identified by Arnvig *et. al.*, DiChiara *et. al.*, Gerrick *et. al.* and Coskun *et. al.*

Name	Northern or PCR size	Location	Surrounding genes	Expression	Study
B11/6C (candidate_1603)	93	4099386-4099478 (-)	Rv3660c-Rv3661	H_2O_2 and pH=5	Arnvig et. al., 2009 & Pellin et. al., 2012
B55 (candidate_84)	61	704187-704247(+)	Rv0609A-Rv0610c	H_2O_2 and MMC	Arnvig et. al., 2009 & Pellin et. al., 2012
C8 (Mcr6, candidate_1621)	58,70,128	4168154-4168281(-)	Rv3722c-Rv3723	TBD	Arnvig et. al., 2009 & DiChiara et.al. 2010 & Pellin et. al., 2012
F6 (Mcr14, candidate_29)	38,58,102	293604-293705(+)	fadA2-fadE5	H_2O_2 and pH=5	Arnvig et. al., 2009 & Hartkoorn et. al., 2012 & Pellin et. al., 2012
G2 (candidate_1269)	67, 214, 229	1914962-1915190(-)	tyrS-lprJ	TBD	Arnvig et. al., 2009 & Pellin et. al., 2012
ASdes (candidate_121)	48, 63, 68, 83, 94, 109, 149, 169, 195	918264-918458(+)	within desA1	TBD	Arnvig et. al., 2009 & Fu et. al., 2018 & Pellin et. al., 2012
ASpks	78, 89, 91, 102, 129, 142, 162	2299745-2299886(+)	within pks12	H_20_2	Arnvig et. al., 2009
AS1726	61, 77, 85, 110, 213	1952291-1952503(-)	within Rv1726	TBD	Arnvig et. al., 2009
AS1890	63, 109, 191, 238	2139419-2139656(+)	within Rv1890	TBD	Arnvig et. al., 2009
MTS2823/Ms1	250,300	4100669-4100968(+)	Rv3661-Rv3662c	in vivo	Arnvig et. al.,2011 & Hnilicova et. al.,2014 & Sikova et. al., 2019
MTS1338/DrrS	100	1960667-1960783(+)	Rv1733c-Rv1735c	NO, stationary phase, in vivo	Arnvig et. al.,2011 & Moores et. al.,2017 & Salina et. al. 2019
MTS0997/Mcr11 (candidate_1693)	115	1413094-1413224(-)	Rv1264-Rv1265	in vivo, stationary phase, low pH or hypoxia	Arnvig et. al. 2011, DiChiara et. al., 2010 & Pellin et. al., 2012 & Girardin et. al.,2019 & Girardin et. al.,2018
Mcr1	>300	2029043-2029087(TBD)	PPE26-PPE27	TBD	DiChiara et. al., 2010
Mcr2	120	1108857-1108824(TBD)	Rv0967-Rv0968	TBD	DiChiara et. al., 2010
Mcr3 (candidate_190)	118	1471619-1471742(+)	murA-rrs	TBD	DiChiara et. al., 2010 & Pellin et. al., 2012
Mcr4 (candidate_1314)	200-250	2137148-2137103(TBD)	fbpB-Rv1887	TBD	DiChiara et. al., 2010 & Pellin et. al., 2012
Mcr5	80	2437823-2437866(-)	within Rv2175c	TBD	DiChiara et. al., 2010
Mcr7	350-400	2692172-2692521(+)	Rv2395-PE_PGRS41	TBD	DiChiara et. al., 2010
Mcr8 (candidate_1935)	200	4073966-4073908(TBD)	Rv3661-3662c	TBD	DiChiara et. al., 2010 & Pellin et. al., 2012
Mcr9 (candidate_1502)	66-82	3317634-3317517(TBD)	ilvB1-cfp6	TBD	DiChiara et. al., 2010 & Pellin et. al., 2012

