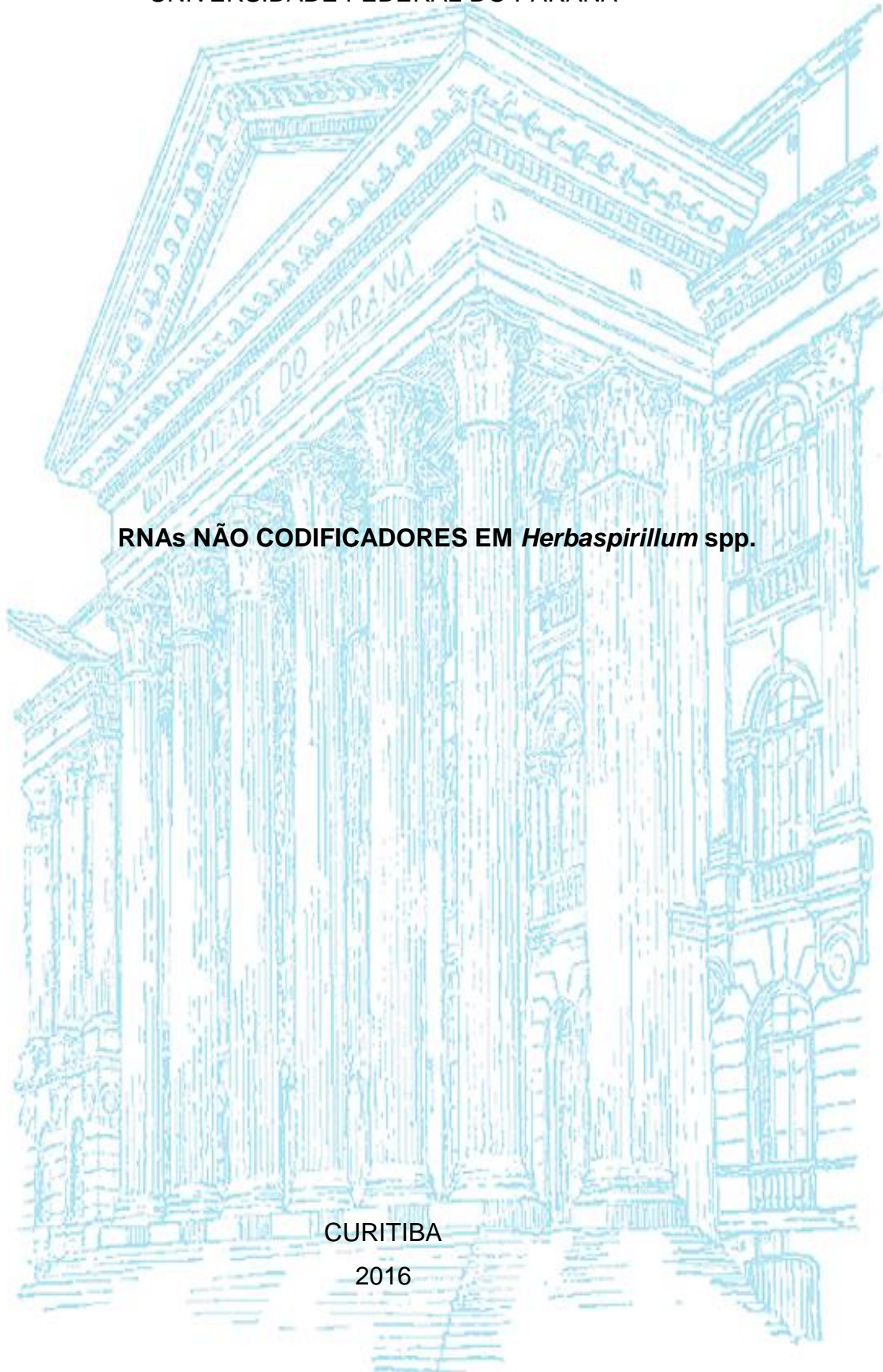


UNIVERSIDADE FEDERAL DO PARANÁ

RNAs NÃO CODIFICADORES EM *Herbaspirillum* spp.

CURITIBA

2016



AMANDA CARVALHO GARCIA

RNAs NÃO CODIFICADORES EM *Herbaspirillum* spp.

Dissertação apresentada como requisito parcial para a obtenção do grau de Mestre em Ciências - Bioquímica pelo Programa de Pós-Graduação em Ciências (Bioquímica) do Departamento de Bioquímica e Biologia Molecular, Setor de Ciências Biológicas da Universidade Federal do Paraná.

Orientador: Prof^a Dr^a Maria Berenice R. Steffens
Co-orientadora: Dr^a Michelle Zibetti Tadra Sfeir

CURITIBA

2016

Universidade Federal do Paraná
Sistema de Bibliotecas

Garcia, Amanda Carvalho
RNAs não codificadores em *Herbaspirillum* spp.. / Amanda Carvalho
Garcia. – Curitiba, 2016.
142 f.: il. ; 30cm.

Orientadora: Maria Berenice Reynaud Steffens
Co-orientadora: Michelle Zibetti Tadra Sfeir
Dissertação (Mestrado) – Universidade Federal do Paraná, Setor de
Ciências Biológicas. Programa de Pós-Graduação em Ciências - Bioquímica.

1. *Herbaspirillum*. 2. Riboswitch. 3. Ácido ribonucleico. I. Título II.
Steffens, Maria Berenice Reynaud. III. Sfeir, Michelle Zibetti Tadra. IV.
Universidade Federal do Paraná. Setor de Ciências Biológicas. Programa de
Pós-Graduação em Ciências - Bioquímica.

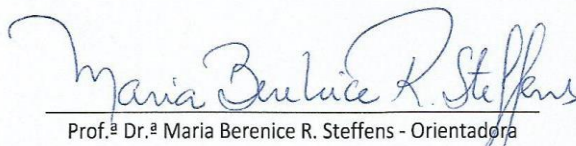
CDD (20. ed.) 574.192

TERMO DE APROVAÇÃO

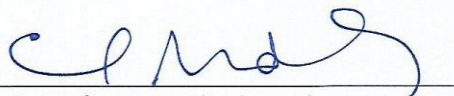
AMANDA CARVALHO GARCIA

RNAs não codificadores em *Herbaspirillum* sp

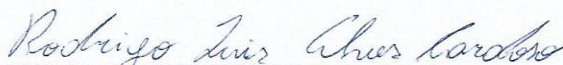
Dissertação aprovada como requisito parcial para obtenção do grau de Mestre no curso de Pós-Graduação em Ciências-Bioquímica, Setor de Ciências Biológicas, Universidade Federal do Paraná, pela seguinte banca examinadora:



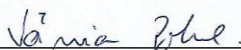
Prof.ª Dr.ª Maria Berenice R. Steffens - Orientadora
Departamento de Bioquímica, UFPR



Prof. Dr. Emanuel Maltempo de Souza
Departamento de Bioquímica, UFPR



Dr. Rodrigo Luis Alves Cardoso
Departamento de Bioinformática, UFPR



Dr.ª Vania Pobre
ITQB – Oeiras – Portugal

Curitiba, 29 de setembro de 2016.

*Dedico este trabalho ao Senhor Jesus Cristo, Maravilhoso, Conselheiro, Deus Forte,
Pai da Eternidade e Príncipe da Paz.*

Aos meus pais Mara Lúcia e Newton Cesar

Aos meus irmãos Manoela e Jair Cesar.

A minha afilhada e sobrinha Maria Eduarda.

Aos meus avós Ana Henrieta e Jair Batista.

A minha família pelo amor incondicional, por quem sou e por tudo o que alcancei.

AGRADECIMENTOS

A Deus pela maravilhosa e inestimável oportunidade de trabalhar com a Prof.^a Dr.^a Maria Berenice R. Steffens, orientadora dedicada a ensinar com temperância. Obrigada por me ensinar os passos da pesquisa científica, pelas discussões e pela amizade.

A Dr.^a Michelle Zibetti Tadra Sfeir pelo apoio, discussões, sugestões e coorientação.

Agradeço ao professor Emanuel M. de Souza pela sua orientação, discussões e ensinamentos fundamentais para minha formação.

Ao professor Fábio de Oliveira Pedrosa pela oportunidade de trabalhar no Núcleo de Fixação Biológica de Nitrogênio e aos demais professores do grupo pelos conselhos sábios.

A Dr.^a Vânia Sofia Pobre Instituto de Tectonologia Química e Biológica da Universidade Nova Lisboa (Oeiras - Portugal), pelos conselhos, discussões e amizade, e por fazer parte da banca externa.

Ao Dr. Rodrigo Alves Cardoso do Programa de Pós-graduação em Bioinformática – UFPR, pela contribuição na realização deste trabalho, pela amizade e por fazer parte da banca externa.

A Coordenação do Curso de Pós-Graduação em Bioquímica por apoiar a realização deste trabalho.

Ao Programa de Pós-Graduação em Bioinformática pela contribuição para a realização deste trabalho.

Agradeço a todos os colegas do grupo e principalmente aos amigos do laboratório 275 e 279 por toda ajuda, paciência, amizade e bons momentos compartilhados.

As agências financiadoras CAPES e CNPq, pela bolsa de estudo e pelo fomento.

“Os que semeiam em lágrimas segarão com alegria. Aquele que leva a preciosa semente, andando e chorando, voltará, sem dúvida, com alegria, trazendo *consigo* os seus molhos”.

Salmos 126:5 ao 6

“A sabedoria é a coisa principal; adquiere, *pois*, a sabedoria; sim, com tudo o que possuis, adquiere o conhecimento. Exalta-a, e ela te exaltará; e, abraçando-a tu, ela te honrará. Dará à tua cabeça um diadema de graça e uma coroa de Glória te entregará”.

Provérbios 4:7 ao 9

RESUMO

Herbaspirillum seropedicae SmR1 é uma bactéria diazotrófica que coloniza, de forma endofítica, várias plantas de interesse econômico. A identificação e caracterização de RNAs não codificadores (ncRNAs) no gênero *Herbaspirillum* é uma etapa importante para o estudo da interação dessas moléculas com mRNAs- ou proteínas-alvo, no processo de regulação pós-transcricional. Neste trabalho foi ampliada a análise estrutural e funcional de ncRNAs com função regulatória em *H. seropedicae* SmR1. Além disso, foram preditos ncRNAs em outras bactérias do gênero *Herbaspirillum*. Utilizando a ferramenta Infernal 1.1.1 foram identificados 55 novos ncRNAs *H. seropedicae* SmR1, com expressão confirmada, que foram somados aos 173 já relatados por Moreno (2013) e Cirino (2014). Os novos ncRNAs foram classificados de acordo com a sua estrutura, como ncRNA *cis-encoded* ou ncRNAs *trans-encoded*. Foram também encontrados *riboswitches* e um representante da nova classe de RNA, a sequência CRISPR. Utilizando a ferramenta TargetRNA2 foram preditos diversos mRNAs-alvos. Dez ncRNAs de *H. seropedicae* SmR1 com expressão confirmada em experimentos de RNAseq foram selecionados e oligonucleotídeos foram desenhados para futura validação por qRT-PCR. Finalmente, estes resultados contribuirão para o entendimento da participação desse tipo de RNA na regulação do metabolismo de bactérias do gênero *Herbaspirillum*.

Palavras-chave: *H. seropedicae* SmR1; RNA não codificador, ncRNA, *cis-encoded* ncRNA, *trans-encoded* ncRNA, *riboswitch*, CRISPR.

ABSTRACT

Herbaspirillum seropedicae SmR1 is a diazotrophic bacterium that colonizes endophytically several plants of economic interest. The identification and characterization of non-coding RNAs (ncRNAs) in the genus *Herbaspirillum* is an important step in the study of the interaction of these molecules with mRNAs- or proteins- targets, in the post-transcriptional regulation process. In this work we continued the study of the structural and functional analysis of ncRNAs with regulatory function in *H. seropedicae* SmR1. In addition, new ncRNAs were predicted in other bacteria of the genus *Herbaspirillum*. Using the Infernal 1.1.1 tool, 55 new ncRNAs, with confirmed expression, were identified in the *H. seropedicae* SmR1 genome. They were added to the 173 ncRNAs already reported by Moreno (2013) and Cirino (2014). The new ncRNAs were classified as cis-encoded ncRNAs or trans-encoded ncRNAs. We also found riboswitches and a representative of the new RNA class, the CRISPR sequence. Using the TargetRNA2 tool, several mRNAs-target were predicted. Ten ncRNAs of *H. seropedicae* SmR1, with confirmed expression in RNAseq experiments, were selected and oligonucleotides were designed for further validation by qRT-PCR. Finally, these results will contribute to the understanding of the participation of this type of RNA in the regulation of the metabolism of bacteria of the genus *Herbaspirillum*.

Key-words: *H. seropedicae* SmR1; RNA non-coding (ncRNA), ncRNA *cis_encoded*, *trans-encoded*, *riboswitches*, CRISPR, mRNA.

LISTA DE FIGURAS

FIGURA 1- SENTIDO DE TRANSCRIÇÃO ncRNAs E MODO DE PAREAMENTO COM mRNAs ALVOS.....	26
FIGURA 2- MECANISMOS DE AÇÃO <i>CIS-ENCODED</i> EM QUE, NO GENOMA, O ncRNA ESTÁ LOCALIZADO PRÓXIMO AO SEU mRNA ALVO.....	27
FIGURA 3- MECANISMOS DE AÇÃO <i>TRANS-ENCODED</i> EM QUE, NO GENOMA, O ncRNA ESTÁ LOCALIZADO DISTANTE DO SEU mRNA ALVO.....	27
FIGURA 4- LOCALIZAÇÃO <i>CIS-ENCODED</i> 5' TERMINAL ureB ncRNA.....	29
FIGURA 5 - MODELO PROVÁVEL DA INTERAÇÃO DA PROTEÍNA Hfq COM O PAR ncRNA/mRNA.....	31
FIGURA 6- EXEMPLOS DE <i>RIBOSWITCHES</i> E SEUS RESPECTIVOS LIGANTES EM PROCARIOTOS.....	33
FIGURA 7- ARRANJO ESTRUTURAL DO <i>RIBOSWITCH</i> E FUNÇÃO REGULATÓRIA DO LIGANTE.....	34
FIGURA 8- MODELO DE ARRANJO DO SISTEMA CRISPR/Cas.....	39
FIGURA 9- IDENTIDADE DE SEQUÊNCIA DOS 12 ncRNAs ANOTADOS NO <i>H. SEROPEDICAE</i> SmR1, EM OUTRAS ESPÉCIES DE BACTÉRIAS COM GENOMAS DISPONÍVEIS AO BANCO DE DADOS NCBI.....	56
FIGURA 10- VISUALIZAÇÃO DA POSIÇÃO E DO PERFIL DE EXPRESSÃO DO ncRNAs 5'_ureB_sRNA NO GENOMA <i>H. SEROPEDICAE</i> SmR1.....	58
FIGURA 11- VISUALIZAÇÃO DOS NÍVEIS DE EXPRESSÃO DO ncRNAs 5'_ureB_sRNA, EM ENSAIOS DE RNA-SEQ, EM <i>H. SEROPEDICAE</i> SmR1.....	59
FIGURA 12- VISUALIZAÇÃO DA POSIÇÃO E DO PERFIL DE EXPRESSÃO DO mRNA 5' UTR <i>cspA</i> NO GENOMA <i>H. SEROPEDICAE</i> SmR1.....	62

