

The Diversity of The Classification of Non-coding RNAs

Amanda CG*

Department of Internal Medicine,
Federal University of Parana, Curitiba-PR, Brazil

Abstract

The genes that encode regulatory RNAs - known as short RNAs (sRNAs) or non-coding sRNAs (ncRNAs) - modulate physiological responses through different mechanisms, such as RNA-RNA interaction or RNA-protein interaction. These molecules are transcribed in trans and in cis relative to the targeted RNA. They are located within the protein coding regions, in the intergenic regions of the genome and show signs of promoter and terminator sequences that are generally Rho-independent. The size of the ncRNA genes ranges from ~ 50 to ~ 500 nucleotides and several transcripts are processed by RNase with smaller end-products. These modulate the physiological responses through different mechanisms, either by RNA-RNA or RNA-protein interactions, and some of the interactions can be stabilized by the Hfq chaperone. The Riboswitches constitute another class of ncRNAs that are located in the 5'UTR region of an mRNA and induce transcriptional regulation through their molecular interactions with linkers. CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) regions have been recently described in prokaryotes, which are based on repeated palindromic sequences. Each replicate consists of small segments of "spacer DNA" taken from exposures prior to the isolation of a bacteriophage virus or exogenous plasmid. CRISPR can be defined as an immune system of resistance to exogenous molecules.

Keywords: ncRNA; cis-encoded ncRNA; Trans-encoded ncRNA; Riboswitch; CRISPR

Received: January 17, 2019; **Accepted:** February 18, 2019; **Published:** February 25, 2019

Introduction

Non-coding RNAs can be classified as short non-messenger RNAs (snmRNAs), small non-coding RNAs (ncRNAs), untranslated RNA molecules, or non-protein-encoding RNAs (npcRNAs) [1]. Short RNAs (small RNAs or sRNAs or non-coding RNAs or ncRNAs) are molecules that modulate physiological responses through different mechanisms involving RNA-RNA interaction or RNA-protein interaction. Some interactions can be stabilized by the chaperone Hfq [2-4], which often occurs in the class of trans-encoded ncRNAs where the formation of the ncRNA:Hfq:mRNA complex may act positively or negatively on post-transcriptional regulation [5].

The most widely studied ncRNAs are the cis-encoded RNAs and the trans-encoded RNAs, the first transcripts being in cis-natural antisense relative to the targeted mRNA and the second transcripts in genomic regions that are far from the targeted mRNA [6]. Hence, only the cis-encoded RNAs have a perfect base pairing with the targets. The Riboswitches constitute another class

of ncRNAs that are located in the 5' UTR region of an mRNA, and induce transcriptional regulation through their interaction with a linker molecule [7-9]. CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) regions have been recently described that are based on repeated palindromic sequences [6,10]. Each replicate consists of small segments of "spacer DNA" from exposures prior to the isolation of a bacteriophage virus or exogenous plasmid. The CRISPR can be defined as an immune system of resistance to exogenous molecules [11].

Computational tools are widely used for the prediction of ncRNAs, such as a) QRNA [12] which conducts a comparative analysis of genomes, and b) ISI [13] which analyzes conserved intergenic sequences between the genomes for the identification of ncRNAs. The RNAz 2.0 computational program [14] is employed to analyse the thermodynamic stability of the conserved RNA structure and the possible existence of a promoter and terminator in the predicted ncRNAs. The prediction tools (sRNAPredict2) [15] and SIPHT [16] evaluate information obtained from the Rho-independent terminator promoters within the database. In

***Corresponding author:** Amanda CG

✉ amanda.ufpr@gmail.com

Department of Internal Medicine, Federal University of Parana, Curitiba-PR, 80.060-240, Brazil.

Tel: +55-41-99833-0124

Citation: Amanda CG (2019) The Diversity of The Classification of Non-coding RNAs. J Genom Gene Study Vol.2 No.1:1

addition, the large-scale experimental strategy involving cDNA sequencing (RNA-seq) has been widely used in the prediction of ncRNAs and is a very efficient method owing to its ability to confirm the gene expression of predicted ncRNAs [17-20]. The aim is to provide an understanding of the role played by this type of RNA in the regulation of bacterial metabolism.

Non-coding RNAs

Scientific and technological advances have led to the discovery of the functions of what can be regarded as the main RNAs: a) The messenger RNA (mRNA), b) Molecular information transfer during protein synthesis; c) The RNA ribosomal (rRNA) component of the structure of protein synthesis and d) The RNA transporter (tRNA) that is capable of transporting amino acids and interacting with proteins [21]. Several structural and functional studies of RNA have been carried out and allowed new classes of RNA to be described, with various functions in the *Archea*, *Bacteria* and *Eukaria* domains [22].

While genomics is concerned with analyzing genomes as coding sequences for mRNAs, rRNAs and tRNAs, RNomics further investigates the RNA-encoding genes that are untranslated but are functional and involved in different cellular processes [23,24]. Non-coding RNAs are involved in several cellular processes, such as the following: chromosome replication in cell division (dif), transcription regulation (6S RNA), RNA processing (mRNA), mRNA stability and translation (antisense sequence - spot42), protein stability (tmRNA) and transport (4.5S or ffs) [25]. There are ncRNAs involved in oxidative stress (oxyS), stationary phase transcription (dsrA, rprA), related to the control of plasmid copy number (RNAI, RNAIII), carbon storage (csrBC) and carbon transport (gcvB) [24].

The class of ncRNAs with regulatory function is involved in several regulatory mechanisms such as gene expression in the modulation of outer membrane surface proteins (Omp) [26]. Some ncRNAs may bind to proteins as a means of modulating their activities, as is the case of the 6S RNA that forms a complex with RNA Polymerase (RNAP) [2,24,27]. These molecules are transcribed in trans or in cis relative to target RNA [4,28]. The genes are located within the protein-coding regions, that is, in the

intergenic regions of the genome and display signs of promoter and terminator sequences, that are generally Rho-independent [29-32].

The size of the ncRNAs genes ranges from ~ 50 to ~ 500 nucleotides and several transcripts are processed by RNase with fewer end-products [20,33-35]. The antisense pairing between the regulatory ncRNA and the target messenger RNA is the most common active mechanism [36,37]. This interaction occurs in small regions where there is an imperfect sequence complementarity, and can be stabilized by the Hfq chaperone protein [26,38,39].

Jager and collaborators [40] demonstrated that the interaction of ncRNA162 targeted RNA in *Archaea* can occur through cis-encoded or trans-encoded action, within two distinct domains. As in bacteria, ncRNAs in *Archaea* are involved in many biological processes, such as metabolic regulation, adaptation to environmental conditions, stress response, regulation of morphology, and cellular behavior. In *Methanosarcinamazei* GO1, Babski et al. [41] identified ncRNAs that are aligned in the 5'UTR region of targeted mRNAs and in *Sulfolobus solfataricus* ncRNAs that are aligned in the 3'UTR region of the targeted mRNAs. In the case of *Pyrobaculum* sp. and *Haloferax volcanii*, there is already evidence that ncRNAs can be aligned with targeted mRNAs at both the 5'UTR end and the 3'UTR end. In addition, it has been noted that the nRNA TRF downregulates translation in response to the RBS binding site.

The mode of regulatory action taken by the ncRNAs depends on their co-localization with their targeted mRNAs. They can be classified as trans-encoded RNA, when encoded far from their mRNA targets, and as cis-encoded when located in the 5'UTR region relative to the target. The riboswitches that are mainly found in the 5'UTR region undergo conformational changes in their secondary structure, owing to the binding of a ligand, and regulate gene expression [27,34]. The mechanism of action taken by the ncRNAs varies in accordance with their function (**Figures 1 and 2**). In the post-transcriptional stage, the cis-encoded ncRNAs can cause mRNA degradation (**Figure 1a**), translation inhibition (**Figure 1b**), cleavage of the targeted mRNA (**Figure 1c**) (5'UTR Overlapping) and transcription termination (**Figure 1d**) [42]. The