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 July 2021

Mcr10	120	1283693-1283815(+)	within Rv1157c	TBD	DiChiara et. al., 2010
Mcr12	118	1228436-1228381(TBD)	Rv1072-Rv1073	TBD	DiChiara et. al., 2010
Mcr13	311	4315154-4315215(TBD)	Rv3866-Rv3867	TBD	DiChiara et. al., 2010
Mcr15	>300	1535417-1535716(-)	Rv1363c-Rv1364c	TBD	DiChiara et. al., 2010
Mcr16	100	2517032-2517134(-)	within fabD	TBD	DiChiara et. al., 2010
Mcr17	82-90	2905457-2905402(TBD)	within Rv2613c	TBD	DiChiara et. al., 2010
Mcr18	82	3466287-3466332(TBD)	within nuoC	TBD	DiChiara et. al., 2010
Mcr19	66-82	575033-575069(+)	within Rv0485	TBD	DiChiara et. al., 2010
ncRv11846/MrsI	100	2096766-2096867(+)	blaI-Rv1847	iron starvation, oxidative stress, and membrane stress	Gerrick et. al.,2018
sncRNA-1	25	4352927- 4352951	esxA-Rv3876	inside macrophages	Coskun et. al., 2020
sncRNA-6	21	786003- 786083	Rv0685-Rv0686	inside macrophages	Coskun et. al., 2020
sncRNA-8	24	1471701- 1471724	murA-rrs	inside macrophages	Coskun et. al., 2020

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 July 2021

Table 2: List of Mtb-encoded proteins that have putative functions in sRNA processing

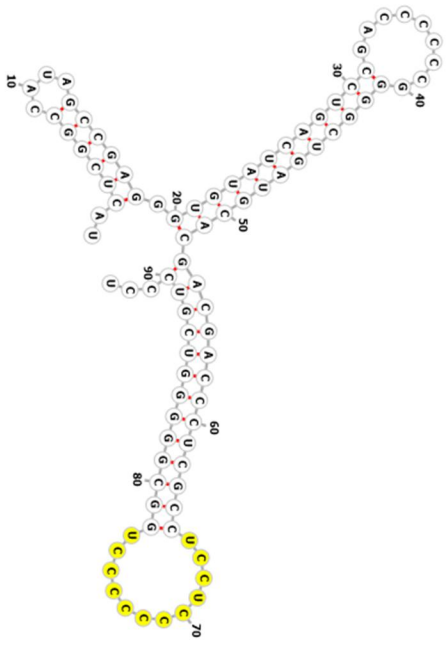
Gene ID	Name	Species	Putative Function
MTB000026	rnpB	*M. bovis, Mtb, M. haemo*	RNA component of RNase P: RNase P catalyzes the removal of the 5'-leader sequence from pre-tRNA to produce the mature 5'terminus.
Rv1340	rphA	*M. marinum, M. leprae, M. bovis, Mtb*	Probable ribonuclease RphA (RNase PH)
Rv2092c	helY	*M. marinum, M. leprae, M. bovis, Mtb, M. abscessus*	DNA helicase activity
Rv2179c	rnt	*Mtb, M. smegmatis, M. leprae, M. marinum*	Conserved hypothetical protein
Rv2228c	Rv2228c	*N/A*	Multifunctional protein. Has RNASE H, alpha-ribazole phosphatase, and acid phosphatase activities.
Rv2407	rnz	*Mtb, M. smegmatis, M. leprae, M. marinum, M. bovis*	Endonucleolytic cleavage of RNA, removing extra 3' nucleotides from tRNA precursor, generating 3' termini of tRNAs.
Rv2444c	rne	*M. bovis, Mtb, M. leprae, M. marinum, M. smegmatis*	Putative RNase E. Plays central role in the maturation of 5S and 16S rRNAs and the majority of tRNAs. Also involved in the degradation of most mRNAs.
Rv2511	orn	*Mtb, M. smegmatis, M. leprae, M. marinum, M. bovis*	Involved in RNA degradation: 3'-to-5' exoribonuclease specific for small oligoribonucleotides.
Rv2681	rnd	*Mtb, M. smegmatis, M. leprae, M. marinum, M. bovis*	Conserved hypothetical protein
Rv2752c	rnj	*M. bovis, Mtb, M. leprae, M. marinum, M. smegmatis*	Conserved hypothetical protein
Rv2783c	gpsl (pnp)	*M. marinum, M. leprae, M. bovis, Mtb, M. smegmatis, M. abscessus*	Involved in mRNA degradation. Hydrolyses single-stranded polyribonucleotides processively in the 3' to 5' direction.
Rv2902c	rnhB	*M. marinum, M. leprae, M.bovis, Mtb, M. abscessus*	Probable ribonuclease HII protein RnhB
Rv2907c	rimM	*M. marinum, M. leprae, M. bovis, Mtb, M. smegmatis, M. abscessus*	Essential for efficient processing of 16S rRNA. Probably part of the 30S subunit prior to or during the final step in the processing of 16S free 30S ribosomal subunits. It could be some accessory protein needed for efficient assembly of the 30S subunit.
Rv2925c	rnc	*M. marinum, M. leprae, M. bovis, Mtb, M. smegmatis, M. abscessus*	Digests double-stranded RNA. Involved in the processing of ribosomal RNA precursors and of some mRNAs
Rv3853	rraA	*M. leprae, M. bovis, Mtb*	Regulator of RNase E activity a RraA
Rv3923c	rnpA	*M. marinum, M. leprae, M. bovis, Mtb, M. smegmatis, M. abscessus*	Ribonuclease P protein component RnpA