FIGURA 13- VISUALIZAÇÃO DOS NÍVEIS DE EXPRESSÃO DO mRNA 5' UTR <i>cspA</i> , EM ENSAIOS DE RNA-seq, EM <i>H. SEROPEDICAE</i> SmR1.....	63
FIGURA 14- VISUALIZAÇÃO DA POSIÇÃO E DO PERFIL DE EXPRESSÃO DO ncRNAs sX4 NO GENOMA <i>H. SEROPEDICAE</i> SmR1.....	64
FIGURA 15- VISUALIZAÇÃO DOS NÍVEIS DE EXPRESSÃO DO ncRNAs sX4, EM ENSAIOS DE RNA-seq, EM <i>H. SEROPEDICAE</i> SmR1.....	65
FIGURA 16- VISUALIZAÇÃO DA POSIÇÃO E DO PERFIL DE EXPRESSÃO DO ncRNAs COBALAMINA NO GENOMA <i>H. SEROPEDICAE</i> SmR1.....	66
FIGURA 17- VISUALIZAÇÃO DOS NÍVEIS DE EXPRESSÃO DO ncRNAs COBALAMINA, EM ENSAIOS DE RNA-seq, EM GENOMA <i>H. SEROPEDICAE</i> SmR1.....	67
FIGURA 18- VISUALIZAÇÃO DA POSIÇÃO E DO PERFIL DE EXPRESSÃO DO ncRNAs MicC NO GENOMA <i>H. SEROPEDICAE</i> SmR1.....	68
FIGURA 19- VISUALIZAÇÃO DOS NÍVEIS DE EXPRESSÃO DO ncRNAs MicC, EM ENSAIOS DE RNA-seq, EM GENOMA <i>H. SEROPEDICAE</i> SmR1.....	69
FIGURA 20- VISUALIZAÇÃO DA POSIÇÃO E DO PERFIL DE EXPRESSÃO DO ncRNAs IsrD NO GENOMA <i>H. SEROPEDICAE</i> SmR1.....	70
FIGURA 21- VISUALIZAÇÃO DOS NÍVEIS DE EXPRESSÃO DO ncRNAs IsrD, EM ENSAIOS DE RNA-seq, EM GENOMA <i>H. SEROPEDICAE</i> SmR1.....	71
FIGURA 22- VISUALIZAÇÃO DA POSIÇÃO E DO PERFIL DE EXPRESSÃO DO ncRNAs IsrG NO GENOMA <i>H. SEROPEDICAE</i> SmR1.....	72
FIGURA 23- VISUALIZAÇÃO DOS NÍVEIS DE EXPRESSÃO DO ncRNAs IsrG, EM ENSAIOS DE RNA-seq, EM GENOMA <i>H. SEROPEDICAE</i> SmR1.....	73
FIGURA 24- VISUALIZAÇÃO DA POSIÇÃO E DO PERFIL DE EXPRESSÃO DO ncRNAs 6S NO GENOMA <i>H. SEROPEDICAE</i> SmR1.....	74

FIGURA 25- VISUALIZAÇÃO DOS NÍVEIS DE EXPRESSÃO DO ncRNAs 6S, EM ENSAIOS DE RNA-seq, EM GENOMA <i>H. SEROPEDICAE</i> SmR1.....	75
FIGURA 26- VISUALIZAÇÃO DO GRÁFICO DE PONTOS E DA ESTRUTURA SECUNDÁRIA DO ncRNAs 6S.....	76
FIGURA 27- ALINHAMENTO MÚLTIPLO DE SEQUÊNCIAS DO GENE DO ncRNA 6S PELA FERRAMENTA <i>CLUSTAW</i>	77
FIGURA 28- VISUALIZAÇÃO DA INTERAÇÃO DO ALVOS ncRNA 4.5S COM O NOME DO GENE GENOMA DE <i>H. SEROPEDICAE</i> SmR1.....	80
FIGURA 29- VISUALIZAÇÃO DO LOCUS CRISPR NO GENOMA DE <i>H. SEROPEDICAE</i> SmR1.....	81
FIGURA 30- VISUALIZAÇÃO DA POSIÇÃO E DO PERFIL DE EXPRESSÃO DO ncRNAs CRISPR NO GENOMA <i>H. SEROPEDICAE</i> SmR1.....	82
FIGURA 31- VISUALIZAÇÃO DOS NÍVEIS DE EXPRESSÃO DO ncRNAs CRISPR, EM ENSAIOS DE RNA-seq, EM <i>H. SEROPEDICAE</i> SmR1.....	83

LISTA DE QUADROS

QUADRO 1- CLASSES DE RNAs E SUAS RESPECTIVAS FUNÇÕES.....	23
QUADRO 2- ncRNAs CODIFICADOS PELO GENOMA DE <i>H. SEROPEDICAE</i> SmR1 E ANOTADOS DE ACORDO COM O RFAM12.0.....	55
QUADRO 3- ncRNAs PREDITOS EM OUTRAS ESPÉCIES DE <i>HERBASPIRILLUM</i> PELA FERRAMENTA COMPUTACIONAL INFERNAL1.1.1.....	85
QUADRO 4- OLIGONUCLEOTÍDEOS DESENHADOS PARA PCR	86
QUADRO 5- OLIGONUCLEOTÍDEOS DESENHADOS PARA RT-PCR.....	87

LISTA DE TABELAS

TABELA 1 - GENES SINTÊNICOS AO 6S RNA.....	38
TABELA 2 - CARACTERÍSTICAS DO GENOMA DE <i>H. SEROPEDICAE</i> SmR1.....	44
TABELA 3 - ncRNAs DE <i>H. SEROPEDICAE</i> SmR1 SELECIONADOS PARA DESENHO DE PRIMERS.....	52

LISTA DE ABREVIATURAS

BLAST	- <i>Basic Local Alignment Search Tool</i>
CDS	- <i>coding sequence</i> (sequência que codifica para uma proteína)
CRISPR	- <i>Clustered Regulatory Interspaced Short Palindromic Repeats</i>
EMBL	- Laboratório de Biologia Molecular Europeu
FASTA	- Formato utilizado para armazenar sequências de bases e de aminoácidos em arquivo texto
GENBANK	- Banco de dados público do National Center for Biological Information, Institutos de Saúde dos Estados Unidos da América
gbK	- Extensão do arquivo que armazena sequência genética
IgR	- Região do genoma que não contém ORFs (Regiões Intergênicas)
mRNA	- RNA mensageiro
NCBI	- <i>National Center for Biotechnology Information</i>
ncRNA	- <i>no coding RNA</i> (RNA não codificador)
ORF	- <i>open reading frame</i> (fase aberta de leitura, possível gene)
PCR	- <i>polymerase chain reaction</i> (reação em cadeia da polimerase)
RBS	- <i>ribosome-binding site</i> (sítio de ligação de ribossomo)
RNase	- ribonuclease
RNA	- Acido Ribonucleico
RPKM	- <i>reads per kilobase of transcript per million mapped reads</i> (Leituras por kilobase de transcritos por milhão de leituras)
rRNA	- RNA ribossomal
SOLiD	- <i>Sequencing by Oligonucleotide Ligation and Detection</i> (Sequenciamento da ligação e detecção de Oligonucleotídeos)
sRNA	- RNA não codificador de proteína com função regulatória
tRNA	- RNA transportador
UTR	- <i>Untranslated Region</i> (Sequência de bases não traduzidas em um mRNA)

SUMÁRIO

1 INTRODUÇÃO	20
2 REVISÃO BIBLIOGRÁFICA	22
2.1 ÁCIDO RIBONUCLEICO RNA.....	22
2.1.1 RNAs NÃO CODIFICADORES EM PROCARIOTOS.....	24
2.1.2 <i>CIS-ENCODED</i> ncRNAs.....	28
2.1.3 <i>TRANS-ENCODED</i> ncRNAs.....	30
2.2 RIBOSWITCHES.....	33
2.3 ncRNAs QUE MODIFICAM A ATIVIDADE DE PROTEÍNAS.....	35
2.4 CRISPR.....	38
2.5 PREDIÇÃO DE ncRNAs E SEUS ALVOS POR BIOINFORMÁTICA.....	40
2.6 GÊNERO <i>HERBASPIRILLUM SEROPEDICAE</i> SmR1.....	43
3 OBJETIVOS	46
4 MATERIAIS E MÉTODOS	47
4.1 CONJUNTO DE DADOS.....	47
4.1.1 GENOMAS DE BACTÉRIAS DO GÊNERO <i>HERBASPIRILLUM</i>	47
4.2 SEQUÊNCIAS DE ncRNAs DE <i>HERBASPIRILLUM SEROPEDICAE</i> SmR1.....	47
4.3 DADOS DE RNA-Seq.....	47
4.4 ANOTAÇÃO DAS SEQUÊNCIAS DE ncRNAs.....	47
4.5 FERRAMENTAS DE BIOINFORMÁTICA.....	48
4.5.1 ARTEMIS.....	48
4.5.2 BLAST.....	48
4.5.3 <i>CLUSTAW</i>	49
4.5.4 INFERNAL 1.1.1.....	49
4.5.5 TargetRNA2.....	50

4.5.6 RNAFold.....	50
4.5.7 SamTools.....	51
4.6 ANÁLISE DO TRANSCRIPTOMA.....	51
4.7 INFERÊNCIA DE ALVOS ncRNAs DE <i>H. SEROPEDICAE</i> SmR1.....	52
4.8 DESENHO DE PRIMERS PARA VALIDAÇÃO FUNCIONAL DE ncRNAs.....	52
4.9 CRISPRmap – CONSERVAÇÃO CRISPR E CRISPRFinder.....	53
5. RESULTADOS E DISCUSSÃO.....	54
5.1 ANOTAÇÃO E CLASSIFICAÇÃO DOS ncRNAs DE <i>H. SEROPEDICAE</i> SmR1...54	54
5.2 PESQUISA DE IDENTIDADE DE SEQUÊNCIAS DE ncRNAs EM BACTÉRIAS DO GÊNERO <i>HERBASPIRILLUM</i>	56
5.3 PREDIÇÃO DE NOVOS ncRNAs NO GÊNERO <i>HERBASPIRILLUM</i>	57
5.3.1 ncRNAs 5'ureB_sRNAs.....	58
5.3.2 mRNA 5'UTR <i>cspA</i>	61
5.3.3 ncRNAs sX4.....	64
5.3.4 <i>Riboswirsch Cobalamin</i>	66
5.3.5 ncRNAs MicC.....	68
5.3.6 ncRNAs IsrD e IsrG.....	70
5.3.7 ncRNA 6S <i>housekeeping</i>	74
5.4 PREDIÇÃO DE ALVOS DOS ncRNAs EM <i>HERBASPIRILLUM SEROPEDICAE</i> SmR1.....	78
5.5 PREDIÇÃO DE CRISPR EM <i>HERBASPIRILLUM SEROPEDICAE</i> SmR1.....	81
5.6 PREDIÇÃO DE ncRNAs EM OUTRAS ESPÉCIES DE <i>HERBASPIRILLUM</i>	84
5.7 SELEÇÃO DE ncRNAs PARA VALIDAÇÃO E DESENHOS DOS PRIMERS PARA PCR E RT-PCR.....	86
6 CONCLUSÃO.....	88

REFERÊNCIAS.....	89
ANEXO I- CONDIÇÕES DE CULTIVO EMPREGADOS NOS EXPERIMENTOS DE RNA-seq com <i>H. seropedicae</i> SmR1.....	117
ANEXO II- ESPÉCIES <i>HERBASPIRILLUM</i> COM CÓDIGO DE E DATA DE ACESSO.....	118
ANEXO III- ncRNAs PREDITOS POR MORENO (2013) E CIRINO (2014) EM <i>H. seropedicae</i> SmR1.....	119
APÊNDICE I- SÊQUENCIA DOS ncRNAs PREDITOS PELA FERRAMENTA INFERNAL 1.1.1 <i>H. seropedicae</i> SmR1.....	121
APÊNDICE II- NÚMERO DE mRNAs ALVOS PREDITOS PARA OS 173 ncRNAs DE <i>H. seropedicae</i> SmR1.....	127
APÊNDICE III- Os 57 NOVOS ncRNAs PREDITOS EM <i>H. seropedicae</i> SmR1.....	131

1. INTRODUÇÃO

Herbaspirillum seropedicae é uma bactéria endofítica, pertencente ao grupo das beta-proteobactérias, com capacidade de fixar nitrogênio que contribui para o crescimento de gramíneas de importância econômica (NOINDORF et al., 2011; BATISTA et al., 2015). O genoma de *H. seropedicae* estirpe SmR1 foi completamente sequenciado e anotado pelo Programa GENOPAR (PEDROSA et al., 2011). Foi observado que esta bactéria possui 4,735 ORFs (*open reading frames*) em um único cromossomo, ocupando assim, cerca de 88,3% do genoma.

Esta anotação não determinou a presença de genes codificadores de RNAs com função regulatória, conhecidos como RNAs curtos (*small RNAs* ou sRNAs) ou RNAs não codificadores (*non-coding RNAs* ou ncRNAs). Estas moléculas modulam respostas fisiológicas através de diferentes mecanismos, através de interação RNA-RNA ou interação RNA-proteína. Algumas interações podem ser estabilizadas pela chaperona Hfq (PICHON e FELDEN, 2007; LIVNY e WALDOR, 2007; LIU e CAMILLI, 2010). É o que ocorre comumente na classe dos ncRNAs *trans-encoded* onde a formação do complexo ncRNA:Hfq:mRNA pode atuar positiva ou negativamente na regulação pós-transcricional (FANER e FEIG, 2013).

Os ncRNAs mais estudados são os *cis-encoded* RNAs e os *trans-encoded* RNAs, sendo os primeiros transcritos *in cis* e antisense em relação ao mRNA alvo e os segundos transcritos em regiões genômicas distantes do mRNA alvo (AZHIKINA et al., 2015). Por isso, somente os *cis-encoded* RNAs apresentam pareamento de bases perfeito com os alvos. Os *Riboswitches* constituem outra classe ncRNAs, localizados na região 5'UTR de um mRNA, que promovem a regulação transcricional através de sua interação com uma molécula ligante (WINKLER, 2005; MEYER et al., 2011; IZAR, MRAHEIL e HAIN, 2011). Recentemente, em procariotos, foram descritas regiões CRISPR (*Clustered Regularly Interspaced Short Palindromic Repeats*) que são repetições de sequências de bases palindrômicas (BRONS et al., 2008; AZHIKINA et al., 2015). Cada repetição consiste em segmentos curtos de "DNA espaçador" de exposições anteriores a um vírus bacteriófago ou plasmídeo exógeno. O sistema CRISPR/cas consiste em um sistema imune de resistência à moléculas exógenas (MARRAFFINI e SONTHEIMER; 2009).

Ferramentas computacionais são muito utilizadas para a predição de ncRNAs, como o QRNA (RIVAS et al., 2001) que realiza a análise comparativa de genomas, o ISI (PICHON e FELDEN, 2003) que analisa sequências intergênicas conservadas entre genomas relacionados para a identificação ncRNAs. O programa computacional RNAz 2.0 (GRUBER et al.; 2010) analisa a estabilidade termodinâmica da estrutura conservada do RNA e a existência de promotor e terminador nos possíveis preditos ncRNAs. As ferramentas de predição sRNAPredict2 (LIVNY et al., 2006) e SIPHT (LIVNY et al., 2008) avaliam as informações dos promotores terminadores Rho-independentes contra banco de dados. Além disso, a abordagem experimental de larga escala que envolve sequenciamento de cDNA (RNA-seq), tem sido amplamente utilizada na predição de ncRNAs e constitui um método bastante eficiente devido à possibilidade de confirmar a expressão de ncRNAs preditos (SHARMA e VOGEL, 2009; SHINHARA et al., 2011; PICHON et al., 2012; SCHROEDER et al., 2015).