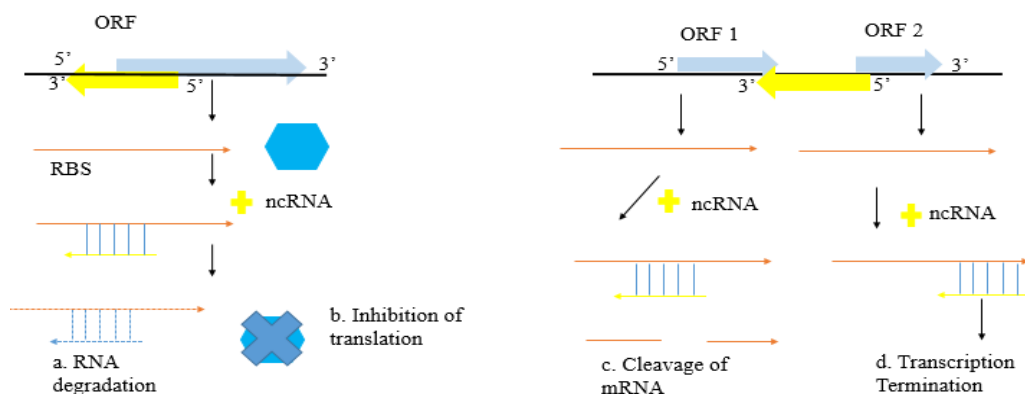


Figure 1 Mechanisms of cis-encoded action in which, the ncRNA in the genome is located close to the gene of its targeted mRNA. The cis-encoded ncRNAs are highlighted in yellow and their target in blue. (**1a**) Degradation of mRNA (**1b**) Inhibition of translation. (**1c**) Cleavage of mRNA and (**1d**) Transcription termination.

interaction of the trans-encoded ncRNA with the targeted mRNA may result in translation inhibition (**Figure 2a**) and degradation of mRNA caused by the negative interaction of 5'UTR that inhibits the ribosome binding site, or in the degradation triggered by the RNase pairing status (**Figure 2b**). The ncRNA that is trans-encoded with its targeted mRNA may form an inhibitory structure that could in some way block the ribosome binding site, thereby impeding translation (**Figure 2c**) [42]. Many cis-encoded RNAs with antisense orientation can form binary trans-encoded complexes with the targeted mRNA, which demonstrates that this is not a strict or definitive classification [43].

The Cis-encoded ncRNAs

There are cis-encoded ncRNAs in different bacterial species, (Gram-positive and Gram-negative) [39,43]. In *E. coli*, Wagner and Simons [44] described the antisense regulatory role played by RNA in the control of mRNA gene expression, phage maturation, mobile transposition and plasmid replication. In addition, they may be involved in regulating the initiation of replication, plasmid conjugation, transposition, mRNA degradation and some cell metabolism pathways. These factors continue to be the subject of investigation by different researchers [4,28,43,45].

The cis-encoded ncRNAs are encoded at the same locus as their targeted mRNA, but in the antisense -sense duplexes, and thus remain fully complementary during the interaction. The mechanism for the post-transcriptional response of gene regulation involves a high degree of sequence complementarity, and this was considered to be an indication that interaction of

the Hfq protein would not be required [23,45]. However, some researchers reported that there was interference of Hfq in the cis-encoded ncRNA pairing [46-48]. In general, these ncRNAs act by complementing the mRNA ribosome-binding site, and then by inhibiting, in turn, the translation [38,45].

An example of a typical cis-encoded ncRNA is the 5'*ureB* of *Helicobacter pylori*, located at 5'- antisense to the *ureB* gene that makes up the *ureAB* operon [49,50]. This ncRNA contains 292 bp and negatively regulates *ureAB* operon expression by blocking the translation in the 5'portion of *ureB* (**Figure 3**). The *ureAB* genes of *H. pylori* are located in the cluster of two *ureAB-ure IEFGH* operons and encode the *UreA* and *UreB* subunits of apoenzyme urease. This enzyme is essential for the survival of the *H. pylori* at low pH since its reaction releases NH_3 and HCO_3^- into the environment, thus allowing homeostasis for bacterial growth [49]. The diversity of the cis-encoded ncRNAs and their regulatory roles vary in accordance with the different organisms. For example, *Salmonella enteric serovar Typhimurium* possesses the cis-encoded ncRNA *lesR-1* where the function is to control replication in eukaryotic cells [51] and *Salmonella serovar Typhi* possesses the cis-encoded ncRNA *AmgR* depending on its virulence in rats [52] and *AsdA* which regulates intracellular replication [53].

The same diversity can be found with regard to regulatory strategies. In *Escherichia coli*, the base pairing between the cis-encoded ncRNA *GadY* and its *gadXW* targeted mRNA, causes the cleavage of the duplex between *gadX* and *gadW*, and leads to increased levels of the *gadX* transcript. The *GadX* product acts

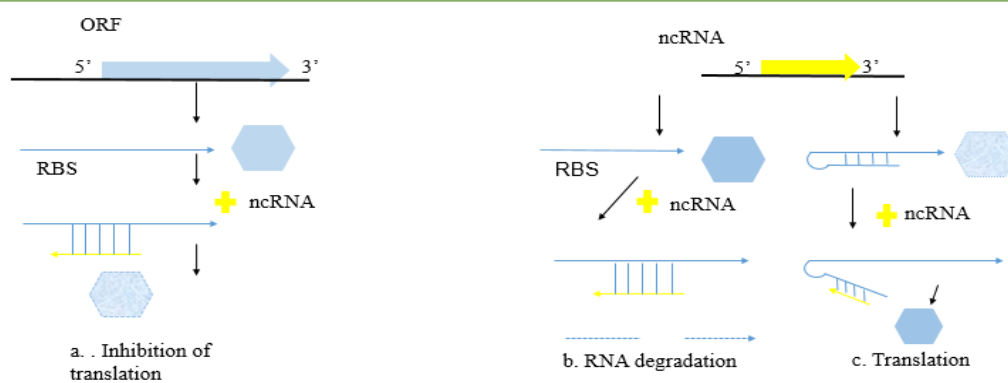


Figure 2 Mechanisms of trans-encoded action in which, the ncRNA in the genome is located far from the gene of its targeted mRNA. The trans-encoded ncRNAs are highlighted in red and their target mRNA in blue. There is limited base pair complementarity. (**2a**) Inhibition of translation (**2b**) Degradation of mRNA (**2c**) Allows translation.

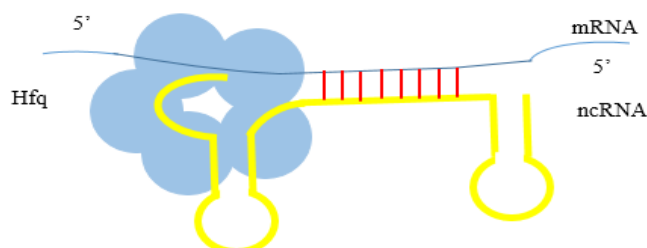


Figure 3 Model of the probable interaction of the Hfq protein with the ncRNA/mRNA pair.

as a transcription factor for the GadAGadB operon during the synthesis of glutamate decarboxylase and this process entails employing an acid stress deflating system in *E. coli* [6,42].

The *SymR-SymE* system in *E. coli* consists of two genes the cis-encoded ncRNA *symR* and the *symE* gene that encode a toxic protein [4,6,38]. The increase in the cellular concentration of the *SymE* protein reduces the synthetic activity of the ribosomes. The cis-encoded ncRNA *symR* negatively regulates the expression of the *symE* gene through the complementarity of the ncRNA/mRNA bases, resulting in inhibition of mRNA *symE* translation and the resumption of synthetic ribosomal activity [6,38,42,50].

In the *Brucella abortus* 2308, Peng et al. [54] identified the cis-encoded ncRNA BsrH that positively regulates the expression of the *hemH* gene, thus providing evidence of the importance of the regulatory expression that ncRNABsrH exerts on its targeted mRNA.

The cis-encoded antisense ncRNAs are also often found in the replication mechanism of plasmids. For example, in the replication control of the ColE1 plasmid, which uses ncRNA instead of proteins to initiate replication at the site of origin, two partially complementary RNAs are transcribed from opposite strands. The larger RNA, with 250-500 nucleotides (RNAII), is transcribed from the sense strand and forms a stable hybrid with a DNA template. This hybrid is then processed by the RNase H to produce a primer design for the DNA polymerase. The smaller RNA, which has 68-108 nucleotides (RNAI), is transcribed from the antisense strand and is complementary to the 5' RNAII region. It functions as a negative regulator for the formation of a primer by forming the RNAI/RNAII duplex that prevents the formation of a hybrid RNAII/DNA template. The concentration of RNAI is proportional to the number of plasmids per cell, and thus constitutes a negative feedback loop that regulates plasmid replication in response to metabolic changes [37,55].

The Trans-encoded ncRNAs

As previously mentioned, the mode of action of trans-encoded ncRNAs differs from cis-encoded ncRNAs since it involves a limited sharing of base pairs with their targeted mRNAs [42,50,56]. This type of ncRNA is encoded in trans and may have targeted mRNAs at different locations in the genome.