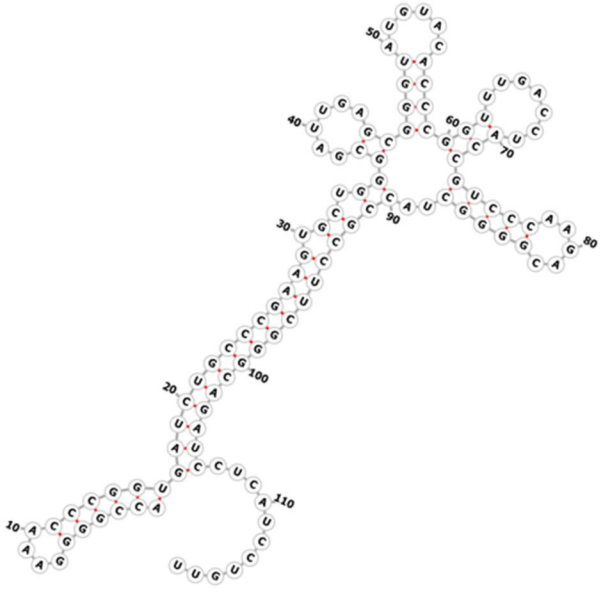
Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 July 2021

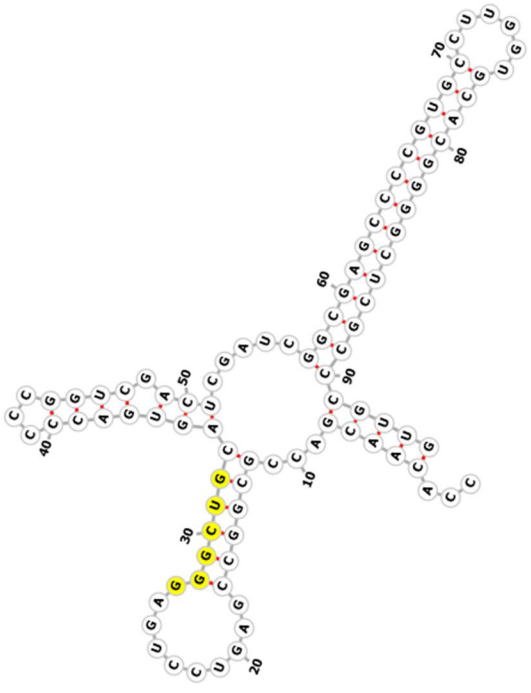
B11/6C