Pesquisas realizadas por Moreno (2013) e Cirino (2014) no genoma de *H. seropedicae* SmR1 resultaram em centenas de ncRNAs preditos por ferramentas computacionais tais como: nocoRNAC, GSalgoritmo, Cufflinks, Rockhopper, TransTermHP. Foi utilizado o Blastn para alinhamento de sequências e para a predição de mRNA alvos, o TargetRNA. A expressão de 173 ncRNAs preditos foi confirmada por sequenciamento de RNA (RNA-seq) obtido em diferentes condições experimentais. Apenas oito ncRNAs, dentre os 173, foram anotados com base no banco de dados de famílias de RNA Rfam 12.0.

Este trabalho tem por objetivo a integração do conjunto de dados gerados pelos trabalhos de Moreno (2013) e Cirino (2014) e a ampliação da análise estrutural e funcional de ncRNAs com função regulatória em *H. seropedicae* SmR1 e em outras bactérias do gênero *Herbaspirillum*. Pretende-se assim contribuir para o entendimento da participação desse tipo de RNA na regulação do metabolismo bacteriano.

2. REVISÃO BIBLIOGRÁFICA

2.1 ÁCIDO RIBONUCLEICO RNA

Há cerca de um século os pesquisadores Phoebus A.T. Levene e Jean L.A. Brachet demonstraram, através de métodos experimentais, que o ácido nucléico de levedura ou ácido zimonucléico ou ácido nucléico pentose (PNA) era na verdade o ácido ribonucléico ou RNA (LEVENE, 1919; HOE et al., 2013; BROSIUS, RAABE; 2016). Os estudos de caracterização estrutural e funcional do RNA progrediram mais lentamente do que os estudos envolvendo DNA devido ao seu envolvimento em diversos processos celulares pouco conhecidos.

O RNA é um ácido ribonucléico devido à presença de uma ribose e é composto também pelas bases nitrogenadas adenina (A), citosina (C), guanina (G) e uracila (U), a qual substitui a base timina (T) presente na molécula de DNA. A molécula de RNA é constituída por polímeros lineares de nucleotídeos ligados entre si por ligações fosfodiéster 5' para 3' (COX et al., 2012). A cadeia de RNA é simples mas pode formar cadeia dupla devido pareamentos internos que levam a estruturas secundárias (TALES, 2013).

O avanço científico e tecnológico proporcionou a descoberta das funções daqueles tidos como principais RNAs: RNA mensageiro (mRNA), molécula de transferência da informação durante a síntese de proteínas; RNA ribossômico (rRNA) componente da estrutura da síntese de proteínas e RNA transportador (tRNA) com capacidade de transportar os aminoácidos e interagir com as proteínas (COX et al., 2012). Diversos estudos estruturais e funcionais do RNA têm sido realizados e possibilitaram a descrição de novas classes de RNA, com funções variadas e nos domínios *Archea*, *Bacteria* e *Eukaria* (QUADRO 1) (BROSIUS e RAABE, 2016).

Enquanto a genômica considera e analisa os genomas como sequências codificadoras de mRNAs, rRNAs e tRNAs, a RNômica investiga, além desses, os genes codificadores de RNAs que são não traduzidos, mas que são RNAs funcionais e estão envolvidos em diferentes processos celulares (HUNTTENHOFER et al., 2005; ULVÉ et al., 2007).

Tipo	Tipo	Tipo	Tipo	Função
Codificador	mRNA (mensageiro)	Presente	Presente	Síntese de proteínas
Não Codificador	rRNA (robossômico)	Presente 70S(50S+30S)	Presente 80S(60S+40S)	Síntese de Proteínas
	tRNA (transportador)	Presente	Presente	Transportador de aminoácido ao local de síntese proteínas
	snoRNA (RNA nucleolar)	Ausente	Presente	Maturação de ribossomos e RNAs transportadores
	snRNA (RNA nuclear)	Ausente	Presente	Síntese do mRNA eucariótico
	sRNA (RNA pequeno)	Presente	Presente	Modula a transcrição e a estabilidade de mRNAs
	CRISPR (RNA CRISPR)	Presente	Presente	Sistema de defesa bacteriana
	miRNA (micro RNA)	Ausente	Presente	Silenciamento gênico pela inibição da tradução e cliva mRNAs
	siRNA (RNA interferência)	Ausente	Presente	Silenciamento gênico
	piRNA (RNA liga-se PIWI)	Ausente	Presente	Silenciamento gênico em células germinativas
	lncRNA (RNA longo)	Ausente	Presente	Regulação biológica e processos de desenvolvimento de doenças

QUADRO 1 – Classes de RNAs e suas respectivas funções. Modificado de ZAHA et al., 2014.

A nomenclatura dos RNAs não codificadores é bastante diversa e foi variando ao longo tempo e de acordo com o avanço do conhecimento a respeito dessas moléculas (HUNTTENHOFER et al., 2005). Em bactérias, inicialmente, os RNAs funcionais que apresentavam entre ~ 50 a 300 nucleotídeos foram denominados RNAs curtos (do inglês *small RNAs*) (BRANTL, 2007). Com o passar do tempo o termo RNA não codificador (*no-coding RNA*) foi adotado e dividido para informar se apresenta ou não função regulatória (MIKULIK, 2003; BROSIUS e TIEDGE, 2004).

Atualmente, os RNAs não codificadores podem ser classificados como RNAs curtos não-mensageiros (snmRNAs, do inglês *small no-messenger RNAs*), RNAs curtos não codificadores (ncRNAs, do inglês *small no-coding RNAs*), RNAs não traduzidos (utRNAs, do inglês *untranslated RNAs*), ou RNAs não codificadores de proteínas (npcRNAs, do inglês *no-protein-coding RNAs*) (HOE et al., 2013).

2.1.1 RNAs NÃO CODIFICADORES EM PROCARIOTOS

Os RNAs não codificadores estão envolvidos em vários processos celulares, como por exemplo: replicação dos cromossomos na divisão celular (*diF*), regulação da transcrição (6S RNA), processamento do RNA (*mprB*), estabilidade do mRNA e tradução (sequência antisense – *spot42*), estabilidade de proteína (tmRNA) e transporte (4.5S ou *ffs*) (MIKULIK, 2003). Existem ncRNAs envolvidos no stress oxidativo (*oxyS*), na transcrição da fase estacionária (*dsrA*, *rprA*), relacionados ao controle do número de cópias de plasmídeos (RNAI, RNAIII), armazenamento de carbono (*csrBC*) e transporte de carbono (*gcvB*) (ULVÉ et al., 2007).

A classe dos ncRNAs com função regulatória está envolvida em diversos mecanismos regulatórios como na expressão gênica, na modulação das proteínas de superfície da membrana externa (Omp) (MASSÉ et al., 2003). Alguns ncRNAs podem ligar-se à proteínas a fim de modular suas atividades, como é o caso do 6S RNA que forma um complexo com a RNA polimerase (RNAP) (ULVÉ et al., 2007; PICHON, FELDEN, 2007; GOTTESMAN e STORZ, 2011).

Essas moléculas são transcritas *in trans* ou em *in cis* em relação ao RNA alvo (STORZ, ALTUVIA e WASSARMAN, 2005; LIU e CAMILLI, 2010). Os genes estão localizados entre as regiões codificadoras de proteínas, ou seja, nas regiões intergênicas do genoma e apresentam sinais de promotores e sequências terminadoras, geralmente Rho-independente (ERMOLAE et al., 2000; WASSARMAN et al., 2001; ARGAMAN et al., 2001; CIAMPI, 2006). O tamanho dos genes de ncRNAs varia entre ~50 até ~500 nucleotídeos e vários transcritos são processados por RNase com produtos finais menores (WASSARMAN e KILEY, 2010; ARNVIG, CORTES e YOUNG; 2014; UPDEGROVE et al., 2015; SCHROEDER et al., 2015).

O pareamento antisense entre o ncRNA regulatório e o RNA mensageiro alvo é o mecanismo de ação mais comum (SHARMA e VOGEL, 2009; MASSÉ e GOTTSMAN, 2002). Essa interação ocorre em regiões curtas e de complementariedade imperfeita de sequência, podendo ser estabilizada pela proteína chaperona Hfq (MASSÉ et al., 2003; REPOILA e DARFEUILLE, 2009; STORZ et al., 2011). Entretanto, Jager e colaboradores (2012) demonstraram que a interação ncRNA₁₆₂ / RNA alvo em *Archaea* pode ocorrer com ação *cis-encoded* ou *trans-encoded*, através de dois domínios distintos.

Assim como em *Bacteria*, os ncRNAs em *Archaea* estão envolvidos em muitos processos biológicos, tais como: regulação metabólica, adaptação a condições ambientais, resposta ao stress, regulação da morfologia e no comportamento celular. Em *Methanosarcina mazei* GO1, Babski e colaboradores (2014) identificaram ncRNAs que alinham na região 5'UTR dos mRNAs alvos e em *Sulfolobus solfataricus* ncRNAs que alinham na região 3'UTR dos mRNAs alvos. Já para *Pyrobaculum* sp. e *Haloferax volcanii*, há evidências que os ncRNAs podem alinhar com mRNAs alvos tanto na extremidade 5'UTR, quanto na extremidade 3'UTR. Além disso, observaram que o ncRNA TRF regula negativamente a tradução em resposta à ligação ao RBS.

O modo de ação dos ncRNAs regulatórios depende da co-localização com os seus mRNA alvos (FIGURA 1) (LI et al., 2012). Podendo ser classificados em *trans-encoded* RNA, codificados distantes dos seus mRNA alvos, e em *cis-encoded* quando localizam-se na região 5'UTR em relação ao alvo. Os *riboswitches*, em sua maioria, encontram-se na região 5'UTR, sofrem alterações conformacionais na sua estrutura secundária, devido a ligação de um ligante, e regulam a expressão gênica (GOTTESMAN e STORZ, 2011; ARNVIG et al., 2014).

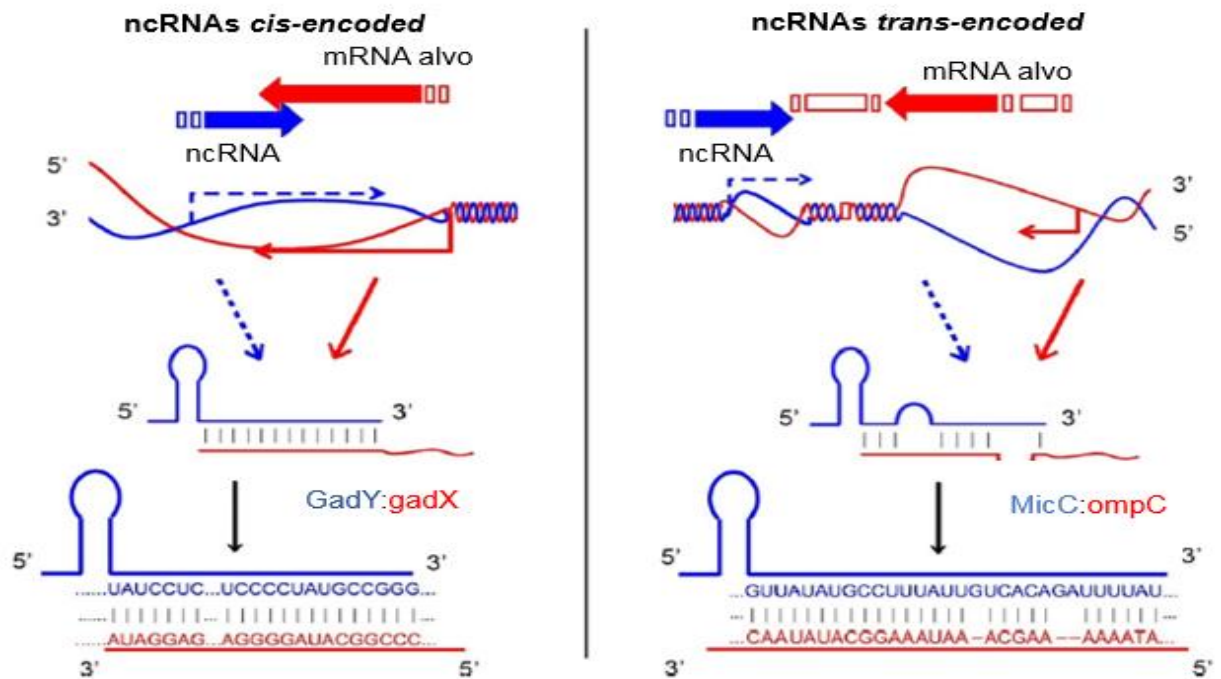


FIGURA 1 – Sentido de transcrição ncRNAs e modo de pareamento com RNAs alvos. Genes de ncRNAs *cis-encoded* estão localizados próximos aos genes codificadores de mRNAs alvos enquanto que genes de ncRNAs *trans-encoded* estão localizados distantes dos seus mRNAs alvos. Após a transcrição os ncRNAs *cis-encoded* fazem pareamento curto e perfeito com seus alvos e os ncRNAs *trans-encoded* fazem pareamento longo e imperfeito. Modificado de LI et al., 2012.

O mecanismo de ação dos ncRNAs varia de acordo com a sua função (FIGURAS 2 e 3). No passo pós-transcricional o ncRNAs *cis-encoded* pode causar a inibição da tradução (FIGURA 2a), a clivagem do mRNA alvo (FIGURA 2b) (*Overlapping* da 5'UTR) e terminação da transcrição (FIGURA. 2c) (WATERS e STORZ, 2009). A interação do ncRNA *trans-encoded* com mRNA alvo pode resultar na inibição da tradução (FIGURA 3a) e na degradação do mRNA pela interação negativa da 5'UTR inibindo o sítio de ligação do ribossoma, ou na degradação pela interação com RNase (FIGURA 3b). O ncRNA *trans-encoded* com o seu mRNA alvo pode formar uma estrutura inibitória que poderia bloquear de alguma forma o local do sítio de ligação do ribossoma, impedindo assim, a tradução (FIGURA 3.c) (WATERS e STORZ, 2009). Muitos RNAs codificados em *cis* com orientação antisense, podem formar complexos binários *trans-encoded* com o mRNA alvo, o que demonstra que esta classificação não é rígida ou decisiva (OLIVA, 2015).

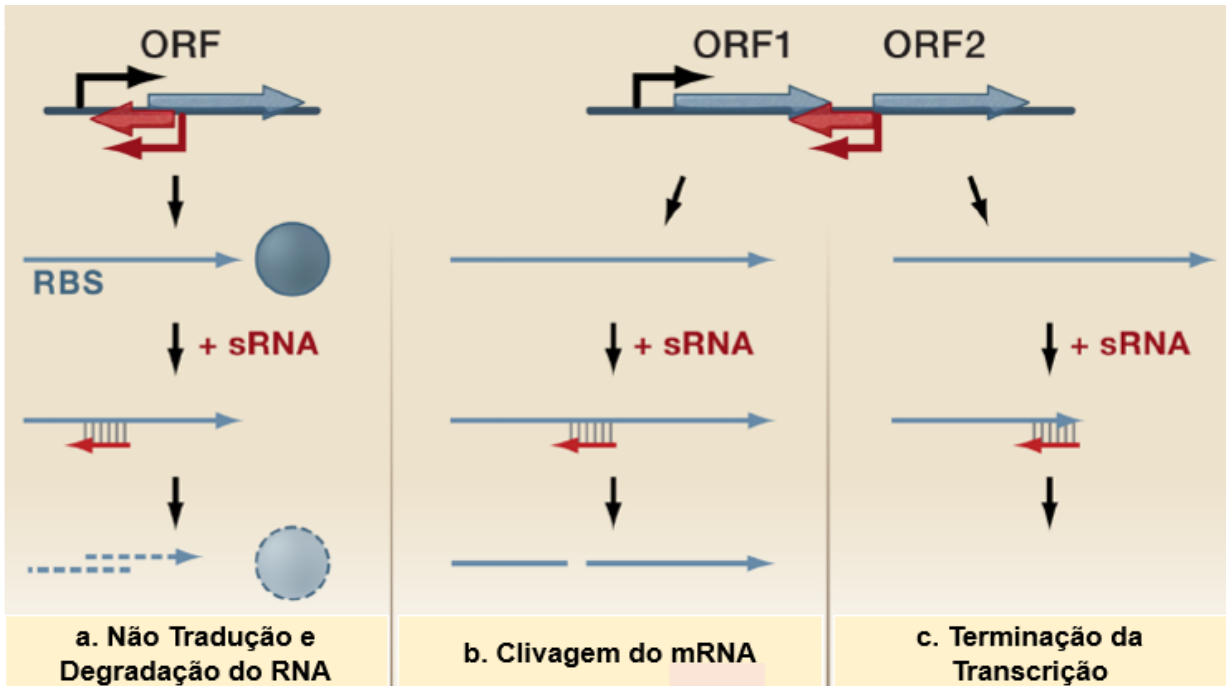


FIGURA 2 – Mecanismos de ação *cis-encoded* em que, no genoma, o ncRNA está localizado próximo ao gene do seu mRNA alvo. O sRNAs *cis-encoded* está representado em vermelho e seu alvo em azul. 2a. Não tradução e degradação do mRNA. 2b. Clivagem do mRNA e 2c. Terminação da transcrição. Modificado de WATERZ e STORZ, 2011.