The transcribed ncRNA generally requires the chaperone Hfq protein to stabilize the targeted ncRNA-RNA interaction due to imperfect base pairing and thus can prevent its eventual degradation by RNase [39,42,56,57]. The most widely studied chaperone is the Hfq protein that in *E. coli* interacts with 40% of the ncRNAs [6]. Schoroeder et al. [58] found that almost 50% of all bacterial species possess trans-encoded ncRNAs that require the chaperone protein Hfq, one exception being *Listeria monocytogenes* here most trans-encoded ncRNAs are independent of Hfq.

The interaction of Hfq with trans-encoded ncRNAs is involved in post-transcriptional regulation in several species of bacteria, and may have either a negative or positive effect on their mRNAs [59,60]. On the basis of studies of crystallography, it has been found that the structure of Hfq has a hexameric protein that is

homologous to Sm proteins that have two motifs (Sm1 and Sm2) [39,54]. Link et al. [61] characterized two RNA binding sites of the Hfq in *E. coli*: a proximal protein that binds to the ncRNA and targeted mRNA and the other distal that binds to the poly (U) tail (Figure 3) [62].

As a result of Fluorescence Resonance Energy Transfer (FRET) studies, Peng et al. [54] found that structural models in the Hfq interact with PAPI, PNPase and RNaseE. Soper et al. [63] described three regulatory ncRNAs, DsrA, RprA, and ArcZ in *E. coli* that positively regulate the translation of the RpoS sigma factor when ncRNA and rpoS mRNA pairing occurs. They detected the formation of an inhibitory clamp in the 5'UTR region and demonstrated that binding to Hfq is important to ensure the stability of the RNA: RNA complex. In the negative regulation, base pairing occurs between the Ribosome Binding Sequence (RBS) and ncRNA that blocks the ribosome binding, or degradation by RNases.

In *E. coli*, regulation of OmpC protein expression involves the MicF ncRNA and a 5'UTR regulation of 22 antisense nucleotides to ompC mRNA [64]. This interaction also involves Hfq and results in translation inhibition. In *Salmonella typhimurium* is the MicC ncRNA, combined with the Hfq protein, which silences the ompD mRNA through the duplex of 12 RNA/RNA base pairs in the protein-coding region. MicC does not inhibit translation initiation in the downstream position, but accelerates RNaseE- dependent activity [65,66].

In *Salmonella enterica serovar Typhimurium*, the trans-encoded ncRNA IsrM, controls the pathogenic factor SPI-1 [67]. In contrast, RybB-1 and RybB-2 are linked to combine with the regulation of oxidative stress response [68]. In *Salmonella enterica serovar Typhi*, the trans-encoded ncRNA RfrA and RfrB play a key role in the regulation of iron homeostasis [69]. In *Chlamydia trachomatis*, the trans-encoded ncRNA IhtA acts as an inhibitor of histone Hc1 protein, whereas in *Neisseria meningitidis*, the trans-encoded ncRNA Nrrf regulates iron homeostasis [70-72]. The following are other examples of trans-encoded ncRNA: AbcR-1 and AbcR-2 in *Brucella abortus* which are related to virulence in rats and are important for the survival of macrophages [73] and Mrc7 in the *Mycobacterium tuberculosis* species which is involved in the regulation of the TAT secretion system [74].

Riboswitches

Riboswitches are the structured elements of non-coding RNA that are considered to be cis-encoded elements of RNA, and are mainly located in the 5'UTR region of a targeted mRNA but less frequent at the 3' end UTR [7,42,50]. However, Loh et al. [75] described a case of a riboswitch that controls in trans the expression of the virulence regulating PrfA protein in *Listeria monocytogenes*. It has the ability to control gene expression at the level of transcription and translation and enable the molecules to acquire different conformations in response to environmental signals, such as high temperatures and the binding of small molecules, such as metabolites or metal ions [9,42,76]. A wide range of riboswitches have recently been detected in prokaryotes. 2% of all the *Bacillus subtilis* genes are regulated by riboswitches that do not bind to intracellular metabolites, such as Flavin Mononucleotide (FMN),

thymine pyrophosphate, S-adenosyl-methionine (SAM), lysine and guanine (**Figure 4**) [43,50,77,78].

The structure of the riboswitch consists of two parts the aptamer region that serves as the binding site for a ligand and the expression platform that provides a suitable conformation for the signal transduction (**Figure 4**) [42,79]. The molecular binding causes conformational changes in the native riboswitch structure, which can regulate the transcription and translation processes [7]. The latter sequence of the expression platform leads to a splicing within the aptamer domain, whether or not it is bound to the linker and may possibly signal a transcription term or control the helical structure at the ribosome-binding site [7,80].

Thus, riboswitches are used to regulate the termination of mRNA transcription (attenuation) and initiate the translation [81]. During the transcription, when the ligand binds to the mRNA in the aptamer region, conformational changes occur that result in the formation of an alternative clamp (**Figure 4a**) [42]. This clamp acts as a transcription terminator that inhibits gene expression. The binding of the linker molecule to the alternative clamp leads to anti-termination (**Figure 4b**).

Moreover binding of a ligand that causes structural changes and it can lead to RBS sequestration, prevents translation (**Figure 4c**). In contrast, the binding of a linker may cause RBS exposure to binder and induce translation (**Figure 4d**).

Riboswitches may play a role in biological cell systems because of the large number of gene families involved. Corbino et al. [82] found methA motifs in *Agrobacterium tumefaciens* similar to the S-adenosylmethionine by decoding the riboswitch (SAM). The SAM-II riboswitch class has a great structural diversity, and low conservation, since it is able to alter the conformational structure of the mRNA [7,83-85].

The structure of the riboswitch may contain several motifs that control gene expression by detecting determined metabolite concentrations, which makes these structures promising targets for antibiotics. Suresch et al. [86] have shown that the S-adenosylmethionine riboswitch-III that is found in anaerobic bacteria is involved in the methionine and SAM biosynthetic pathway regulation process. In either the presence or absence of the ligand, S-adenosylmethionine riboswitch-III plays a dual

role, which facilitates the conformational change between the partially and fully folded state, and forms a stable duplex structure, which strengthens the interactions between the Shine-Dalgarno nucleotides (SD) and anti-Shine-Dalgarno (aDS).

Perez et al. [87] employed an *in vivo* gene expression system to validate the cobalamin - riboswitch of the Cyneobacterium, *Synechococcus sp.* strain PCC 7002 and concluded that methionine biosynthesis is probably the only means of using cobalamin in this strain of cyanobacteria.

The ncRNAs That Induce Protein Activity

The first ncRNAs characterized in the *E. coli* enterobacterium, include 4.5S, 6S, tmRNA (RNA transfer-messenger) and Spot42 [42,88]. In *E. coli*, Spot42 RNA was discovered almost 40 years ago as a small and unstable RNA molecule encoded by the *spf* gene that is also present in *Vibrionaceae*, a class of γ -proteobacteria. *Salmonicidal alivivium* encodes a Spot42 RNA ncRNA with 84% identity compared with that of *E. coli*. Deletion of the *spf* gene results in a 25% decrease in the action of RNA polymerase I (Pol I) [89]. This ncRNA optimizes carbon dioxide uptake and metabolism by binding to the targeted mRNAs related to this metabolism, such as galactose operon galK. The ncRNA Spot42 is an example of ncRNA that acts in a traditional way through the ncRNA/mRNA interaction, but there are cases, such as the CsrB and the 6SRNA that regulate targeted protein activity and usually have specific recognition sequences [39].