MTS1338/DHS

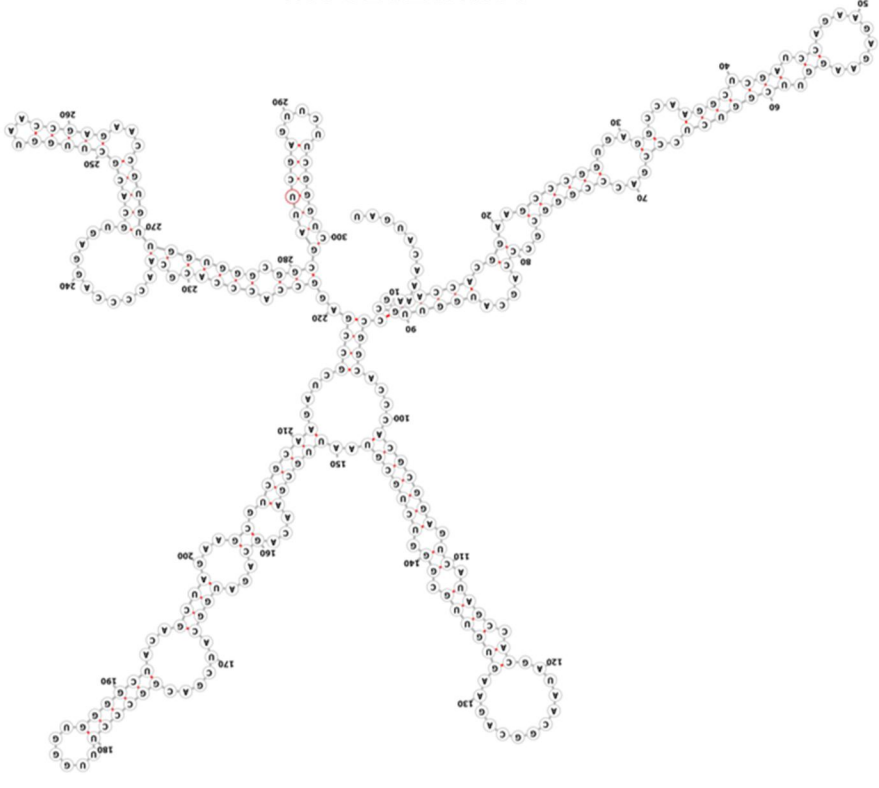
ncRV11846/MrsI



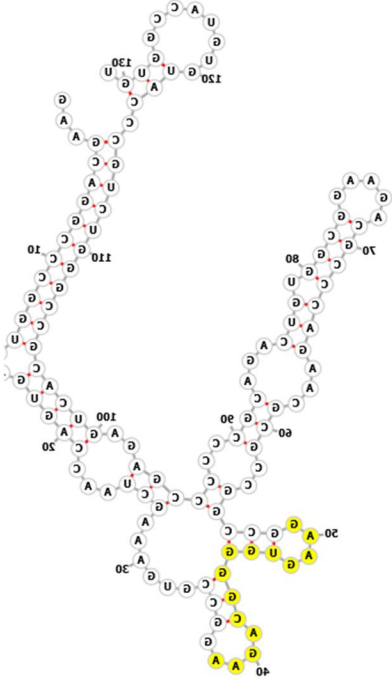




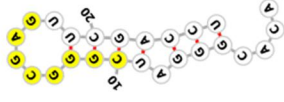