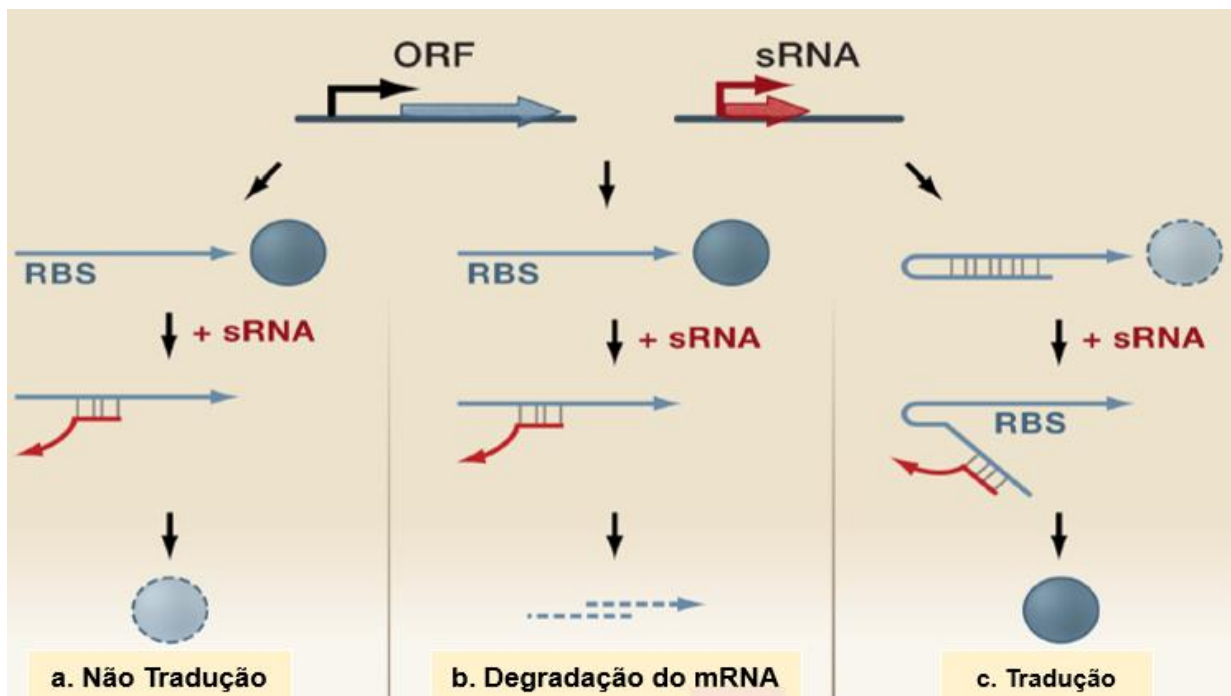


FIGURA 3 – Mecanismos de ação *trans-encoded* em que, no genoma, o ncRNA está localizado distante do gene do seu mRNA alvo. O ncRNAs *trans-encoded* está representado em vermelho e o seu mRNA alvo em azul. Complementariedade de pares de bases é limitada. 3a. Inibição da tradução. 3b. Degradação do mRNA. 3c. Permite a tradução. Modificado WATERZ e STORZ, 2011.

2.1.2 *cis-encoded* ncRNAs

Os *cis-encoded* ncRNAs estão presentes em diferentes espécies bacterianas, Gram-positivas e Gram-negativas (STORZ, ALTUVIA e WASSARMAN; 2005; OLIVA, SAHR e BUCHRIESER, 2015). Em *E. coli*, Wagner e Simons (1994) descreveram o papel regulatório antisense do RNA no controle da expressão gênica de mRNAs, na maturação de fagos, na transposição de elementos móveis e na replicação de plasmídeos. Além disso, podem participar da regulação do início da replicação, conjugação de plasmídeos, transposição, degradação do mRNA e também de algumas vias do metabolismo celular. Estes eventos continuam sendo objeto de investigação de diferentes pesquisadores (STORZ, ALTUVIA e WASSARMAN; 2005; BRANTL, 2007; LIU e CAMILLI, 2010; OLIVA, SAHR e BUCHRIESER, 2015).

Os *cis-encoded* ncRNAs são codificados no mesmo locus do seu mRNA alvo, mas no sentido antisense do duplex, sendo assim, totalmente complementares durante a interação. O mecanismo de resposta pós-transcricional da regulação gênica envolve um elevado grau de complementariedade de sequência, e isso foi considerado um indício de que não seria necessária a interação da proteína Hfq (HÜTTENHOFER, SCHATTNER e POLACEK, 2005; BRANTL, 2007). Entretanto, alguns pesquisadores reportaram a interferência de Hfq no pareamento ncRNA *cis-encoded*/mRNA alvo (SITTKA et al., 2008; LORENZ et al., 2010; ROSSI et al., 2016). De forma geral, esses ncRNAs atuam através do pareamento complementar ao sítio de ligação do ribossoma do mRNA, inibindo, por sua vez, a tradução (BRANTL, 2007; REPOILA e DARFEUILLE, 2009).

Um exemplo de *cis-encoded* ncRNA típico é o 5'*ureB* de *Helicobacter pylori*, localizado 5'-antisense ao gene *ureB* que compõe o operon *ureAB* (WEN, FENG e SACHS, 2013; HAN et al., 2013). Esse ncRNA contém 292 pb e regula negativamente a expressão do operon *ureAB* ao bloquear a tradução na porção 5' do *ureB* (FIGURA 4). Os genes *ureAB* de *H. pylori* localizam-se no cluster de dois operons *ureAB-ureIEFGH* e codificam as subunidades UreA e UreB da apoenzima urease. Essa enzima é essencial para a sobrevivência da *H. pylori* em baixo pH uma vez que sua reação produz NH₃ e HCO₃ para o ambiente, permitindo assim, a homeostasia para o crescimento bacteriano (WEN, FENG e SACHS, 2013).

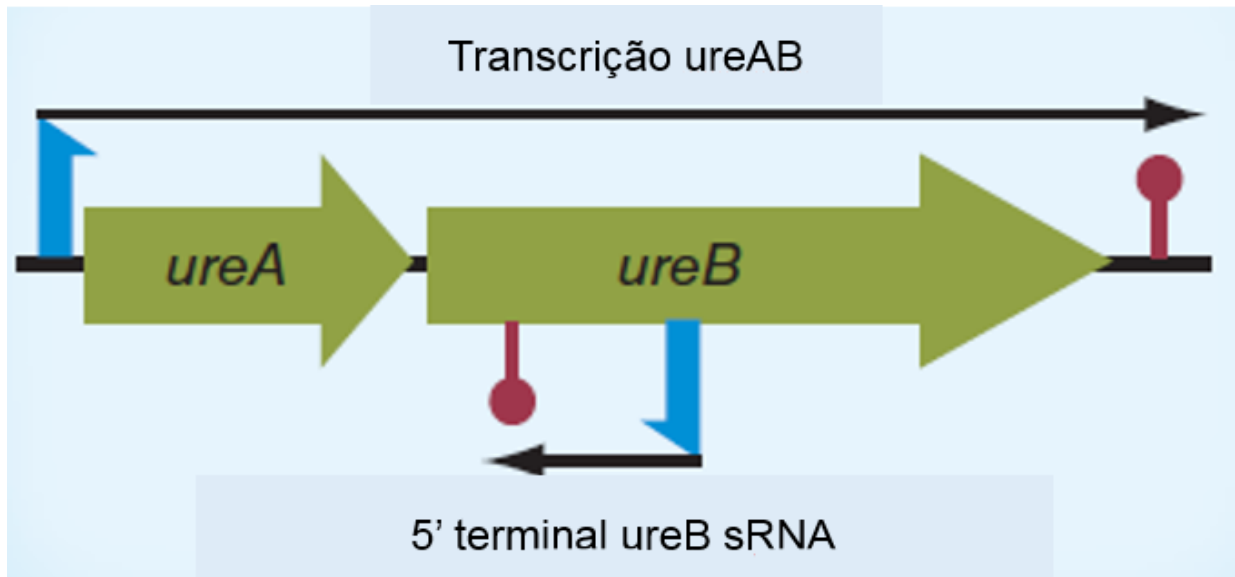


FIGURA 4 – Localização do *cis-encodado* 5' terminal ureB ncRNA. Esse ncRNA está localizado na região 5'-antisense do gene ureB que compõem o operon ureAB. Promotores representados em azul e terminadores tipo grampo representado em vermelho. Modificado de HAN et al., 2013.

A diversidade do *cis-encodado* ncRNA e seus papéis regulatórios variam de acordo com cada organismo como. Por exemplo, *Salmonella enterica serovar Typhimurium* possui o *cis-encodado* ncRNA lesR-1 cuja a função é controlar a replicação em células de eucarioto (GONZALO-ASENSIO et al., 2013) e *Salmonella entérica serovar Typi* possui o *cis-encodado* ncRNA AmgR envolvido na virulência em ratos (LEE e GROISMAN, 2010) e o AsdA que é regulador da replicação intracelular (DADZIE et al., 2013).

A mesma diversidade pode ser observada com relação às estratégias regulatórias. Em *Escherichia coli* o emparelhamento de bases entre o *cis-encodado* ncRNA GadY e seu mRNA alvo *gadXW* causa a clivagem do duplex entre *gadX* e *gadW*, proporcionando o aumento dos níveis do transcrito *gadX*. O produto GadX funciona como fator de transcrição do operon GadAGadB durante a síntese da glutamato descarboxilase e este processo consiste em um sistema de defesa contra estresse ácido em *E. coli* (WATERS e STORZ, 2009; AZHIKINA et al., 2015).

O sistema *SymR-SymE* em *E. coli* consiste em dois genes, o *cis-encodado* ncRNA *symR* e o gene *symE* que codificam uma proteína tóxica (REPOILA e DARFEUILLE, 2009; LIU e CAMILL, 2010; AZHIKINA et al., 2015). O aumento da concentração celular da proteína SymE diminui a atividade sintética dos ribossomos. O *cis-encodado*

ncRNA *symR* regula negativamente a expressão do gene *symE* pela complementariedade de bases ncRNA/mRNA, resultando na inibição da tradução de mRNAs *symE* e na retomada da atividade sintética ribossomal (REPOILA e DARFEUILLE, 2009; WATERS e STORZ, 2009; HAN et al., 2013; AZHIKINA et al., 2015).

Peng e colaboradores (2015) identificaram, em *Brucella abortus* 2308, o *cis-encoded* ncRNA BsrH que regula positivamente a expressão do gene *hemH*, evidenciando, com isso, a importância da expressão regulatória que ncRNA BsrH exerce sobre seu mRNA alvo.

Os *cis-encoded* ncRNA antisense são também comuns no mecanismo de replicação de plasmídeos. Por exemplo, no controle da replicação do plasmídeo ColE1, que usa ncRNA em vez de proteínas para iniciar a replicação no sítio de origem, são transcritos dois RNAs parcialmente complementares a partir de fitas opostas. O RNA maior, com 250-500 nucleotídeos (RNAII), é transcrito a partir da fita sense e forma um híbrido estável com o DNA molde. Este híbrido é então processado pela RNase H para gerar um primer para a DNA polimerase. O RNA menor, com 68-108 nucleotídeos (RNAI), é transcrito a partir da fita antisense e é complementar à região 5' do RNAII. Ele funciona como um regulador negativo da formação do primer pela formação do duplex RNAI/RNAII que impede a formação do híbrido RNAII/DNA molde. A concentração de RNAI é proporcional ao número de plasmídeos por célula, constituindo assim uma alça de *feedback* negativo que regula a replicação plasmidial em resposta às alterações metabólicas (MASSÉ, MAJDALANI e GOTTSEMAN, 2003; CAMPS, 2010).

2.1.3 *trans-encoded* ncRNAs

Como já mencionado, o modo de ação dos *trans-encoded* ncRNAs difere dos *cis-encoded* ncRNAs pelo limitado compartilhamento de pares de bases com os seus mRNA alvos (WATERS e STORZ, 2009; HAN et al., 2013; AZHIKINA et al., 2013). Esse tipo de ncRNA é codificado *in trans* e pode ter mRNAs alvos em diferentes locais no genoma. O ncRNA transcrito necessita geralmente da proteína chaperona Hfq para estabilizar a interação ncRNA-RNA alvo, devido ao pareamento de bases imperfeito e, assim, evita-se sua eventual degradação pela RNase (RICHARDS e BELASCO,

2008; STORZ et al., 2011, WATERS e STORZ, 2009). A chaperona mais estudada é a proteína Hfq que em *E. coli* interage com 40% dos ncRNAs (AZHIKINA et al., 2015). Schoroeder e colaboradores (2016) relatam que quase 50% de todas as espécies bacterianas possuem *trans-encoded* ncRNAs que necessitam da proteína chaperona Hfq, uma exceção é *Listeria monocytogenes* em que a maioria dos *trans-encoded* ncRNAs são independentes da Hfq.

A interação de Hfq com *trans-encoded* ncRNAs está envolvida na regulação pós-transcricional em várias espécies de bactérias, e pode exercer efeito negativo ou positivo sobre os seus mRNAs (RICHARDS, VARDERPOOL, 2011; CECH et al., 2016). A estrutura de Hfq, baseada em estudos de cristalografia, mostra uma proteína hexamérica, homóloga a proteínas Sm que possuem dois motivos (Sm1 e Sm2) (STORZ et al., 2011; PENG, DONG e WU; 2015). Em *E. coli*, Link e colaboradores (2009) caracterizaram dois sítios de ligação de RNA da proteína Hfq: proximal que se liga ao ncRNA e mRNA alvo e o outro distal que se liga a cauda poli U (FIGURA 5) (HOLMQVIST et al., 2016).