The *E. Coli* CsrB ncRNA has 22 GGA sequences that are regarded as binding sites for the CsrA protein [88]. By binding to the protein, it is able to regulate mRNA stability and translation. The CsrA protein regulates carbon storage [90] through its binding to the mRNAs and is the essential component of the Csr regulatory system. This system is responsible for a) The repression of a wide range of stationary -phase genes, b) Negatively regulating gluconeogenesis, glycogen biosynthesis and catabolism, and c) Biofilm formation. In addition, it activates the glycolytic pathway, acetate metabolism and flagellum biosynthesis. CsrA acts post-transcriptionally by repressing the gene expression of enzymes that are essential for the metabolism of carbohydrates such as

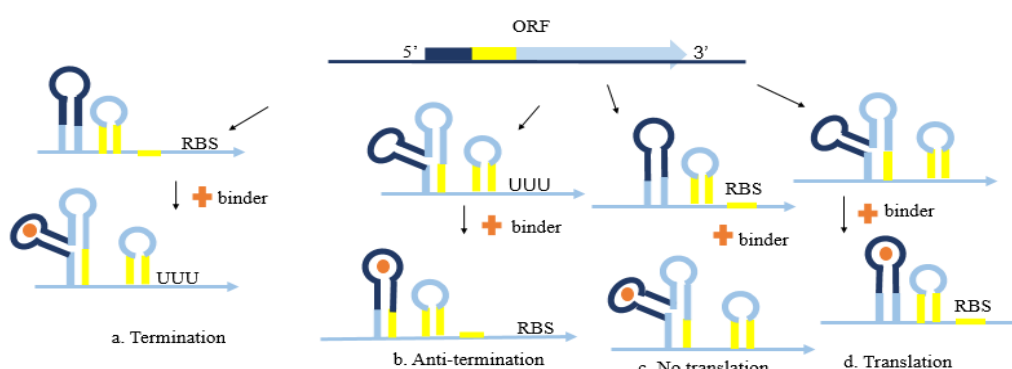


Figure 4 Riboswitch structural arrangement and regulatory function for the binding. The aptamer region (pink) and an expression platform (yellow) on the 5' UTR of the respective mRNA (blue), form the Riboswitch. Molecular binding has a regulatory function in the processes of transcription (a and b) and translation.

ADP-glucose pyrophosphorylase (glgC), glycogen synthase (glgA), glycogen branching enzyme (glgB), and glycogen phosphorylase (glgP).

Thus, there is a decrease in the intracellular levels of biosynthetic glycogen enzymes and a reduction of glycogen synthesis [91]. CsrA destabilizes its targeted mRNAs by binding to the nucleotide-coding region of nucleotides -18 and +3, which includes (RBS), and thus preventing the translation of the mRNA and leading to its degradation by RNase. The *csrB* ncRNA competes directly with the targeted mRNA of CsrA and the dimeric protein CsrA recognizes the GGA motif of the structure of the *nsRB* ncRNA and the targeted mRNA. The CsrC and CsrB ncRNAs in *E. coli* modulate the activity of CsrA and the binding of the targeted mRNA to the protein [39,42]. Thus, intracellular levels of CsrA are also regulated by the CsrB and CsrC ncRNA switch, and act as antagonists to capture CsrA. In *Erwinia spp*, the RsmA protein that is, homologous to CsrA, regulates several genes involved in plant disease. In *Pseudomonas aeruginosa*, Csr (*Rms*) controls several systems like Rhl and several virulence factors [92].

Hindley identified 6S RNA in 1967 in *Escherichia coli* [25]. Since then, 6S RNA has been predicted through computational tools and experimentally analyzed in several bacterial species [93]. Wehner et al. [94] analyzed 1611 bacterial genomes and determined a set of 1,750 RNA 6S genes, 1,367 of which were new. In the Rfam database, the 6S RNA is described by two entries - the first RF00013 with 153 sequences, and the second RF01685 with 89 sequences. It has been demonstrated in *E. coli* that 6S RNA (the *ssrS* gene) can act in the transcription process by combining with the sigma 70 factor that is dependent on RNA polymerase (6S RNA:RNAP), so that it can modulate its function. This ncRNA sequesters almost all of the RNAP holoenzyme in late stationary phase and thus assists in the transcriptional adaptation of the stationary phase of growth [87,93,94]. With regard to the position of 6S RNA in the genome, there is a synthetic pattern for the Bacteria domain and this pattern is most closely followed in Gamma-Proteobacteria. Among the common synaptic genes are the following: *ygfA* (5-formyltetrahydrofolate cyclo ligase), *zapA* (ZapA protein), AAA (+) (ATPase superfamily) and *peptidase_M24* (PF00557 family of metallopeptidases) [95,96]. Gutiérrez et al.

[92] identified the ZapA protein in *Bacillus subtilis* together with the FtsZ-tubulin protein.

CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)

In 2007, Barrangou [97] and colleagues demonstrated the change in resistance to phage infection found in bacterial cells, by removing or adding spacer sequences similar to the sequences of invading phages. This strategy defined an adaptive immune defense system that uses short RNA and is called Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR). The CRISPR locus contains hundreds of spacer sequences (~ 26-72 nucleotides) in the genome, flanked by repeats of 21 to 47 nucleotides. They are usually combined with proteins (*cas*), that form the CRISPR system, and this confers RNA-mediated resistance to nucleic acids from, for example, bacteriophage, plasmids or mobile genetic elements [1,43,98,99]. The CRISPR locus can be found in ~ 40% of bacterial genomes and 90% of archaeal genomes [42,100,101]. Most of the prokaryotes containing CRISPR have multiple groups of 2 to 20 loci, each arranged in tandem and containing up to 100 identical repeats among 25 and 50 base pairs [101]. Cas proteins provide the enzymatic machinery required to acquire new spacers and create recognition marks of invading elements [102].

The arrangement and processing of CRISPR RNAs involves a number of phases (Figure 5) [102]:

- Adhesion: invasive DNA is integrated into the CRISPR locus, resulting in new modes of replication in the matrix (Figure 5a);
- Expression: transcription of the written record containing the RNA phage (new spacer) and its processing by the Cas proteins to form crRNA (crRNA matrix formed of single spacer sequences flanked by short repeats) and adjacent coded Cas proteins (Figure 5b);
- Interference: this is the formation of the complex CrRNAs/Cas proteins and later, the invasion of the targeted DNA takes place and leads to its respective degradation (Figure 5c).

CRISPR may be expressed in species such as *Legionella pneumophila*, where sRNA is the Cas2-dependent crRNA involved

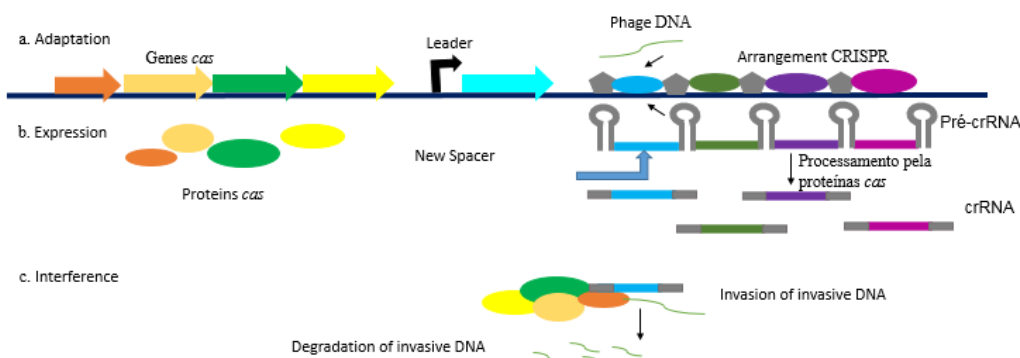


Figure 5 Array model of the CRISPR system\Cas. The. Adaptation: DNA phage integrated into the CRISPR matrix and formation of the new spacer. Expression: Processing of *cas* proteins. W. Interference: Invasion of targeted DNA and its respective degradation.

in the stress response [103-105]. In *Listeria monocytogenes*, the RliB-CRISPR sRNA also makes use of the CRISPR\Cas system in antiviral resistance [106,107]. *Francisellanoviciida* possesses the ncRNAs, Cas9-dependent crRNA, tracrRNA and scaRNA, with CRISPR\Cas property in the regulation of endogenous viral factors [105]. According to Marraffini [108], the strains of *Streptococcus pyogenes* have the CRISPR\Cas9 system and their function is to mediate the anti-phagocytic resistance mechanisms. In *Campylobacter jejuni* PT14, the CRISPR\Cas14 system protects against the invasion/presence of bacteriophage [109].

Amitai and Sorek [110] studied the mechanisms involved in the adaptation of CRISPR-Cas, with regard to the perception that once the exogenous DNA has been identified by the CRISPRC as a system, it is integrated into the genome of the host cell, thus creating immunological memory, which is a natural ability based on the information contained in the DNA. This kind of system must be properly manipulated in the formation of new bases of information storage in living organisms, something that has been extensively investigated in recent years.

Prediction of ncRNAs and ats Targets

Bioinformatics is devoted to designing computer programs for the handling of biological data and the identification of gene sequences, the prediction of three-dimensional protein structure, the identification of enzyme inhibitors, the system of protein grouping, the establishment of phylogenetic trees and experiments in gene expression analysis [111]. It provides the tools required for Genomics, Transcriptomics, Proteomics and Metabolomics [112].