MTS2823/Ms1



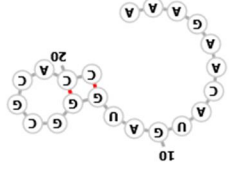
MTS0997/Mcr11



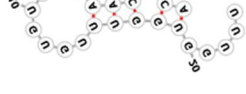
sncRNA-1



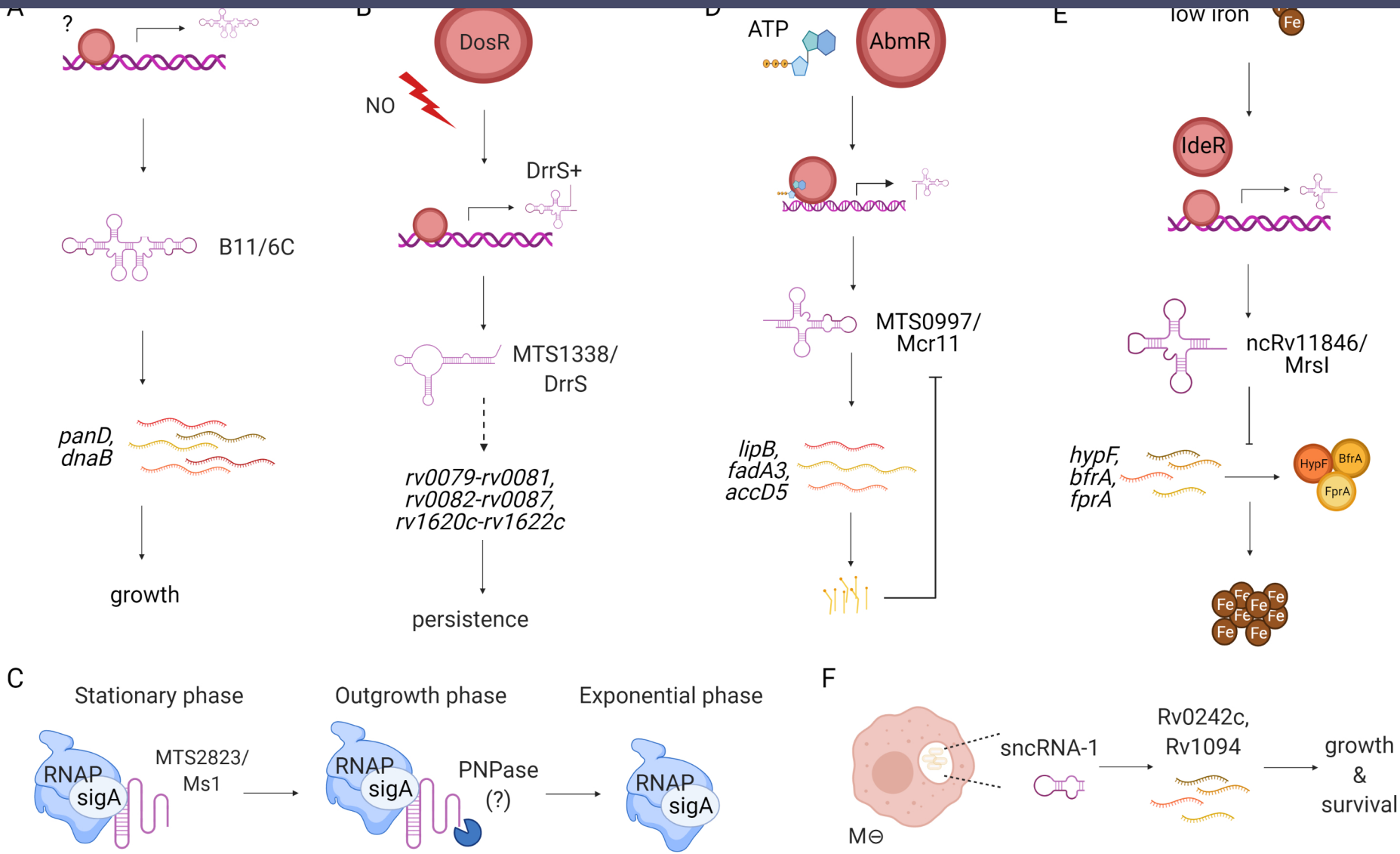