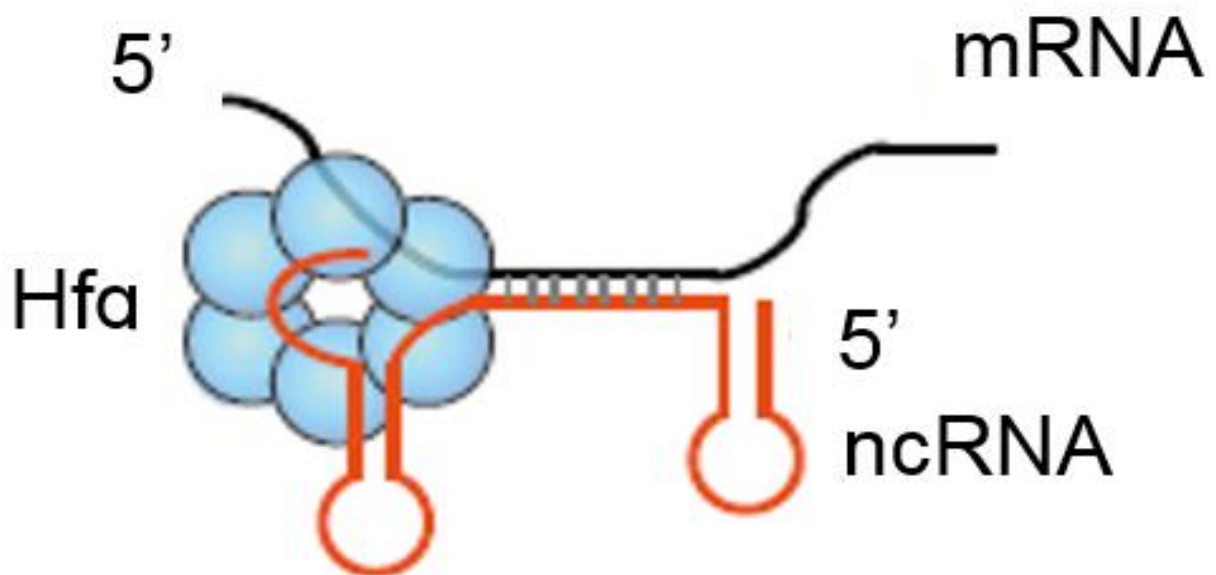


FIGURA 5 - Modelo provável da interação da proteína Hfq com o par ncRNA/mRNA. Modificado de HOLMQVIST et al., 2016.

Peng e colaboradores (2015) reportaram, através de estudos de transferência de energia por ressonância fluorescência (FRET), modelos estruturais em a Hfq interage com PAPI, PNPase e RNaseE. Soper e colaboradores (2010) descreveram três

ncRNAs regulatórios DsrA, RprA, ArcZ em *E. coli* que regulam positivamente a tradução do fator sigma RpoS quando ocorre o emparelhamento de bases ncRNA e mRNA *rpoS*. Eles identificaram a formação de um grampo inibitório na região 5'UTR e demonstraram que a ligação à Hfq é importante para a estabilidade do complexo RNA:RNA. Já na regulação negativa, o emparelhamento de bases ocorre entre a sequência de ligação ao ribossomo (RBS) e o ncRNA, bloqueando a ligação ao ribossomo, ou a degradação pela RNases.

Em *E. coli*, a regulação da expressão da proteína OmpC envolve o ncRNA MicF e uma região 5' UTR de 22 nucleotídeos antisense ao mRNA *ompC* (CHEN et al., 2004). Esta interação envolve também a Hfq e resulta na inibição da tradução. Já em *Salmonella typhimurium* é o ncRNA MicC, associado a proteína Hfq, que silencia o mRNA *ompD* através do duplex de RNA/RNA de 12 pares de bases na região codificadora de proteína. O MicC não inibe a iniciação da tradução na posição a jusante, mas acelera a atividade de RNaseE-dependente (PFEIFFER et al., 2009; WROBLEWSKA e OLEJINICZAK, 2016).

Em *Salmonella enterica serovar Typhimurium* o *trans*-encoded ncRNA IsrM, controla o fator patogênico SPI-1 (GONG et al., 2011). Já os RybB-1 e RybB-2 estão associados com a regulação da resposta ao stress oxidativo (CALDERON et al., 2014). Em *Salmonella enterica serovar Typi* os *trans*-encoded ncRNA RfrA e RfrB exercem função na regulação da homeostasia do ferro (LECLERC et al., 2013). Em *Chlamydia trachomatis* o *trans*-encoded ncRNA IhtA exerce a função de inibidor da proteína histona Hc1, já em *Neisseria meningitidis* o *trans*-encoded ncRNA Nrrf possui propriedade de regulação do homeostasia do ferro (TATTERSALL et al., 2012; MELLIN et al., 2007. MELLIN et al., 2010).

Outros exemplos de *trans*-encoded ncRNA: AbcR-1 e AbcR-2 em *Brucella abortus* estão relacionados com a virulência em ratos e são importantes para a sobrevivência de macrófagos (CASWELL et al., 2012) e Mrc7 na espécie *Mycobacterium tuberculosis* está envolvido na regulação do sistema de secreção da TAT (SOLANS et al., 2014).

2.2 RIBOSWITCHES

Riboswitches são elementos estruturados de RNA não codificante considerados elementos *cis-encoded* de RNA, localizados em sua maioria na região 5'UTR de um mRNA alvo e menos frequentes na extremidade 3' UTR (WINKLER, 2005; WATERZ e STORZ, 2009; HAN et al., 2013). Entretanto, Loh e colaboradores (2009) descreveram um caso de *riboswitch* que controla *in trans* a expressão da proteína reguladora da virulência PrfA em *Listeria monocytogenes*. Possuem a capacidade de controlar a expressão gênica a nível de transcrição e tradução e de adquirir diferentes conformações em resposta a sinais ambientais, como elevadas temperaturas e a ligação de pequenas moléculas, como metabólitos ou íons metálicos (SMITH et al., 2010; WATERS e STORZ, 2009; IZAR, MRAHEIL e HAIN, 2011). Recentemente, uma grande variedade de *riboswitches* foi identificada em procariotos. *Bacillus subtilis* possui 2% de todos os seus genes regulados por *riboswitches* que sem ligam a metabólitos intracelulares, como Flavina mononucleotídeo (FMN), pirofosfato de timina, S-adenosil-metionina (SAM), lisina e guanina (FIGURA 6) (MIRONOV et al., 2002; HENKIN, 2008; HAN et al., 2013; OLIVA, SAHR e BUCHRIESER, 2015).

Riboswitch	Ligante	Descrição
FMN	Flavina mononucleotídeo (FMN)	MIRONOV et al.; 2002; WINKLER et al.; 2002
THI box	Timina pirofosfato (TPP)	MIRONOV et al.; 2002; WINKLER et al.; 2002
B12	Adensilcobalamina	NAHVI et al.; 2002
S box (SAM-II)	S-adenosilmetionina (SAM)	EPSHEIN et al.; 2003; McDANIEL et al.; 2003; WINKLER et al.; 2003
SAM-III	S-adenosilmetionina (SAM)	CORBINO et al.; 2005
Smk box (SAM-III)	S-adenosilmetionina (SAM)	FUCHS et al.; 2006
SAH	S-adenosilhomocisteína (SAH)	WANG et al.; 2008
L box	Lsina	GRUNDY et al.; 2003
Glicina	Glicina	MANDAL et al.; 2004
Purina	Guanina e Adenina	MANDAL et al.; 2003; MANDAL , BREAKER et al.; 2004
dG	Deoxiguanosina	KIM et al.; 2004
Cíclico di-GMP	Cíclico di-GMP	SUDARSAN et al.; 2008
glmS	Glucosamina-6-fosfato	WINKLER et al.; 2004
preQ1	7-Aminoetil 7-deazaguanina	ROTH et al.; 2007
Mg	Magnésio	CROMIE et al.; 2006; DANN et al.; 2007

FIGURA 6 – Exemplos de *Riboswitches* e seus respectivos ligantes. Modificado de HENKIN, 2008.

A estrutura do *riboswitch* é composta por duas partes, a região de aptâmero que serve de sítio de ligação para um ligante e a plataforma de expressão que assume uma conformação adequada para a tradução (FIGURA 7) (WATERS e STORZ,

2009; ANTHONY et al., 2011). A ligação da molécula ligante provoca alterações conformacionais na estrutura nativa do *riboswitch*, podendo regular os processos de transcrição e tradução (WINKLER, 2005). A última sequência da plataforma de expressão forma um emperalhamento com o domínio aptâmero, dependendo se estiver ligado ou não ao ligante e pode, eventualmente, sinalizar um término de transcrição ou controlar a estrutura helicoidal no sítio de ligação do ribossoma (WEINBERG et al., 2011).

Assim os *riboswitches* são utilizados para regulação da terminação da transcrição (atenuação) do mRNA e a na iniciação da tradução (BARRICK et al., 2004). Na transcrição quando o ligante se liga ao mRNA na região do aptâmero, ocorrem alterações conformacionais, resultando na formação de um grupo alternativo (FIGURA 7a) (WATERS e STORZ, 2009). Este grupo atua como um terminador da transcrição que inibe a expressão gênica. A ligação da molécula ligante ao grupo alternativo leva à anti-terminação (FIGURA 7b). A ligação de um ligante que provoca alteração estrutural e leva ao sequestro do RBS impede a tradução (FIGURA 7c). Ao contrário, a ligação de um ligante pode causar a exposição do RBS e promover a tradução (FIGURA 7d).

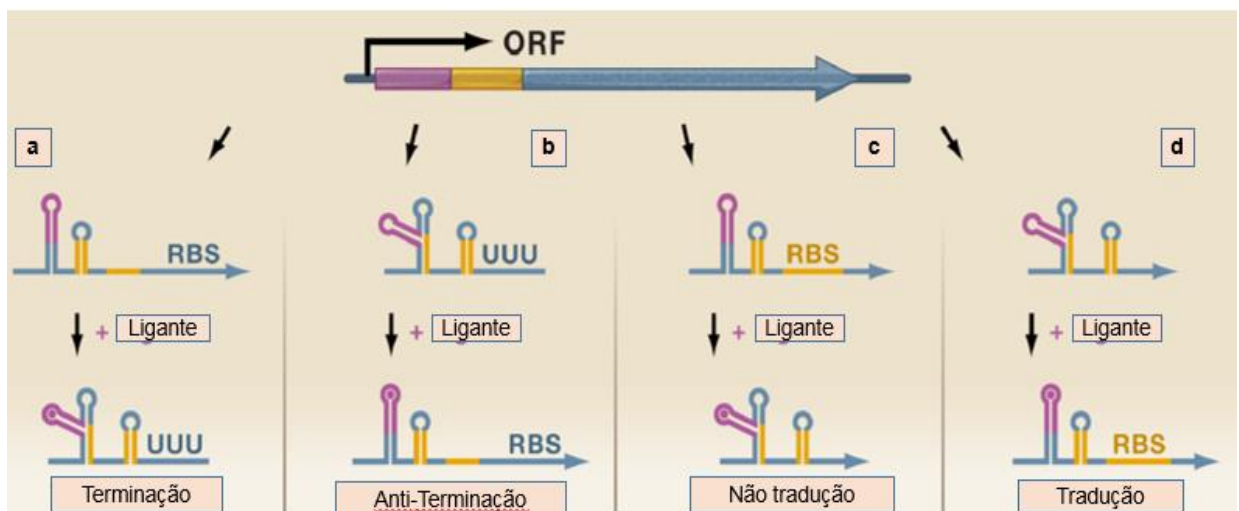


FIGURA 7 – Arranjo Estrutural do Riboswitch e Função Regulatória do Ligante. Riboswitch é formado pela região aptâmero (rosa) e uma plataforma de expressão (amarelo) na 5'UTR do respectivo mRNA (azul). Molécula ligante exerce função regulatória nos processos da transcrição (A e B) e da tradução. Modificado de WATERS e STORZ, 2009.

Os *riboswitches* podem desempenhar função em sistemas biológicos da célula, devido o grande número de famílias de genes envolvidos. Corbino e colaboradores (2005) descobriram motivos *metA* em *Agrobacterium tumefaciens* semelhante ao *riboswitch* dectetor de S-adenosilmetionina (SAM). A classe de *riboswitch* SAM-II apresenta grande diversidade estrutural, de baixa conservação, podendo alterar a estrutura conformacional do mRNA (EPSHEIN, MIRONOV e NUDLER, 2003; McDANIEL et al., 2003; WINKLER et al., 2003; LALAOUNA et al., 2013).

A estrutura do *riboswitch* pode conter diversos motivos que controlam a expressão do gene, através da detecção de concentrações dos metabólitos específicos, o que torna estas estruturas promissores alvo para antibióticos. Suresch e colaboradores (2016) demonstraram que o S-adenosilmetionina *riboswitch*-III, presente em bactérias anaeróbicas, está envolvido no processo de regulação da metionina e das vias biossintéticas SAM. Em presença ou ausência do ligante, o S-adenosilmetionina *riboswitch*-III exerce uma dupla função, o que facilita a mudança conformacional entre o estado parcial e totalmente dobrado, formando uma estrutura duplex estável, a qual fortalece as interações entre os nucleotídeos Shine-Dalgarno (SD) e anti-Shine-Dalgarno (aDS).

Pérez e colaboradores (2016) validaram, através de um sistema de expressão gênica *in vivo*, o cobalamina-*riboswitch* da cianobactéria *Synechococcus* sp. estirpe PCC 7002 e propuseram que a biossíntese de metionina é provavelmente o único meio de uso de cobalamina nesta estirpe de cianobactéria.

2.3 Os ncRNAs QUE MODIFICAM A ATIVIDADE DE PROTEÍNAS

Os primeiros ncRNAs caracterizados na enterobactéria *E. coli* incluem 4.5S, 6S, tmRNA (do inglês *transfer-messenger RNA*) e Spot42 (GOTTESMAN e STORZ, 2011; WATERS e STORZ, 2009). Em *E. coli*, o Spot42 RNA foi descoberto há quase 40 anos como uma molécula de RNA pequena e instável, codificada pelo gene *spf* presente também em *Vibrionaceae*, uma classe das γ -proteobactérias. *Aliivibrio salmonicida* codifica um ncRNA Spot42 RNA com 84% de identidade com o de *E. coli*. A deleção do gene *spf* resulta na diminuição de 25% da ação da RNA polimerase I (Pol I) (TROTOCHAUD e WASSARMAN, 2004). Este ncRNA otimiza a captação e o metabolismo do carbono ao ligar-se a mRNAs-alvo relacionados a este metabolismo, como o *galK* do operon da galactose. O ncRNA Spot42 é um exemplo de ncRNA que

atua tradicionalmente pela interação ncRNA/mRNA, mas existem casos, como o CsrB e o 6SRNA que regulam a atividade de proteínas-alvo e normalmente possuem sequências de reconhecimento específicas (STORZ et al., 2011).

O ncRNA CsrB de *E. coli* tem 22 sequências GGA consideradas sítios de ligação para a proteína CsrA (VAKULSKAS et al., 2016). Ao ligar-se à proteína ele regula a estabilidade do mRNA e a tradução. A proteína CsrA regula o armazenamento de carbono (LIU et al., 1997) através da ligação a mRNAs e é o componente essencial do sistema regulatório Csr. Esse sistema é responsável pela repressão de uma variedade de genes da fase estacionária, regulando negativamente a gluconeogênese, a biossíntese e catabolismo do glicogênio e a formação de biofilme. Além disso, atua via glicolítica, metabolismo de acetato e a biossíntese de flagelo. CsrA atua pós-transcricionalmente reprimindo a expressão gênica de enzimas essenciais para o metabolismo dos carboidratos como ADP-glucose pirofosforilase (*glgC*), glicogênio sintase (*glgA*), enzima de ramificação do glicogênio (*glgB*), glicogênio-fosforilase (*glgP*). Assim, ocorre a diminuição nos níveis intracelulares das enzimas biossintéticas de glicogênio e a diminuição intracelular da síntese de glicogênio (HOLMQVIST et al., 2016). CsrA desestabiliza seus mRNAs-alvo através da ligação na região codificadora de nucleotídeos -18 e +3, que inclui (RBS), impedindo a tradução do mRNA e promovendo sua degradação pela RNase. O ncRNA CsrB compete diretamente com o mRNA-alvo da CsrA e a proteína dimérica CsrA reconhece o motivo GGA da estrutura do ncRNA CsrB e o mRNA alvo. Os ncRNAs CsrC e CsrB em *E. coli* modulam a atividade de CsrA e a ligação do mRNA-alvo à proteína (STORZ et al., 2011; WATERZ e STORZ, 2011). Desta forma, os níveis intracelulares de CsrA também são regulados pelos ncRNAs CsrB e CsrC que atuam como antagonistas e capturam CsrA. Em *Erwinia* spp a proteína RsmA, homóloga da CsrA, regula vários genes envolvidos na doença de planta (GUTIÉRREZ et al., 2005). Em *Pseudomonas aeruginosa* Csr (Rms) controla vários sistemas como Rhl e vários fatores de virulência (GUTIÉRREZ et al., 2005).