With regard to the prediction of ncRNAs, several computational approaches have been adopted for the identification of genes in the intergenic regions of prokaryotic genomes [113,114]. Many of these are based on the search for transcriptional signals, conserved promoter sequences, rho-independent terminators, the transcription factor binding site, such as sRNAPredict [115], a putative predicting algorithm for analysis using the TransTermHP database [116,117] or TRANSFER [118]. sRNAPredict3 and SIPHT are recent computational versions for the prediction of ncRNAs in bacteria [16]. sRNAscanner and sRNAfinder [119] were designed to overcome the limitations of the predictive capacity of the available transcription signals in all the genomic sequences and proved to be efficient.

The nocoRNAC (non-coding RNA characterization) is a computational tool that was developed to study the interactions between ncRNA-mRNA in conjunction with the prediction of ncRNAs in bacterial genomes. This program uses transcription termination signals that are predicted through the TransTermHP tool; the promoters are identified by the Stress Induced Duplex Destabilization (SIDD) model and determine any possible regions that can be destabilized [120,121]. Thus, it is possible to detect the regions of the genome flanked by the promoter sequence and Rho-independent terminator sequence, which is the candidate selected to encode the ncRNAs. In an alternative approach, the Cufflinks tool locates regions of the genome that have considerable levels of transcription and free ORFs [122].

Comparative genomic analyses have been conducted to predict new ncRNAs, and conserved sequences identified for the first time in the Intergenic Region (IGR), and shared by clustering; a comparison is made with the multiple alignments that are classified as ncRNAs. Programs such as QRNA [12], ERPIN [123], ISI [13], INFERNAL.1.1.1 [124], MSARI [125] and RNAz [126] are compared in terms of their thermodynamic stability through a prediction of their conserved stable RNA structures, and ncRNAs in bacteria [14,127]. In the case of secondary structures, the RNAfold carries out a statistical analysis of the RNA folding, through the disturbance of the thermodynamic parameters for equilibrium, to assess its predictive capacity [119,128,129].

With regard to the target prediction, it is very important to develop models with the aim of integrating bioinformatics for prediction with experimental validation for the confirmation of mRNA targets. The classification of ncRNAs provides information on the complementarity of the bases (whether perfect or imperfect) with their targeted mRNAs and eventual binding to proteins, by altering their activity [2,130-133]. The targetRNA2/targeting mRNA prediction tool is currently one of the most widely used [134]. It employs various features including a) The ability to conserve ncRNA in other bacteria, b) The secondary structure of both the ncRNA and each target candidate, and c) The hybridization energy levels between the interactions. It also has the ability to integrate data from RNA-seq material when available.

Another computational approach used to predict targeted mRNAs is the IntaRNA tool [135], which is also regarded as quite efficient and rapid in predicting the interactions between ncRNA/mRNA. It uses energy-free hybridization and is integrated with the CopraRNA tool, which predicts ncRNAs by comparing the query-sequence with available sequences in the program [135,136]. Other approaches include direct detection by means of microarrays, and northern blotting, [3,30,111,137,138].

Pnek et al. [139] have found predicted ncRNA in *Streptomyces* that are based on a study of sequence conservation in intergenic regions, the location of the transcription termination factor, and the genomic arrangement of syngenic genes for ncRNAs. They detected the expression of 20 ncRNAs by microarray and RT-PCR, and adopted a computational approach to determine their secondary structure; as a result, they identified 6S ncRNA. Voss et al. [140] used a cyanobacteria model to predict ncRNAs from transcriptome and proteome data and identified the Yfr2a-Yfr2c ncRNA, a conserved structure that can be found among cyanobacteria. In an attempt to predict the existence of the 5'-operon-leader, the Rho-independent 3'-transcription terminator and riboswitches, these authors used the TransTermHP, ClustalW, RNAz, and RNAfold computational tools for validation with the aid of Northern Blot [141].

Modi et al. [137] carried out a functional characterization of ncRNAs in *E. coli* through network inference, based on a compendium of gene expression profiles with functional prediction and on the regulatory interactions of ncRNAs. These authors experimentally validated the functions attributed to three ncRNAs, IsrA and GlmZ, involved in DNA damage response, and GcvB, involved in the regulation of amino acid availability.

Khoo and collaborators [142] integrated several computational methods that entailed the prediction and analysis of ncRNAs and identified 29 in *Burkholderia pseudomallei* among which [8] were believed to be new. Ignatov and colleagues [143] used material and transcriptomic analysis to reveal that *Mycobacterium avium* and *M. tuberculosis* contain different sets of ncRNAs in the intergenic regions and suggested that this characteristic may be the basis of the observed physiological differences between the two species. Schroeder et al. [20] observed that in the genus *Rickettsia*, ncRNAs are the main post-transcriptional regulators involved in virulence, survival, plasmid expression, primary and secondary metabolism. This means that they can presumably encode trans-encoded ncRNAs involved in the pathogen interaction and host.

Conclusion and Future Perspectives

Any knowledge of the post-transcriptional regulation that ncRNAs exert on a given targeted mRNA, are related to their position, matched in ncRNA: mRNA, and depend on which sRNA class is being worked on. There are different classes of sRNAs and it is

necessary to know how these can regulate several mechanisms when by forming a regulatory network, which plays a key role in the regulatory circuit of the genome under study. There have been considerable advances in computational technology aimed at providing knowledge of the role and functionality of non-coding RNAs in the various domains of life. The contribution of non-coding RNAs in bacteria is a relatively new area of research, based predominantly on the recent findings that its expression and function are frequently studied in the interaction between plant and bacterium, mainly in the diazotrophic class, of socioeconomic importance. Experimental approaches to provide biochemical understanding in detailed macromolecular modulations induced by post-transcriptional modifications. This will undoubtedly open new avenues of research for a better prognosis of the disease and therapeutic interventions.

Funding

This study was financed in part by the Coordination of Improvement of Higher Education Personnel—Brazil (CAPES)—Finance Code 001.

References

- Hoe CH, Raabe CA, Rozhdestvensky TS, Tang TH (2013) Bacterial sRNAs: Regulation in stress. *Int J Med Microbiol* 303: 217-229.
- Pichon C, Felden B (2007) Proteins that interact with bacterial small RNA regulators. *Fems Microbiol Rev* 31: 614-625.
- Livny J, Waldor MK (2007) Identification of small RNAs in diverse bacterial species. *Curr Opin Microbiol* 10: 1096-1101.
- Liu M, Camilli A (2010) A broadening world of bacterial small RNAs. *Curr Opin Microbiol* 13: 18-23.
- Faner MA, Feig AL (2013) Identifying and characterizing Hfq-RNA interactions. *Methods* 63: 144-159.
- Azhikina TL, Ignatov DV, Salina EG, Fursov MV, Kaprelyants AS (2015) Role of Small Noncoding RNAs in Bacterial Metabolism. *Biochemistry* 80: 1633-1646.
- Winkler WC (2005) Riboswitches and the role of noncoding RNAs in bacterial metabolic control. *Current Opinion in Chemical Biology* 9: 594-602.
- Meyer MM, Hammond MC, Salinas Y, Roth A, Sudarsan N, et al. (2011) Challenges of ligand identification for riboswitch candidates. *RNA Biology* 8: 5-10.
- Izar B, Mraheil MA, Hain T (2011) Identification and Role of Regulatory Non-Coding RNAs in *Listeria monocytogenes*. *Int J Mol Sci* 12: 5070-5079.
- Brosius J, Tiedge H (2004) Nomenclature. *RNA Biol* 1: 81-83.
- Marraffini LA, Sontheimer EJ (2008) CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* 322: 843-1845.
- Rivas E, Klein RJ, Jones TA, Eddy SR (2001) Computational identification of noncoding RNAs in *E. coli* by comparative genomics. *Curr Biol* 11: 1369-1373.
- Pichon C, Felden B (2003) Intergenic sequence inspector: searching and identifying bacterial RNAs. *Bioinformatics* 19: 1707-1709.
- Gruber AR, Findeiß S, Washietl S, Hofacker IL, Stadler PF (2010) RNAz 2.0: improved noncoding RNA detection. *Pac Symp Biocomput* 15: 69-79.
- Livny J, Brencic A, Lory S, Matthew KW (2006) Identification of *Pseudomonas aeruginosa* sRNAs and prediction of sRNA-encoding genes in 10 diverse pathogens using the bioinformatic tool sRNAPredict 2. *Nucleic Acids Res* 34: 3484-3493.
- Livny J, Teonadi H, Livny M, Matthew KW (2008) High-Throughput, Kingdom wide prediction and annotation of bacterial non-coding RNAs. *PLoS One* 3: e3197.
- Sharma CM, Vogel J (2009) Experimental approaches for the discovery and characterization of regulatory small RNA. *Curr Opin Microbiol* 12: 536-546.
- Shinhara A, Matsui M, Hiraoka K, Nomura W, Hirano R (2011) Deep sequencing reveals as-yet-undiscovered small RNAs in *Escherichia coli*. *BMC Genomics* 12: 1471-2164.
- Pichon C, du Merle L, Caliot ME, Trieu-Cuot P, Le Bouguéne C (2012) An in silico model for identification of small RNAs in whole bacterial genomes: characterization of antisense RNAs in pathogenic *Escherichia coli* and *Streptococcus agalactiae* strains. *Nucleic Acids Res* 40: 2846-2861.
- Schroeder CLC, Narra HP, Mark M, Sahni A, Patel J, et al. (2015) Bacterial small RNAs in the Genus *Rickettsia*. *BMC Genomics* 16: 1-18.
- Cox MM, Doudna JA, O'donnell M (2012) *Biologia Molecular-Princípios e Técnicas*. Artmed.
- Brosius J, Raabe CA (2016) What is an RNA? A top layer for RNA classification. *RNA biology* 13: 140-144.
- Hüttenhofer A, Schattner P, Polacek P (2005) Non-coding RNAs: hope or hype? *Trends Genet* 21: 289-297.