sncRNA-6



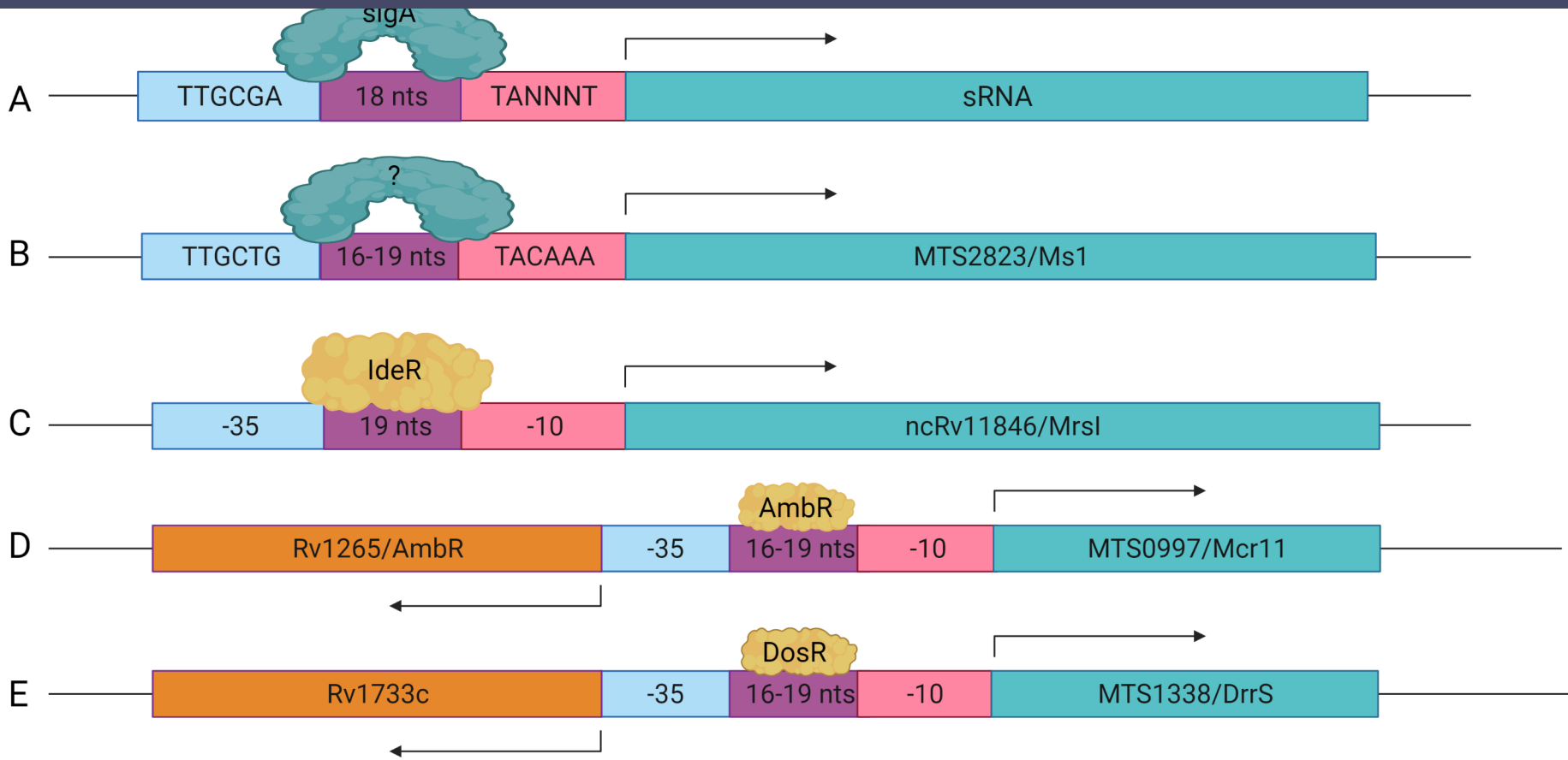
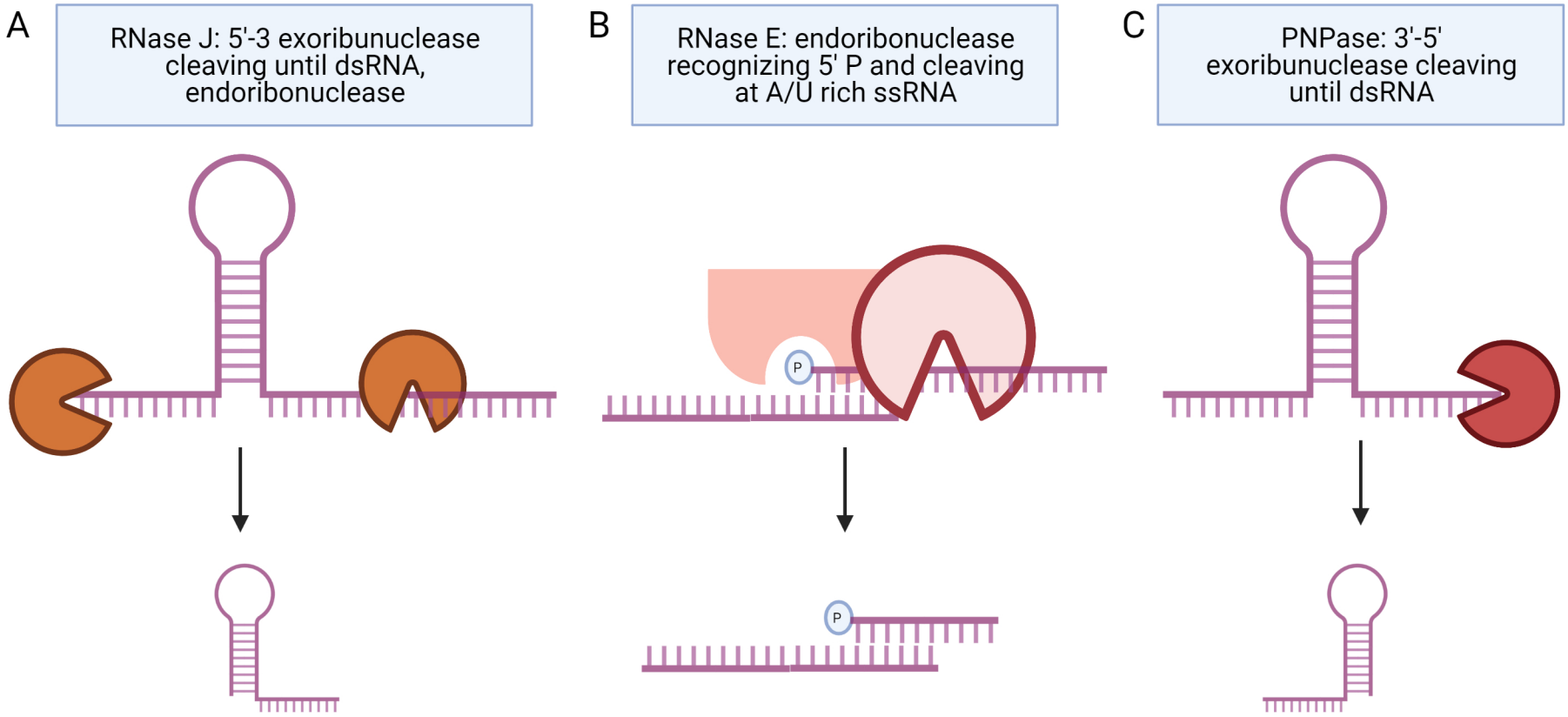
sncRNA-8



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 July 2021



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 July 2021



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 July 2021