O 6S RNA foi identificado por Hindley em 1967 em *Escherichia coli* (MIKULÍK, 2003). Desde então, 6S RNA foi predito através de ferramentas computacionais e experimentalmente analisado em diversas espécies bacterianas (STEUTEN et al., 2014). Wehner e colaboradores (2014) analisaram 1.611 genomas bacterianos e determinaram um conjunto de 1.750 genes de 6S RNA, sendo 1.367 novos. No banco de dados *Rfam*, o 6S RNA é descrito por duas entradas, a primeira RF00013 com 153 sequências, e a segunda RF01685 com 89 sequências.

Foi demonstrado em *E. coli* que o 6S RNA (gene *ssrS*) pode atuar no processo da transcrição ao associar-se a RNA polimerase dependente do fator sigma 70 (6S RNA:RNAP), modulando sua função. Esse ncRNA sequestra quase toda RNAP holoenzima na fase estacionária tardia e com isso contribui para adaptação transcricional do crescimento da fase estacionária (TROTOCHAUD e WASSARMAN, 2004; WENHER et al., 2014; STEUTEN et al., 2014).

Em relação à posição do 6S RNA no genoma, foi observado um padrão sintênico para o domínio *Bacteria*, sendo que este padrão é mais amplamente observado em gama-Proteobacteria. Dentre os genes sintênicos comuns estão: *ygfA* (5-formiltetrahydrofolato ciclo ligase), *zapA* (proteína ZapA), AAA (+) (superfamília de ATPases) e *peptidase_M24* (PF00557 família de metalopeptidases) (TABELA 1) (FREDERIC e LOSICH, 2002; WEHNER et al., 2014). Frederic e Losich (2002) identificaram a proteína ZapA em *Bacillus subtilis* junto da proteína tubulina FtsZ.

	PF01812	PF05164	PF00004	PF00557	
	#6S	5-FTHF_cyc-lig	ZapA	AAA(+)	Peptidase_M24
<i>Fusobacteria</i>	5			1	
<i>Firmicutes</i>	468	6		316	3
<i>Chloroflexi</i>	6				
<i>Cyanobacteria</i>	71			3	
<i>Actinobacteria</i>	7+1?			6	
<i>Spirochaetes</i>	27				3
<i>Planctomycetes</i>	5	1		1	
<i>Bacteroidetes</i>	62		33		2
<i>Chlorobi</i>	11		11		
<i>Aquificae</i>	8+2?	8			
<i>δ-/ε-Proteobacteria</i>	133	7	17	2	1
<i>Acidobacteria</i>	8		5		
<i>α-Proteobacteria</i>	214	29	147	10	3
<i>β-Proteobacteria</i>	127	1	125		
<i>γ-Proteobacteria</i>	431	283	389	13	319
<i>Chrysiogenetes</i>	1	1			
<i>Deferribacteres</i>	4	4		4	
<i>Synergistetes</i>	4			4	
Total		340	727	360	331

TABELA 1 – Genes sintênicos ao 6S RNA: *ygfA* (5-formiltetrahidrofolato ciclo ligase), *zapA* (proteína de divisão celular ZapA), AAA (+) (superfamília de ATPases) e peptidase_M24 (PF00557 família de metalopeptidases). O retângulo vermelho destaca as classes γ -Proteobacteria, β -Proteobacteria e α -Proteobacteria em que se observa maior sintonia entre 6S RNA e o gene *zapA*. Modificado de WEHNER et al., 2014.

2.4 CRISPR (*Clustered Regularly Interspaced Short Palindromic Repeats*)

Em 2007, Barrangou e colaboradores demonstraram, em células bacterianas, a alteração da resistência à infecção por fagos através da remoção ou adição de sequências espaçadoras similares às sequências dos fagos invasores. Essa estratégia definiu um sistema de defesa imune adaptativo que usa RNA curto e denominado *Clustered Regularly Interspaced Short Palindromic Repeats* (CRISPR). No genoma, o locus CRISPR contém centenas de sequências espaçadoras (~26-72 nucleotídeos) flanqueadas por repetições de 21 a 47 nucleotídeos. Geralmente estão associados a proteínas (*cas*), formando o sistema CRISPR\cas, que confere resistência mediada por RNA contra ácidos nucléicos provenientes, por exemplo, de bacteriófagos, plasmídeos ou elementos genéticos móveis (BRONS et al., 2008;

EVGUENIEVA-HACKENBERG e KLUG, 2011; OLIVA, SAHR e BUCHRIESER, 2015; HOE et al., 2013). O locus CRISPR está presente em ~40% dos genomas bacterianos e em 90% dos genomas de arqueias (MARRAFFINI e SONTIMMER, 2008; WATERZ e STORZ, 2011; BABU et al., 2011).

A maioria dos procariotos contendo CRISPR possui múltiplos grupos 2 a 20 loci, cada um organizado *in tandem* contendo até 100 repetições idênticas entre 25 a 50 pares de bases (BABU et al., 2011). As proteínas Cas fornecem a maquinaria enzimática necessária para aquisição de novos espaçadores e criando marcas de reconhecimento de elementos invasores (RICHTER et al, 2012).

O arranjo e o processamento dos RNAs CRISPR estão representados na (FIGURA 8) e envolve as fases (HOE et al., 2013):

Adaptação: o DNA invasor está integrado ao locus do CRISPR, resultando em novas repetições na matriz (FIGURA 8a);

Expressão: transcrição do transcrito que contém o RNA do fago (novo espaçador) e seu processamento pelas proteínas Cas para formar crRNA (matriz crRNA; composto de sequências únicas espaçadoras flanqueadas por repetições curtas) e proteínas codificadas Cas adjacente (FIGURA 8b);

Interferência: é a formação do complexo crRNAs / proteínas Cas e posteriormente, acontece a invasão ao DNA alvo e sua respectiva degradação (FIGURA 8.c).

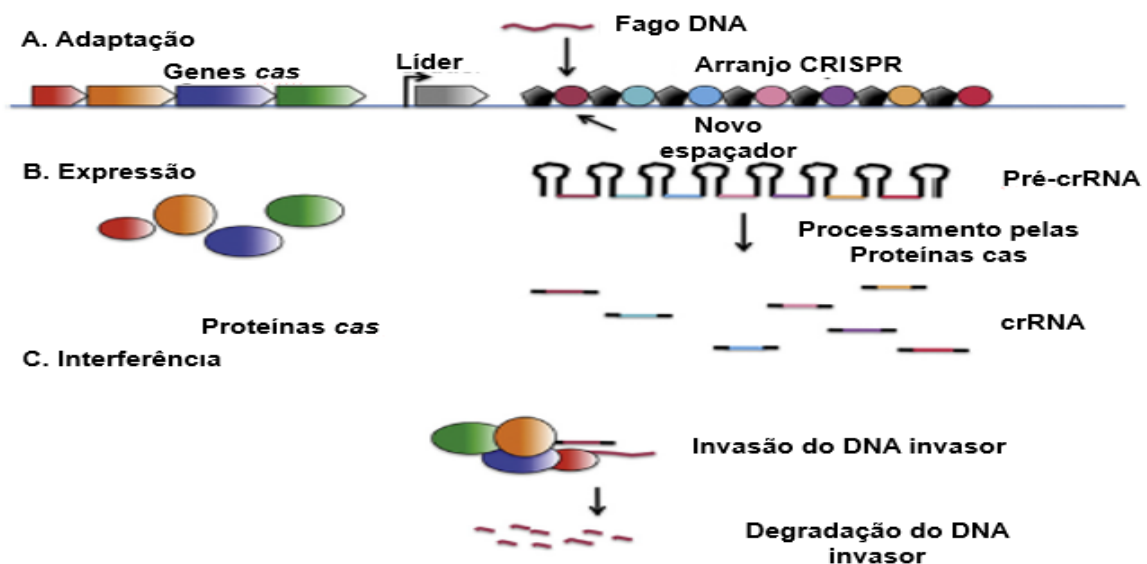


FIGURA 8 – Modelo de arranjo do sistema CRISPR\Cas. a. Adaptação: fago de DNA é integrado a matriz CRISPR e a formação do novo espaçador. b. Expressão: Processamento das proteínas cas. c. Interferência: invasão ao DNA alvo e sua respectiva degradação. Modificado de HOE et al., 2013.

O CRISPR pode ser expresso em espécies como *Legionella pneumophila*, onde o sRNA é o Cas2-dependente crRNA envolvido na resposta do stress (SAHR et al., 2009; GUNDERSON et al., CIANCOTTO et al., 2013). Em *Listeria monocytogenes* o sRNA RliB-CRISPR também faz uso do sistema CRISPR\Cas na resistência antiviral (TOLEDO et al., 2009; SESTO et al., 2014). *Francisella novicida* possui os ncRNAs, Cas9-dependente crRNA, tracrRNA e scaRNA, com propriedade CRISPR\Cas na regulação de fatores endógenos virais (SAMPSON et al., 2013). Segundo Marraffini (2016) a estirpe *Streptococcus pyogenes* possui o sistema CRISPR\Cas9 cuja função é intermediar a resistência antifagocítica. Na bactéria *Campylobacter jejuni* PT14 o sistema CRISPR\Cas14 protege-a da invasão/presença do bacteriófago (HOOTON et al., 2016).

Recentemente Amitai e Sorek (2016) estudaram os mecanismos envolvidos na adaptação de CRISPR-Cas, na percepção de que o DNA exógeno uma vez identificado pelo sistema CRISPR/Cas é integrado ao genoma da célula hospedeira, podendo assim, gerar memória imunológica, uma capacidade natural baseada na informação contida no DNA. Tal sistema pode ser manipulado adequadamente na formação de bases novas de armazenamento de informações em organismos vivos, algo que vem sendo investigado nos últimos anos.

2.5 PREDIÇÃO DE ncRNAs E SEUS ALVOS

A bioinformática dedica-se ao desenvolvimento de programas computacionais para o tratamento de dados biológicos e a identificação de sequências de genes, a predição configuração tridimensional de proteínas, identificar inibidores enzimáticos, promover o agrupamento de proteínas, para estabelecer árvores filogenéticas e analisar experimentos da expressão gênica (ARAÚJO et al., 2008). Fornece ferramentas para o desenvolvimento da Genômica, Transcriptômica, Proteômica e Metabolômica (BINNECK, 2004).

Em relação a predição de ncRNAs, várias abordagens computacionais tem sido desenvolvidas para a identificação de genes em regiões intergênicas de genomas procarióticos (ALTUVIA, 2007; SRIDHAR et al., 2010). Muitas delas baseiam-se na busca de sinais transcrpcionais, sequências promotoras conservadas, terminador rho-

independente, sítio de fator de transcrição, como sRNAPredict (LIVNY et al., 2005), análise algorítmica de predição putativa utilizando base de dados da TransTermHP (KINGSFORD et al., 2007; TRAN et al., 2009) ou TRANSFARC (WINGENDER et al., 2000). sRNAPredict3 e SIPHT são versões computacionais recentes de predição de ncRNAs em bactérias (LIVNY et al., 2008). sRNAscanner e sRNAfinder (SRIDHAR et al., 2010) foram desenvolvidos para suprir a limitação da predição de sinais de transcrição disponíveis em todas as sequências genômicas e mostrou-se eficiente.

O nocoRNAc (non-coding RNA characterization) é uma ferramenta computacional desenvolvida para estudar as interações entre ncRNA-mRNA em conjunto com a predição de ncRNAs em genomas bacterianos. Esse programa utiliza sinais terminadores de transcrição preditos através da ferramenta TransTermHP, os promotores são identificados pelo modelo SIDD (*Stress Induced Duplex Destabilization*) e verifica regiões possíveis a serem desestabilizadas (HERBIG e NIESELT, 2011; SRIDHAR e GUNASEKARAN, 2013). Assim são identificadas regiões do genoma flanqueadas por sequência promotora e sequência terminadora Rho-independentes, candidatas a codificar ncRNAs. Em uma abordagem distinta, a ferramenta *Cufflinks* localiza regiões do genoma com consideráveis níveis de transcrição e livres de ORFs (TRAPNELL et al., 2010).

Análises comparativas genômicas têm sido aplicadas para predição de novos ncRNAs, e sequências conservadas identificadas pela primeira vez na região intergênica (IGR), compartilhadas pelo agrupamento e comparação dos alinhamentos múltiplos são classificadas como ncRNAs. Programas como QRNA (RIVAS et al., 2001), ERPIN (GAUTHERET e LAMBERT, 2001), ISI (PICHON e FELDEN, 2003), INFERNAL (NAWROCKI et al., 2009), MSARI (COVENTRY et al., 2004); RNAModif (MACKE et al., 2001) e RNAz (PEDERSEN et al., 2006) comparam estabilidade termodinâmica através da predição de estruturas conservadas de RNA estáveis, predizendo ncRNAs em bactérias (WASHIED et al., 2005; GRUBER et al., 2010).

No caso de estruturas secundárias, o RNAFold realiza análise estatística algorítmica do dobramento do RNA, através da perturbação dos parâmetros termodinâmicos, para sua predição (LAYTON e BUNDSHUH, 2005; LORENZ et al., 2011; SRIDHAR e GUNASEKARAN, 2013).

Com relação à predição de alvos, o desenvolvimento de modelos é muito importante, pois integra a bioinformática para a predição e a validação experimental para confirmação de mRNA alvos. A classificação dos ncRNAs fornece informações

sobre a complementariedade de bases (perfeita ou imperfeita) com seus mRNAs alvos e a eventual ligação à proteínas, alterando sua atividade (TJADEN et al., 2006; PICHON e FELDEN, 2008; TJADEN, 2008; CAO, XU, CHEN; 2014; LI *et al.*, 2012).

A ferramenta de predição de mRNA alvos TargetRNA2 é uma das mais utilizadas atualmente (KERY et al., 2014). Utiliza vários recursos incluindo a conservação de ncRNA em outras bactérias, a estrutura secundária tanto do ncRNA e de cada candidato alvo, e a energia de hibridização entre as interações. Possui a capacidade de integrar dados de material RNA-seq quando disponíveis. Outra abordagem computacional utilizada para predição de alvos mRNAs, a ferramenta *IntaRNA* (WRIGHT et al., 2014), também é considerada bastante eficiente e rápida na predição das interações entre ncRNA / RNAm. Utiliza a energia livre de hibridização e possui integração com a ferramenta CopraRNA, que prediz ncRNAs comparando a sequência-consulta com sequências disponíveis no programa (BUSCH et al., 2008; WRIGHT et al., 2014).

Outras abordagens incluem a detecção direta por meio de *microarrays*, *northern blotting*, (WASSARMAN et al., 2001; LIVNY e WALDOR, 2007; ALTUVIA, 2007; MODI et al., 2011) e *Clip-seq* (CLARK et al., 2014).