- 24 Ulvé VM, Sevin EW, Chéron A, Barloy-Hubler F (2007) Identification of chromosomal alpha-proteobacterial small RNAs by comparative genome analysis and detection in *Sinorhizobium meliloti* strain 1021. *BMC Genomics* 8: 1-16.
- 25 Mikulík K (2003) Structure and functional properties of prokaryotic small noncoding RNAs. *Folia Microbiol (Praha)* 48: 443-468.
- 26 Massé E, Majdalani N, Gottesman S (2003) Regulatory roles for small RNAs in bacteria. *Curr Opin Microbiol* 6: 120-124.
- 27 Gottesman S, Storz G (2001) Bacterial Small RNA Regulators: Versatile Roles and Rapidly Evolving Variations. *Cold Spring Harb Perspect Biol* 3: a003798.
- 28 Storz G, Altuvia S, Wassarman KM (2005) An abundance of RNA regulators. *Annu Rev Biochem* 74: 199-217.
- 29 Ermolaeva MD, Khalak HG, White O, Smith HO, Salzberg SL (2000) Prediction of Transcription Terminators in Bacterial Genomes. *Mol Biol* 301: 27-33.
- 30 Wassarman KM, Repoila F, Rosenow C, Storz G, Gottesman S (2001) Identification of novel small RNAs using comparative genomics and microarrays. *Genes Dev* 15: 1637-1651.
- 31 Argaman L, Hershberg R, Vogel J, Bejerano G, Wagner EG, et al. (2001) Novel small RNA-encoding genes in the intergenic region of *Escherichia coli*. *Curr Biol* 11: 941-950.
- 32 Ciampi MS (2006) Rho-dependent terminators and transcription termination. *Microbiology* 152: 2515-2528.
- 33 Wassarman KM, Kiley PJ (2010) Global approaches for finding small RNA and small open reading frame functions. *J Bacteriol* 192: 26-28.
- 34 Arnvig KB, Cortes T, Young DB (2014) Noncoding RNA in *Mycobacteria*. *Microbiol Spectrum* 2: 1-24.
- 35 Updegrove TB, Zhang A, Storz G (2016) Hfq: the flexible RNA matchmaker. *Curr Opin Microbiol* 30: 133-138.
- 36 Sharma CM, Vogel J (2009) Experimental approaches for the discovery and characterization of regulatory small RNA. *Curr Opin Microbiol* 12: 536-546.
- 37 Massé E, Gottesman S (2002) A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*. *Proc Natl Acad Sci* 9: 4620-4625.
- 38 Massé E, Majdalani N, Gottesman S (2003) Regulatory roles for small RNAs in bacteria. *Curr Opin Microbiol* 6: 120-124.
- 39 Repoila F, Darfeuille F (2009) Small regulatory non-coding RNAs in bacteria: physiology and mechanistic aspects *Biol Cell* 101: 117-131.
- 40 Jäger D, Pernitzsch SR, Richter AS, Backofen R, Sharma CM, et al. (2012) An archaeal sRNA targeting cis- and trans-encoded mRNAs via two distinct domains. *Nucleic Acids Res* 40: 10964-10979.
- 41 Babski J, Maier LK, Heyer R, Jaschinski K, Prasse D, et al. (2014) Small regulatory RNAs in Archaea. *RNA Biol* 11: 484-493.
- 42 Waters LS, Storz G (2009) Regulatory RNAs in bacteria. *Cell* 136: 615-628.
- 43 Oliva G, Sahr T, Buchrieser C (2015) Small RNAs, 5'UTR elements and RNA-binding proteins in intracellular bacteria: impact on metabolism and virulence. *FEMS Microbiol Rev* 39: 331-349.
- 44 Wagner EG, Simons RW (1994) Antisense RNA control in bacteria, phages, and plasmids. *Annu Rev Microbiol* 48: 713-42.
- 45 Brantl S (2007) Regulatory mechanisms employed by cis-encoded antisense RNAs. *Current Opinion In Microbiology* 10: 102-109.
- 46 Sittka A, Lucchini S, Papenfort K, Sharma CM, Rolle K, et al. (2008) Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator Hfq. *PLoS Genet* 22: 8-e1000163.
- 47 Lorenz R, Bernhart SH, Siederdisen CH, Tafer H, Flamm C, et al. (2011) ViennaRNA Package 2.0. *Algorithms Mol Biol* 6: 26.
- 48 Rossi CC, Bossé JT, Yanwen Li, Witney AA, Gould KA, et al. (2016) A computational strategy for the search of regulatory small RNAs in *Actinobacillus pleuropneumoniae*. *RNA* 22: 1373-1385.
- 49 Wen Y, Feng J, Sachs G (2013) *Helicobacter pylori* 5'ureB-sRNA, a cis-encoded antisense small RNA, negatively regulates ureAB expression by transcription termination. *J Bacteriol* 195: 444-452.
- 50 Han Y, Liu L, Fang N, Yang R, Zhou D (2013) Regulation of pathogenicity by noncoding RNAs in bacteria. *Future. Microbiol* 8: 579-591.
- 51 Gonzalo-Asensio J, Ortega AD, Rico-Pérez G, Pucciarelli MG, García-Del Portillo F (2013) A novel antisense RNA from the *Salmonella* virulence plasmid pSLT expressed by non-growing bacteria inside eukaryotic cells. *PLoS One* 8: e77939.
- 52 Lee EJ, Groisman EA (2010) An antisense RNA that governs the expression kinetics of a multifunctional virulence gene. *Mol Microbiol* 76: 1020-1033.
- 53 Dadzie I, Xu S, Ni B, Zhang X, Zhang H, et al. (2013) Identification and characterization of a cis-encoded antisense RNA associated with the replication process of *Salmonella enterica* serovar Typhi. *PLoS One* 8: e61308.
- 54 Peng X, Dong H, Wu Q (2015) A new cis-encoded sRNA, BsrH, regulating the expression of hemH gene in *Brucella abortus*, 2308. *FEMS Microbiol Lett* 362: 1-7.
- 55 Camps M (2010) Modulation of ColE1-like plasmid replication for recombinant gene expression. *Recent Pat DNA Gene Seq* 4: 58-73.
- 56 Azhikina T, Ignatov DV, Salina EG, Fursov MV, Kaprelyants AS (2015) Role of Small Noncoding RNAs in Bacterial Metabolism. *Biochemistry* 80: 1633-1646.
- 57 Richards J, Belasco JG (2008) A new window onto translational repression by bacterial sRNAs. *Mol Cell* 32: 751-753.
- 58 Schroeder CL, Narra HP, Sahni A, Rojas M, Khanipov K, et al. (2016) Identification and Characterization of Novel Small RNAs in *Rickettsia prowazekii*. *Front Microbiol* 7: 1-14.
- 59 Richards GR, Vanderpool CK (2011) Molecular call and response: the physiology of bacterial small RNAs. *Biochim Biophys Acta* 1809: 525-531.
- 60 Cech GM, Szalewska-Pałasz A, Kubiak K, Malabirade A, Grange W, et al. (2016) The *Escherichia coli* Hfq Protein: An Unattended DNA-Transactions Regulator. *Front Mol Biosci* 3:1-6.
- 61 Link TM, Valentin-Hansen P, Brennan RG (2009) Structure of *Escherichia coli* Hfq bound to polyribadenylate RNA. *PNAS* 106: 19292-19297.
- 62 Holmqvist E, Wright PR, Li L, Bischler T, Barquist L, et al. (2016) Global RNA recognition patterns of post-transcriptional regulators Hfq and CsrA revealed by UV crosslinking in vivo. *EMBO J* 35: 991-1011.
- 63 Soper T, Mandin P, Majdalani N, Gottesman S, Woodson SA (2010) Positive regulation by small RNAs and the role of Hfq. *Proc Natl Acad Sci* 107: 9602-9607.
- 64 Chen S, Zhang A, Blyn LB, Storz G (2004) MicC, a second small-RNA