A seguir serão apresentados exemplos de predições de ncRNAs e seus alvos em diferentes microrganismos. Pánek e colaboradores (2008) identificaram preditos ncRNA em *Streptomyces* baseados no estudo da conservação de sequência em regiões intergênicas, localização do fator terminador de transcrição e arranjo genômico de genes sintênicos ao ncRNAs. Detectaram a expressão de 20 ncRNAs por microarrays e RT-PCR, utilizaram abordagem computacional para determinação de estrutura secundária e identificaram o ncRNA 6S. Voss e colaboradores (2009) utilizaram modelo de *cyanobacteria* para predição de ncRNAs a partir de dados de transcriptoma e proteoma e identificaram os Yfr2a-Yfr2c ncRNA, de estrutura conservada entre as cianobactérias. Para prever a existência do 5'-operon-líder, do 3'-terminador de transcrição Rho-independente e de *riboswitches*, esses autores utilizaram as ferramentas computacionais TransTermHP, ClustaW, RNAz, RNAfold e para validação utilizaram Northern Blot (BRENES-ÁVAREZ et al., 2016). Modi e colaboradores (2011) realizaram a caracterização funcional de ncRNAs em *E. coli* através de inferência de rede, baseados em um compêndio de perfis de expressão de genes com predição funcional e nas interações regulatórias dos ncRNAs. Esses autores validaram experimentalmente as funções atribuídas a três ncRNAs, IsrA e

GlmZ, envolvidos na resposta ao dano do DNA, e GcvB, envolvido na regulação da disponibilidade de aminoácidos. Khoo e colaboradores (2012) integraram vários métodos computacionais de predição e análise de ncRNAs e identificaram 29 em *Burkholderia pseudomallei* dentre os quais 8 foram considerados novos. Ignatov e colaboradores (2013) utilizaram material e análise transcriptômica para revelar que *Mycobacterium avium* e *M. tuberculosis* contém diferentes conjuntos de ncRNAs nas regiões intergênicas e sugeriram que essa característica pode ser a base das diferenças fisiológicas observadas entre as duas espécies. Schroeder e colaboradores (2015) observaram que no gênero *Rickettsia* os ncRNAs são os principais reguladores pós-transcricionais envolvidos em processos de virulência, sobrevivência, expressão de plasmídeos, metabolismo primário e secundário e presumidamente codificam ncRNA *trans*-encoded envolvidos na interação patógeno e hospedeiro.

2.6 GÊNERO *Herbaspirillum*

O gênero *Herbaspirillum* foi descrito por Baldani e colaboradores em 1986 e incluiu a espécie *seropedicae* com diversas estirpes isoladas de diferentes gramíneas (BALDANI et al., 1986). *H. seropedicae* se caracteriza por fixar nitrogênio (diazotrófica) e colonizar plantas. Uma segunda espécie diazotrófica, *Pseudomonas rubrisubalbicans* foi reclassificada com o nome de *Herbaspirillum rubrisubalbicans* e também apresenta capacidade de colonizar plantas (BALDANI et al., 1996). A partir da década de 1990, o gênero *Herbaspirillum* incluiu novas espécies: *H. autotrophicum* (syn. *Aquaspirillum autotrophicum*) e *H. huttiense* (syn. *Pseudomonas huttiense*) (DING e YAKOTA, 2004) *H. frisingense* (KIRCHHOF et al., 2001), *H. lusitanum* (VALVERDE et al., 2003), *H. cholorophenicum* (IM et al., 2004), *H. hiltnei* (ROTHBALLER et al., 2006), *H. rhizosphaerae* (JUNG et al., 2007), *H. aquaticum* ((DOBRITSA, REDDY e SAMADPOUR, 2010), *H. massiliense* (LAGIE et al., 2012), *H. canariense*, *H. aurantiacum* e *H. soli* (CARRO et al., 2012) e *H. seropedicae* AU14040 (MARQUES et al., 2015). A espécie *H. putei* (DING e YOKOTA, 2004) foi reclassificada como *H. huttiense* subespécie *putei* (DOBRITSA, REDDY e SAMADPOUR, 2010).

H. seropedicae pertence à classe Beta das Proteobacterias, é Gram-negativa e apresenta crescimento característico em meio de cultura semi-sólido e sem nitrogênio (BALDANI, 1984; BALDANI et al., 1986). É uma bactéria que coloniza espaços intracelulares e xilema de raízes de gramíneas como cana-de-açúcar e sorgo (JAMES et al., 1997; BALDANI et al., 1986) e espaços intracelulares em sementes e raízes de arroz (BALDANI et al., 1992; JAMES et al., 1997). Também pode ser isolada a partir de caules folhas de diversas gramíneas (BALDANI et al., 1992; DOBEREINER, BALDANI e BALDANI, 1995; OLIVARES *et al.*, 1996).

O genoma da estirpe SmR1 de *H. seropedicae* foi completamente sequenciado e anotado pelo consórcio GENOPAR (Genoma do Paraná). Foram anotadas 4.735 *Open Reading frames* (ORFs) que ocupam cerca de 88,3% do genoma, em um único cromossomo circular. Nesta anotação não foram considerados os genes de RNAs não codificadores, exceto os rRNAs e tRNAs. Uma nova anotação foi disponibilizada pelo NCBI em 2015. As características principais do genoma de *H. seropedicae* SmR1 estão listadas na TABELA 2.

TABELA 2: CARACTERÍSTICAS DO GENOMA DE *HERBASPIRILLUM SEROPEDICAE* SmR1

	Pedrosa et al 2011	NCBI (2016)
Tamanho em pares de bases	5.513.887	5.513.889
%GC	63,4	63,4
Total de genes	4.804	4.796
Total de CDS	4.735	4.687
rRNA operons	3	9
tRNAs	55	55
Outro RNA	Na	1
Pseudogenes	Na	44

Na – não anotado

Devido ao envolvimento dos ncRNAs na regulação de importantes vias metabólicas importantes, WILLEMANN (2010) e MORENO (2013) realizaram buscas por RNAs não codificadores no genoma de *H. seropedicae* SmR1, através de abordagem bioinformática e análise de seqüenciamento de RNA. A aplicação da ferramenta integradora sRNATool permitiu identificar 200 seqüências candidatas no genoma de *H. seropedicae* SmR1 (WILLEMANN, 2010). Um conjunto de 98 ncRNAs

preditos foram confirmados a partir de dados obtidos em três condições experimentais (cultivo controle, cultivo na presença de naringenina e bactéria associado à planta) foi confirmado por RNA-seq (MORENO, 2013). Este número foi ampliado em 75 novos ncRNAs por CIRINO (2014) que usou os mesmos critérios propostos por MORENO (2013) para analisar dados de transcriptoma (RNA-Seq) de *H. seropedicae* SmR1 cultivado em 11 condições diferentes (ANEXO I).

Devido à importância regulatória dos ncRNAs não codificadores em *H. seropediace* SmR1, o presente trabalho teve por objetivo detalhar e ampliar os estudos de Moreno (2013) e Cirino (2014) para o gênero *Herbaspirillum*.

3 OBJETIVO

Este trabalho teve por objetivo aprofundar a análise dos ncRNAs com ênfase na classificação, análise de estrutura secundária e predição de alvos por abordagem bioinformática. A estratégia para o alcance deste objetivo consistiu de:

- Atualizar a anotação dos 173 ncRNAs identificados no genoma de *H. seropedicae* SmR1, utilizando o banco de dados de famílias de RNA Rfam 12.0.
- Pesquisar a existência dos ncRNAs identificados em *H. seropedicae* SmR1 em outras bactérias do gênero *Herbaspirillum*.
- Realizar a predição de novos ncRNAs no gênero *Herbaspirillum* utilizando a ferramenta Infernal 1.1.1
- Comparar os ncRNAs identificados com ncRNAs Hfq-dependentes já descritos na literatura.
- Predizer os alvos dos ncRNAs por Bioinformática.

4 MATERIAIS E MÉTODOS

4.1. Conjunto de dados

4.1.1. Genomas de bactérias do gênero *Herbaspirillum*

Os genomas de *H. seropedicae* SmR1, *H. seropedicae* Z67, *H. lusitanum* P6-12, *H. hiltineri* N3, *H. frisingense* GSF30, *Herbaspirillum* spp. estirpe B65, *H. rubrisubalbicans* M1, *H. autotrophicum* IAM 14942, *H. huttiense* subsp. *puttei*, *Herbaspirillum* spp. estirpe B501, *Herbaspirillum* spp. estirpe GW103, *Herbaspirillum* spp. estirpe RV1423, *H. rhizosphaerae* UMS-37, *H. rubrisubalbicans* spp. estirpe Os34, *H. rubrisubalbicans* spp. estirpe Os45 e *Herbaspirillum* AU14040 foram obtidos do banco de dados do *National Center for Biotechnology Information Search* (NCBI). Os respectivos números de referência e as datas de acesso estão listados no ANEXO II.

4.2 Sequências de ncRNAs de *H. seropedicae* SmR1.

As 173 sequências de prováveis ncRNAs *H. seropedicae* SmR1 com expressão confirmada foram coletadas das dissertações de mestrado de Moreno (2013) e Cirino (2014).

4.3 Dados de RNA-seq

Dados de expressão gênica de *H. seropedicae* SmR1 foram gentilmente cedidos pelos pesquisadores do Núcleo de Fixação Biológica de Nitrogênio. As condições de cultivo e/ou inoculação *in planta* estão descritas no (ANEXO I).

4.4 Anotação das sequências de ncRNA

As sequências de ncRNAs foram anotados utilizando o banco de dados online Rfam¹ versões 12.0 e 12.1. Este banco contém uma coleção de famílias de ncRNA representada por sequências de alinhamentos manualmente curadas, estruturas secundárias consenso, anotações colhidas de fontes de taxonomia e ontologia

¹ <http://rfam.xfam.org>

(NAWROCKI et al., 2014). A base é uma fonte ampla e diversificada de ncRNAs, inclui 2.474 famílias de ncRNAs, com informações sobre diversos tipos de ncRNAs em todo os três domínios da vida e vírus. Infernal 1.1.1 foi usado para realização de alinhamentos múltiplos de sequências pelo modelo de covariância e, além da anotação de ncRNAs, o Rfam 12.1 classifica os ncRNAs e fornece referências bibliográficas para cada família, *links* para o PDB (do inglês *Protein Data Bank*), miRBase, ENA e Gene Ontology (GO).

4.5 Ferramentas de Bioinformática

As seguintes ferramentas de bioinformática foram utilizadas:

4.5.1 *Artemis*²

Ferramenta desenvolvida pelo Instituto Sanger de anotação de genoma, que permite a visualização de sequências nucleotídicas e anotações nos formatos EMBL e GENBANK ou FASTA (RUTHERFORD et al., 2000). Este programa foi escrito em *Java* e está disponível para diversas plataformas, como UNIX, Linux, Windows. É possível visualizar os alinhamentos das leituras de RNA-seq em um arquivo no formato BAM. Adicionalmente, valores de RPKM podem ser calculados para cada ncRNA candidato (CARVER et al. 2010).

4.5.2 *BLAST*³

O algoritmo utilizado pelo programa BLAST (do inglês *Basic Local Alignment Search Tool*), compara sequências de nucleotídeos ou proteínas com sequências de base de dados e calcula a significância estatística dos alinhamentos.

² <http://www.sanger.ac.uk/science/tools/artemis/>

³ <https://blast.ncbi.nlm.nih.gov/>

Este algoritmo permite diferentes comparações de sequências: Blastp compara sequências de aminoácidos (*query*) contra um banco de dados de sequências de proteínas (*subject*); Blastn compara sequências de nucleotídeos contra banco de dados de sequências de nucleotídeos; Blastx compara sequências de nucleotídeos traduzidas para todas as fases de leitura possíveis contra um banco de dados de sequências de proteínas; tBlastn compara sequências de aminoácido contra um banco de dados de sequências de nucleotídeos traduzidas e tBlastx compara as seis fases de leitura de uma sequência de nucleotídeos contra um banco de dados de nucleotídeos traduzido para possíveis fases de leitura (ALTSCHUL et al., 1990).

4.5.3 *Clustaw*⁴

Programa usado para o alinhamento global múltiplo de sequências de DNA ou proteínas. Pode ser utilizado na identificação de regiões conservadas das sequências e para inferências filogenéticas (HIGGINS et al., 1988).

4.5.4 *Infernal 1.1.1*⁵ (*INFERENCE of RNA Alignment*)

Ferramenta computacional que utiliza modelo de covariância para gerar perfis probabilísticos de sequências e estruturas secundárias de famílias de RNA (NAWROCKI et al., 2013). O modelo de covariância é um caso especial de modelo probabilístico que valoriza uma combinação da sequência consenso e de estrutura secundária consenso de um dado RNA de modo que, em muitos casos, é capaz de identificar RNAs homólogos que tem a estrutura secundária conservada, mas baixa conservação de sequência primária (NAWROCKI et al., 2013). O Infernal 1.1.1 é composto por cinco programas: *cmbuild*, *cmcalibrate*, *cmsearch*, *cmscan* e *cmalign*.

⁴<http://www.ebi.ac.uk/Tools/msa/clustalw2/>

⁵<http://eddylab.org/infernal/>

Para os genomas estudados neste trabalho, a seguinte linha de comando foi utilizada:

```
cmScan -o sequenceSMR1.out -tblout sequenceSMR1.tbl -T 24 -notrunc Rfam.cm
sequenceSMR1
```

Onde *sequenceSMR1*, que representa o genoma da bactéria *H. seropedicae* SmR1, foi substituído pelos genomas listados no item 4.1.1.

4.5.5 TargetRNA2⁶

Ferramenta de predição de mRNAs alvos que fazem paramento de bases com ncRNAs (TJADEN et al., 2006; KERY et al., 2014). Calcula a pontuação de hibridização e a significância estatística das interações entre mRNA-ncRNA (TJADEN, 2008; BACKOFEN e HESS, 2010). Utiliza uma variedade de recursos para identificar mRNAs alvos:

- a) Conservação do ncRNA - comparação da sequência disponível no GenBank de acordo com o genoma (*replicon*) depositado e indicação de regiões com maior conservação de sequências propensas a sofrerem interações mRNA-ncRNA;
- b) Acessibilidade - do ncRNA estrutura do mRNA é ponderado pela sua estabilidade e considerada nas regiões de interações de ncRNAs;
- c) Energia de hibridização - regiões do mRNAs com baixo índice de energia hibridização em uma ou mais regiões de interações com alvos potenciais (KERY et al., 2014).

4.5.6 RNAfold⁷

Ferramenta componente do Pacote ViennaRNA 2.0 prediz estruturas secundárias através da determinação da energia livre mínima. O registro de saída padrão (*output*) é uma representação da estrutura (LORENZ et al., 2011).

⁶ <http://cs.wellesley.edu/~btjaden/TargetRNA2/>

⁷ <http://rna.tbi.univie.ac.at/cgi-bin/RNAfold>

4.5.7 SamTools

É um pacote para a manipulação de mapeamentos em formato SAM (LI et al., 2009). Neste trabalho, foram utilizados os seguintes comandos:

a) Conversão do mapeamento de SAM para BAM

samtools view -bS arquivo.SAM > arquivo.BAM

b) Alinhamento das coordenadas

samtools sort arquivo.BAM arquivobam.sorted

c) Classifica o índice para acelerar o acesso ao arquivo

samtools index arquivo.sorted.bam

Todos os programas foram executados na plataforma Linux®.

4.6 ANÁLISE DO TRANSCRIPTOMA

A normalização das amostras sequenciadas na plataforma SoliD foi realizada pelo método de RPKM (*reads per kilobase of transcript per million mapped reads*) que estima o valor de expressão de um gene onde mede a densidade das leituras em uma região gênica de interesse, normalizando a contagem das leituras em suas regiões exônicas comparadas ao tamanho original do gene ou éxon (MORTAZAVI et al., 2008).