- regulator of Omp protein expression in *Escherichia coli*. *J Bacteriol* 186: 6689-6697.
- 65 Pfeiffer V, Papenfort K, Lucchini S, Hinton JC, Vogel J (2009) Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. *Nat Struct Mol Biol* 16: 840-846.
 - 66 Wroblewska Z, Olejniczak M (2016) Hfq assists small RNAs in binding to the coding sequence of ompD mRNA and in rearranging its structure. *RNA* 22: 979-994.
 - 67 Gong H, Gia-Phong V, Yong B, Elton C, Ruobin W, et al. (2001) A *Salmonella* small non-coding RNA facilitates bacterial invasion and intracellular replication by modulating the expression of virulence factors. *PLoS Pathog* 7: e1002120.
 - 68 Calderón PF, Morales EH, Acuña LG, Fuentes DN, Gil F, et al. (2014) The small RNA RyhB homologs from *Salmonella typhimurium* participate in the response to S-nitrosoglutathione-induced stress. *Biochem Biophys Res Commun* 450: 641-645.
 - 69 Leclerc JM, Dozois CM, Daigle F (2013) Role of the *Salmonella enterica* serovar Typhi Fur regulator and small RNAs RfrA and RfrB in iron homeostasis and interaction with host cells. *Microbiology* 159: 591- 602.
 - 70 Tattersall J, Rao GV, Runac J, Hackstadt T, Grieshaber SS, et al. (2012) Translation inhibition of the developmental cycle protein HctA by the small RNA IhtA is conserved across *Chlamydia*. *PLoS One* 7: e47439.
 - 71 Mellin JR, Goswami S, Grogan S, Brian Tjaden B, Genco CA (2007) Novel fur- and iron-regulated small RNA, NrrF, is required for indirect fur-mediated regulation of the sdhA and sdhC genes in *Neisseria meningitidis*. *J Bacteriol* 189: 3686-3694.
 - 72 Mellin JR, McClure R, Lopez D, Green O, Reinhard B, et al. (2010) Role of Hfq in iron-dependent and -independent gene regulation in *Neisseria meningitidis*. *Microbiology* 156: 2316-2326.
 - 73 Caswell CC, Gaines JM, Ciborowski P, Smith D, Borchers CH, et al. (2012) Identification of two small regulatory RNAs linked to virulence in *Brucella abortus* 2308. *Mol Microbiol* 85: 345-360.
 - 74 Solans L, Gonzalo-Asensio J, Sala C, Benjak A, Swapna U, et al. (2014) The PhoP-dependent ncRNA Mcr7 modulates the TAT secretion system in *Mycobacterium tuberculosis*. *PLoS Pathog* 10: e1004183.
 - 75 Loh E, Dussurget O, Gripenland J, Vaitkevicius K, Tiensuu T, et al. (2009) A trans-acting riboswitch controls expression of the virulence regulator PrfA in *Listeria monocytogenes*. *Cell* 139: 770-779.
 - 76 Smith NA, Eamens AL, Wang MB (2010) The presence of high-molecular-weight viral RNAs interferes with the detection of viral small RNAs. *RNA* 16: 1062-1067.
 - 77 Mironov AS, Gusarov I, Rafikov R, Lopez LE, Shatalin K, et al. (2002) Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria. *Cell* 11: 747-756.
 - 78 Henkin TM (2008) Riboswitch RNAs: using RNA to sense cellular metabolism. *Genes Dev* 22: 3383- 3390.
 - 79 Anthony PC, Perez CF, García-García C, Steven M Block (2012) Folding energy landscape of the thiamine pyrophosphate riboswitch aptamer. *PNAS* 2012 109: 1485-1489.
 - 80 Weinberg Z, Barrick JE, Yao Z, Roth A, Kim JN, et al. (2007) Identification of 22 candidate structured RNAs in bacteria using the CMfinder comparative genomics pipeline. *Nucleic Acids Res* 35:4809-4819.
 - 81 Barrick JE, Corbino KA, Winkler WC, Nahvi A, Mandal M, et al. (2004) New RNA motifs suggest an expanded scope for riboswitches in bacterial genetic control. *PNAS* 101: 6421-6426.
 - 82 Corbino KA1, Barrick JE, Lim J, Welz R, Tucker BJ, et al. (2005) Evidence for a second class of S-adenosylmethionine riboswitches and other regulatory RNA motifs in alpha-proteobacteria. *Genome Biol* 6: R70.
 - 83 Epshtein V, Mironov AS, Nudler E (2003) The riboswitch-mediated control of sulfur metabolism in bacteria. *Proc Natl Acad Sci* 100: 5052-5056.
 - 84 McDaniel BA, Grundy FJ, Artsimovitch I, Henkin TM (2003) Transcription termination control of the S box system: direct measurement of S-adenosylmethionine by the leader RNA. *Proc Natl Acad Sci* 100: 3083-3088.
 - 85 Lalaouna D, Simoneau-Roy M, Lafontaine D, Massé E (2013) Regulatory RNAs and target mRNA decay in prokaryotes. *Biochim Biophys Acta* 1829: 742-747.
 - 86 Suresh G, Srinivasan H, Nanda S, Priyakumar UD (2016) Ligand-Induced Stabilization of a Duplex-like Architecture Is Crucial for the Switching Mechanism of the SAM-III Riboswitch. *Biochemistry* 55: 3349-3360.
 - 87 Pérez AA, Rodionov DA, Bryant DA (2016) Identification and Regulation of Genes for Cobalamin Transport in the Cyanobacterium *Synechococcus* sp. Strain PCC 7002. *J Bacteriol* 198: 2753-2761.
 - 88 Gottesman S, Storz G (2011) Bacterial Small RNA Regulators: Versatile Roles and Rapidly Evolving Variations. *Cold Spring Harb Perspect Biol* 3:a003798.
 - 89 Trotochaud AE, Wassarman KM (2004) 6S RNA function enhances long-term cell survival. *J Bacteriol* 186: 4978-4985.
 - 90 Vakulskas CA, Leng Y, Abe H, Amaki T, Okayama A, et al. (2016) Antagonistic control of the turnover pathway for the global regulatory sRNA CsrB by the CsrA and CsrD proteins. *Nucleic Acids Res* 44: 7896-7910.
 - 91 Liu MY, Gui G, Wei B, Preston JF, Oakford L, et al. (1997) The RNA molecule CsrB binds to the global regulatory protein CsrA and antagonizes its activity in *Escherichia coli*. *J Biol Chem* 272: 17502-17510.
 - 92 Gutiérrez P, Li Y, Osborne MJ, Pomerantseva E, Liu Q, et al. (2005) Solution structure of the carbon storage regulator protein CsrA from *Escherichia coli*. *J Bacteriol* 187: 3496-3501.
 - 93 Steuten B, Hoch PG, Damm K, Schneider S, Köhler K, et al. (2014) Regulation of transcription by 6S RNAs: insights from the *Escherichia coli* and *Bacillus subtilis* model systems. *RNA Biol* 11: 508-521.
 - 94 Wehner S, Damm K, Hartmann RK, Marz M (2014) Dissemination of 6S RNA among bacteria. *RNA Biol* 11: 1467-1478.
 - 95 Oliveira IFF, Borges AS, Kooij V, Jentys JB, Luirink J, et al. (2010) Characterization of ftsZ mutations that render *Bacillus subtilis* resistant to MinC. *PLoS One* 5: e12048.
 - 96 Barrangou R, Marraffini LA (2014) CRISPR-Cas systems: Prokaryotes upgrade to adaptive immunity. *Mol Cell* 54: 234-244.
 - 97 Rath D, Amlinger L, Rath A, Lundgren M (2015) The CRISPR-Cas immune system: biology, mechanisms and applications. *Biochimie* 117: 119-28.
 - 98 Evguenieva-Hackenberg E, Klug G (2011) New aspects of RNA processing in prokaryotes. *Curr Opin Microbiol* 14: 587-92.
 - 99 Marraffini LA, Sontheimer EJ (2008) CRISPR interference limits

- horizontal gene transfer in staphylococci by targeting DNA. *Science* 322: 1843-1845.
- 100 Babu M, Beloglazova N, Flick R, Graham C, Skarina T, et al. (2011) A dual function of the CRISPR-Cas system in bacterial antiviral immunity and DNA repair. *Mol Microbiol* 79: 484-502.
- 101 Richter C, Chang JT, Fineran PC (2012) Function and regulation of clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR associated (Cas) systems. *Viruses* 4: 2291-311.
- 102 Sahr T, Brüggemann H, Jules M, Lomma M, Albert-Weissenberger C, et al. (2009) Two small ncRNAs jointly govern virulence and transmission in *Legionella pneumophila*. *Mol Microbiol* 72: 741-762.
- 103 Gunderson FF, Cianciotto NP (2013) The CRISPR-associated gene Cas2 of *Legionella pneumophila* is required for intracellular infection of amoebae. *M Bio* 4: e00074-13.
- 104 Gunderson FF, Mallama CA, Fairbairn SG, Cianciotto NP (2015) Nuclease activity of *Legionella pneumophila* Cas2 promotes intracellular infection of amoebal host cells. *Infect Immun* 83: 1008-18.
- 105 Toledo-Arana A, Dussurget O, Nikitas G, Sesto N, Guet-Revillet H, et al. (2009) The *Listeria* transcriptional landscape from saprophytism to virulence. *Nature* 459: 950-956.
- 106 Sesto N, Touchon M, Andrade JM, Kondo J, Rocha EPC, et al. (2014) PNPase dependent CRISPR System in *Listeria*. *PLoS Genet* 10:e1004065
- 107 Sampson TR, Saroj SD, Llewellyn AC, Tzeng YL (2013) Weiss DS A CRISPR/Cas system mediates bacterial innate immune evasion and virulence. *Nature* 497:254-257.
- 108 Marraffini AL (2016) The CRISPR-Cas system of *Streptococcus pyogenes* Function and applications. University of Oklahoma Health Sciences Centre.
- 109 Hooton SP, Brathwaite KJ, Connerton IF (2016) The Bacteriophage Carrier State of *Campylobacter jejuni* Features Changes in Host Non-coding RNAs and the Acquisition of New Host-derived CRISPR Spacer Sequences. *Front Microbiol* 7: 1-8.
- 110 Amitai G, Sorek R (2016) CRISPR-Cas adaptation: insights into the mechanism of action. *Nat Rev Microbiol* 14: 67-76.
- 111 Araújo NDD (2008) The effects and applications of bioinformatics on the biomedical area. *Estud Biol* 30: 143-148.
- 112 Binneck E (2004) The omics: integrating bioinformation Magazine. *Biotechnology*.
- 113 Altuvia S (2007) Identification of bacterial small non-coding RNAs; experimental approaches. *Curr Opin Microbiol* 10: 257-261.
- 114 Sridhar J, Narmada SR, Sabarinathan R, Ou HY, Deng Z, et al. (2010) sRNAscanner: a computational tool for intergenic small RNA detection in bacterial genomes. *PLoS One* 5: e11970.
- 115 Livny J, Fogel MA, Davis BM, Waldor MK (2005) sRNAPredict: an integrative computational approach to identify sRNAs in bacterial genomes. *Nucleic Acids Res* 33: 4096-4105.
- 116 Kingsford CL, Ayanbule K, Salzberg SL (2007) Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake. *Genome Biol* 8: R22.
- 117 Tran TT, Zhou F, Marshburn S, Stead M, Kushner SR, et al. (2009) De novo computational prediction of non-coding RNA genes in prokaryotic genomes. *Bioinformatics* 25: 2897-2905.
- 118 Wingender E, Chen X, Hehl R, Karas H, Liebich I, et al. (2000) TRANSFAC: an integrated system for gene expression regulation. *Nucleic Acids Res* 28: 316-319.
- 119 Sridhar J, Narmada SR, Sabiranathan R, Ou HY, Deng Z, et al. (2010) sRNAscanner: A computational tool for intergenic small RNA detection in bacterial genomes. *Plos One* 5: e11970.
- 120 Herbig A, Nieselt K (2011) nocoRNAc: characterization of non-coding RNAs in prokaryotes. *BMC Bioinformatics* 12: 1-13.
- 121 Sridhar J, Gunasekaran P (2013) Computational small RNA prediction in bacteria. *Bioinform Biol Insights* 7: 83-95.
- 122 Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, et al. (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* 28: 511-5.
- 123 Gautheret D, Lambert A (2001) Direct RNA motif definition and identification from multiple sequence alignments using secondary structure profiles. *J Mol Biol* 313: 1003-1011.
- 124 Nawrocki EP, Kolbe DL, Eddy SR (2009) Infernal 1.0: inference of RNA alignments. *Bioinformatics* 25: 1335-1337.
- 125 Coventry A, Kleitman DJ, Berger B (2004) MSARI: multiple sequence alignments for statistical detection of RNA secondary structure. *PNAS* 101: 12102-12107.
- 126 Pedersen JS, Bejerano G, Rosenbloom K, Lindblad-Toh K, Lander ES, et al. (2006) Identification and classification of conserved RNA secondary structures in the human genome. *PLoS Comput Biol* 4: e33.
- 127 Fallmann J, Will S, Engelhardt J, Grüning B, Backofen R, et al. (2017) Recent advances in RNA folding. *J Biotechnol* 10: 97-104.
- 128 Layton DM, Bundschuh R (2005) A statistical analysis of RNA folding algorithms through thermodynamic parameter perturbation. *Nucleic Acids Res* 33: 519-524.
- 129 Lorenz R, Bernhart SH, Siederdisen CHZ, Tafer H, Flamm C, et al. (2011) ViennaRNA Package 2.0. Algorithms. *Mol Biol* 6: 26.
- 130 Tjaden B, Goodwin SS, Opdyke JA, Guillier M, Fu DX, et al. (2006) Target prediction for small, noncoding RNAs in bacteria. *Nucleic Acids Res* 34: 2791-2802.
- 131 Tjaden B (2008) Target RNA: a tool for predicting targets of small RNA action in bacteria. *Nucleic Acids Res* 36: 109-113.
- 132 Cao S, Xu X, Chen SJ (2014) Predicting structure and stability for RNA complexes with intermolecular loop-loop base-pairing. *RNA* 20: 835-845.
- 133 Li W, Ying X, Lu Q, Chen L (2012) Predicting sRNAs and their targets in bacteria. *Genomics Proteomics Bioinformatics* 10: 276-284.
- 134 Kery MB, Feldman M, Livny J, Tjaden B (2014) TargetRNA2: identifying targets of small regulatory RNAs in bacteria. *Nucleic Acids Res* 42: 124-129.
- 135 Wright PR, Georg J, Mann M, Sorescu DA, Richter AS, et al. (2014) CopraRNA and IntaRNA: predicting small RNA targets, networks and interaction domains. *Nucleic Acids Res* 42: 119-123.
- 136 Busch A, Richter AS, Backofen R (2008) IntaRNA: efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. *Bioinformatics* 24: 2849-2856.
- 137 Mod SR, Camacho DM, Kohanski MA, Walker GC, Collins JJ (2011) Functional characterization of bacterial sRNAs using a network biology approach. *PNAS* 108: 15522-15527.

- 138 Clark PM, Loher P, Quann K, Brody J, Londin ER, et al. (2014) Argonaute CLIP-Seq reveals miRNA targetome diversity across tissue types. *Sci Rep* 4: 1-11.
- 139 Pánek J, Bobek J, Mikulík K, Basler M, Vohradský J (2008) Biocomputational prediction of small non-coding RNAs in *Streptomyces*. *BMC genomics* 9: 1-11.
- 140 Voss B, Georg J, Schön V, Ude S, Hess WR (2009) Biocomputational prediction of non-coding RNAs in model cyanobacteria. *BMC Genomics* 10: 1-15.
- 141 Brenes-Álvarez M, Olmedo-Verd E, Vioque A, Muro-Pastor AM (2016) Identification of conserved and potentially regulatory small RNAs in Heterocystous cyanobacteria. *Front Microbiol* 7: 1-11.
- 142 Khoo JS, Chai SF, Mohamed R, Nathan S, Firdaus-Raih M (2012) Computational discovery and RT-PCR validation of novel *Burkholderia* conserved and *Burkholderia pseudomallei* unique sRNAs. *BMC Genomics* 13: 1-14.
- 143 Ignatov D, Malakho S, Majorov K, Skvortsov T, Apt A, et al. (2013) RNA-seq analysis of *Mycobacterium avium* non-coding transcriptome. *Plos One* 8: e74209.