Este método considera o tamanho da biblioteca, o tamanho dos genes e o número de leituras por gene, de acordo com a fórmula:

$$R = \frac{10^9 C}{NL}$$

C = número de leituras por gene

L = tamanho do gene (Kb)

N = tamanho da biblioteca (número total de leituras por replicata biológica)

Para determinar o valor de cobertura, a seguinte expressão matemática foi utilizada:

$$x = \frac{C * 35}{L}$$

C = número de leituras por gene

L = tamanho do gene (Kb)

35 = tamanho da leitura de RNA-seq após o processo de *trimming*.

4.7 INFERÊNCIA DE ALVOS DOS ncRNAs DE *H. seropedicae* SmR1

Os ncRNAs com função regulatória exercem resposta pós-transcricional sobre os seus mRNAs alvos, através de interações por pareamento de bases. Esta interação de RNA:RNA resulta na formação de um circuito regulatório ou rede regulatória em que alguns ncRNAs podem estar envolvidos em diferentes mecanismos regulatórios celulares. A inferência de alvos dos ncRNAs de *H. seropedicae* SmR1 foi realizada utilizando a ferramenta TargetRNA2 (ITEM 4.5.5). As sequências dos 173 ncRNAs identificados por Moreno (2013) e Cirino (2014) foram analisadas pela ferramenta citada e os resultados foram analisados.

4.8 DESENHO DE PRIMERS PARA VALIDAÇÃO FUNCIONAL de ncRNAs

Alguns ncRNAs, dentre os 173 com expressão confirmada, foram selecionados para validação por PCR em tempo real. Para o critério de seleção considerou-se 10 ncRNAs com expressão confirmada nas condições Baixo O₂, Alto O₂ e Presença de Nitrato, dentre as 11 condições experimentais avaliadas. A lista de ncRNAs selecionados está apresentada na TABELA 3. Alguns desses ncRNAs selecionados foram estavam expressos em mais de uma condição.

NcRNA	Posição Inicial	Posição Final	Nucleotídeos	Rfam 12.0	Cobertura >=5
4.5S	41088917	4109042	146	Bactéria small SRP	AD1, AD3, AltoO2, BaixoO2, Nitrato, PS1, PS3, PT, WR
CUFF.1109	3382646	3382739	94	SucA	AD1, AD3, AltoO2, BaixoO2, Nitrato, PS1, PS3, PT, WR
Hs_noco_732	4239106	4239615	510	TmRNA	AD1, AD3, AltoO2, BaixoO2, Nitrato, PS1, PS3, PT, WR
smallRNA_49.2	4896565	4896712	148	Ausência	AD1, AD3, AltoO2, BaixoO2, Nitrato, PS1, PS3, PT, WR
CUFF.1011	2800595	2800705	111	crcB	AD1, AltoO2, Nitrato, PT
Hs_noco_530	957549	957626	78	Ausência	AD1, Nitrato
Hs_noco_517	831492	831611	120	Ausência	AltoO2, Nitrato
Hs_noco_820	5397562	5398115	525	Ausência	BaixoO2
Hs_noco_824	5433895	5434068	174	Ausência	Nitrato
Hs_noco_746	4428208	4428599	392	Ausência	Ausência de Expressão

TABELA 3 – ncRNAs de *H. seropedicae* SmR1 selecionados para desenho de *primers*. FONTE: AUTOR (2016).

4.9 CRISPRmap⁸ - CONSERVAÇÃO CRISPR e CRISPRFinder⁹

Ferramenta computacional que permitem identificar a conservação de repetição CRISPR. Permite o agrupamento independente e determinar as famílias de sequência conservadas, motivos de estrutura aptas para endoribonucleases e relações evolutivas (LANGE et al., 2013). CRISPRFinder é uma interface que proporciona a análise detalhada das sequências genômicas CRISPR (GRISSA et al., 2007).

⁸ <http://rna.informatik.uni-freiburg.de/CRISPRmap/>

⁹ <http://crispr.u-psud.fr/Server/>

5 RESULTADOS E DISCUSSÃO

A pesquisa sobre RNAs não codificadores na bactéria *H. seropedicae* SmR1 iniciou com Moreno (2013) seguido por Cirino (2014). Como resultados foram identificados 173 ncRNAs expressos em 11 condições experimentais distintas (ANEXOS I e II). O objetivo do presente trabalho foi ampliar o conhecimento sobre ncRNAs com função regulatória, utilizando novas ferramentas computacionais para identificar novos ncRNAs candidatos no gênero *Herbaspirillum*, predizer seus alvos e realizar análise de estrutura secundária. Além disso, foi iniciada etapa de caracterização *in vitro* pela seleção de alguns ncRNAs de interesse e a construção de oligonucleotídeos iniciadores para ensaios de PCR quantitativa.

5.1.1 Anotação e classificação dos ncRNAs de *H. seropedicae* SmR1

No início deste trabalho foi realizada uma atualização da identidade e classificação dos 173 ncRNAs, previamente identificados por Moreno e Cirino, no genoma de *H. seropedicae* SmR1 (ANEXO III). Para isso, foi utilizado o banco de dados de famílias de RNA Rfam 12.0 (ITEM 4.4). Entre os 173 ncRNAs anotados, 6 foram classificados pelo Rfam 12.0 como: ncRNAs *trans-encoded* 6S (RF00013); ncRNAs *trans-encoded* 4.5S (RF00169); ncRNAs *trans-encoded* tmRNA (RF00023); *riboswitch* TPP (RF00059); ncRNAs *cis-encoded* Intron_gpl (RF00028) e *riboswitch* FMN (RF00050), e outros 2 classificados pelo banco de dados RIBEX como *riboswitches* (ABREU-GOODGER e MERINO, 2005) (QUADRO 2).

No presente trabalho foram anotados mais 4 ncRNAs (QUADRO 2), sendo um ncRNAs *trans-encoded* *betaproteobactéria_toxic_sRNAs* (RF02278) e três *riboswitches* (*crcB* motif ncRNA RF01734, *sucA* RF0170 e *ykkC-ykkD* RF00442).

ncRNAs	Rfam 12.0	Classificação
Moreno (2013)		
6S	RF00013	<i>trans-encoded</i>
4.5S	RF00169	<i>trans-encoded</i>
TmRNA	RF00023	<i>trans-encoded</i>
TPP	RF00059	Riboswitch
Intron_gpl	RF00028	<i>cis-encoded</i>
FMN	RF00050	<i>Riboswitch</i>
RLE0100	RIBEX	<i>Riboswitch</i>
RLE0038	RIBEX	<i>Riboswitch</i>
Este trabalho		
Betaproteobactéria_toxic_sRNAs	RF02278	<i>trans-encoded</i>
crcB motif ncRNA	RF01734	<i>Riboswitch</i>
SucA	RF01070	<i>Riboswitch</i>
ykkC-yxkD	RF00442	<i>Riboswitch</i>

QUADRO 2 – ncRNAs codificados pelo genoma de *H. seropedicae* SmR1 e anotados de acordo com o Rfam 12.0.

Betaproteobactéria_toxic_sRNA é um pequeno RNA não codificante encontrado exclusivamente em betaproteobactérias e sem função definida (KIMELMAN et al., 2012). Com base na sequência e na conservação da estrutura, sugere-se que pode se ligar ao sítio de ligação de ribossoma na região 5'UTR de vários genes. Kimelman e colaboradores (2012) reportaram que a expressão de ncRNA tóxicos em células de *E. coli* inibiam seu crescimento devido à interferência desses ncRNAs na ligação da *dnaA* no DnaA box.

O *riboswitch crcB* é um ncRNAs que responde a íons fluoreto e tem um domínio conservado, o motivo crb (BREAKER, 2012). Está presente em várias bactérias e arqueias. O *riboswitch crcB* é uma sequência 5'UTR de diferentes genes, incluindo genes envolvidos no reparo de DNA (*mutS*), no transporte de íons (K⁺/Cl⁻) e o gene que codifica a enzima formato hidrogênio liase (WEINBERG et al., 2010).

O *sucA* é um *riboswitch* conservado e localizado na região 5'UTR do gene *sucA* (WEINBERG et al.; 2007). Está presente, em sua maioria, em bactérias da ordem *Burkholderiales* das β -proteobacterias; sua sequência se sobrepõe ao sítio de ligação do ribossomo do gene *sucA* (WEINBERG et al.; 2007; SASS et al.; 2015).

O *riboswitch ykkC-yxkD* (ou *ykkC-yxkD leader*) localiza-se a montante dos genes *ykkC* e *yxkD* em *Bacillus subtilis* e genes relacionados em outras bactérias (MEYER

et al., 2011). A região 3' conservada alinha com a haste rica em GC de um terminador transcricional em bactérias Gram-positivas e sua função está relacionada ação de bombas de efluxo e sistemas de desintoxicação em respostas a moléculas exógenas (BARRICK et al., 2004).

5.2 Pesquisa de identidade de sequências de ncRNAs em bactérias do gênero *Herbaspirillum*

Realizou-se a busca por identidade de sequência dos 12 ncRNAs de *H. seropedicae* SmR1 com anotação (QUADRO 2) em outras espécies de bactérias com genomas disponíveis no banco de dados NCBI. O resultado está apresentado na FIGURA 9. Dentre as bactérias estão o próprio gênero *Herbaspirillum* e os gêneros *Janthinobacterium* e *Herminimonas*.

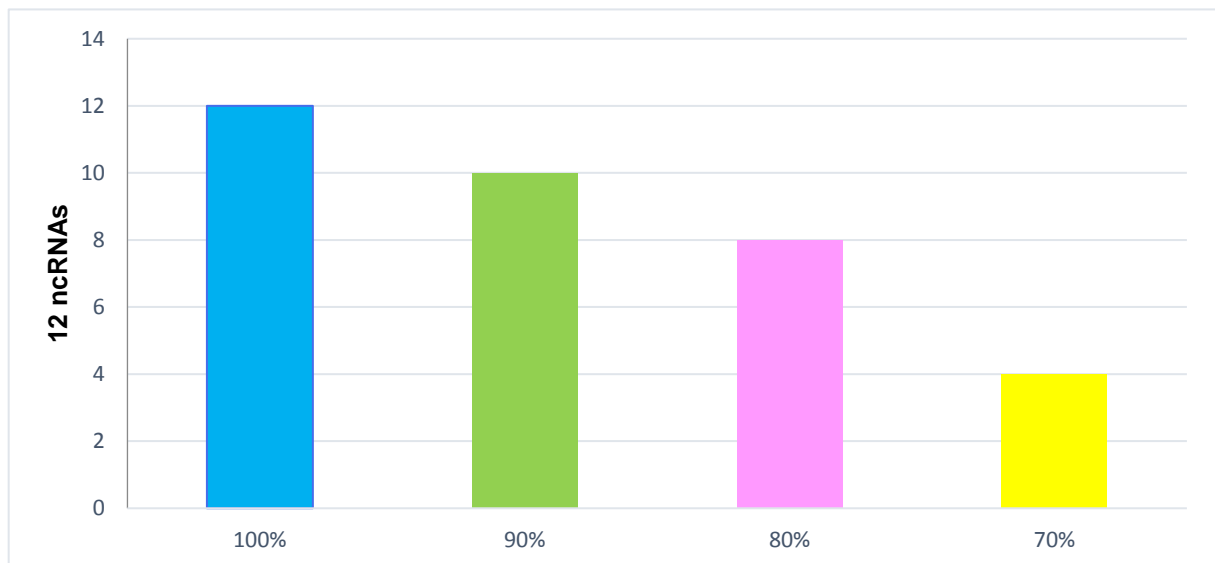


FIGURA 9 – Identidade de sequência dos 12 ncRNAs anotados no *H. seropedicae* SmR1, em outras espécies de bactérias com genomas disponíveis ao banco de dados NCBI. Azul: gênero *Herbaspirillum* com 100% de identidade, verde e rosa: gêneros *Janthinobacterium* e *Herminimonas* com 90% e 80% de identidade, respectivamente. Amarelo: demais gêneros com 70% de identidade. Os ncRNAs do gênero *Herbaspirillum* apresentaram 90% e 80% de similaridade, com os genes *Janthinobacterium* e *Herminimonas*, respectivamente. FONTE: AUTOR (2016).

5.3 Predição de novos ncRNAs no gênero *Herbaspirillum*

Uma vez que se dispunha de uma ferramenta de predição de ncRNAs que não havia sido utilizada por Moreno (2013) e de uma nova versão do banco de famílias de RNA Rfam (Rfam 12.0), realizou-se uma nova predição de ncRNAs para o genoma de diferentes bactérias do gênero *Herbaspirillum*. Foi utilizada a ferramenta Infernal 1.1.1 que tem a vantagem de fornecer a identidade do ncRNA (ITEM 4.5.4). Para *H. seropedicae* SmR1 foram preditos 57 novos ncRNAs com identidade de sequências depositadas no banco de família de RNA Rfam 12.0 (ITEM 4.4). Os ncRNAs foram classificados de acordo com a função regulatória em *cis-encoded*, *trans-encoded* ou *riboswitch*. Os resultados estão disponíveis no APÊNDICE I e APÊNDICE III. Para cada um dos novos ncRNAs foram anotadas as seguintes informações: código de acesso no Rfam 12.0 (ITEM 4.4), tamanho da sequência nucleotídica, posição inicial e final no genoma, sentido de transcrição (ITEM 4.5.2), predição de alvos pelo TargetRNA2 (ITEM 4.5.5), descrição e classificação dos ncRNAs, co-localização, condições de expressão nas 11 condições anteriormente citadas.

Os ncRNAs preditos que obtiveram o valor de cobertura igual ou superior a cinco, nas bibliotecas de RNA-seq, foram classificados como prováveis ncRNAs. Ao contrário, aqueles que apresentaram valores de cobertura inferior a cinco foram considerados somente como possíveis preditos ncRNAs, mas não expressos nas condições de cultivo testadas. Observou-se que haviam ncRNAs expressos em todas as 11 condições e que outros eram expressos em somente em algumas condições.

Essas informações constam no APÊNDICE II. Alguns exemplos de ncRNAs preditos em *H. seropedicae* SmR1 pela ferramenta Infernal 1.1.1 estão descritos abaixo.

5.3.1 ncRNA 5'_ureB_sRNA

O ncRNA 5'_ureB_sRNA possui 286 nucleotídeos e é classificado como *cis-encoded* antisense ncRNA de acordo com a sua função e localização no genoma. Apresenta sequência nucleotídica extraída pelo programa Infernal 1.1.1 (ITEM 4.5.4), descrita abaixo:

5_ureB_sRNA - undefined product 5415884:5416169 reverse

```
5'CTTCCTCGATCTGCTGCGGACAGATGAAGTGGATGTGGCTGTCGATGCCGCC
GGCGGTGACGATCATGCCTTCACCCGCGATGATTTCCGGTAGCGCCACCGATGG
CCATGGTCACGCCAGGCTGGATGTCGGGATTGCCGGCCTTGCCGATGCCGGC
GATCTTGCCCGACTTGATGCCGATGTCGGCCTTGACGATGCCCCAGTGATCCA
CGATCAGGGCATTGGTGGATGACGGTATCCATCACGTCCGCATGGGCGCGTTGC
GACTGGCCCATGCCGTCGCGGAT3'
```

O ncRNA 5_ureB_sRNA está co-localizado com os genes *ureA*, *ureB*, *ureC*, *ureD*, *ureE* e *ureF* (no sentido sense). A localização e o perfil de expressão observado no RNAseq estão mostrados na FIGURA 10.

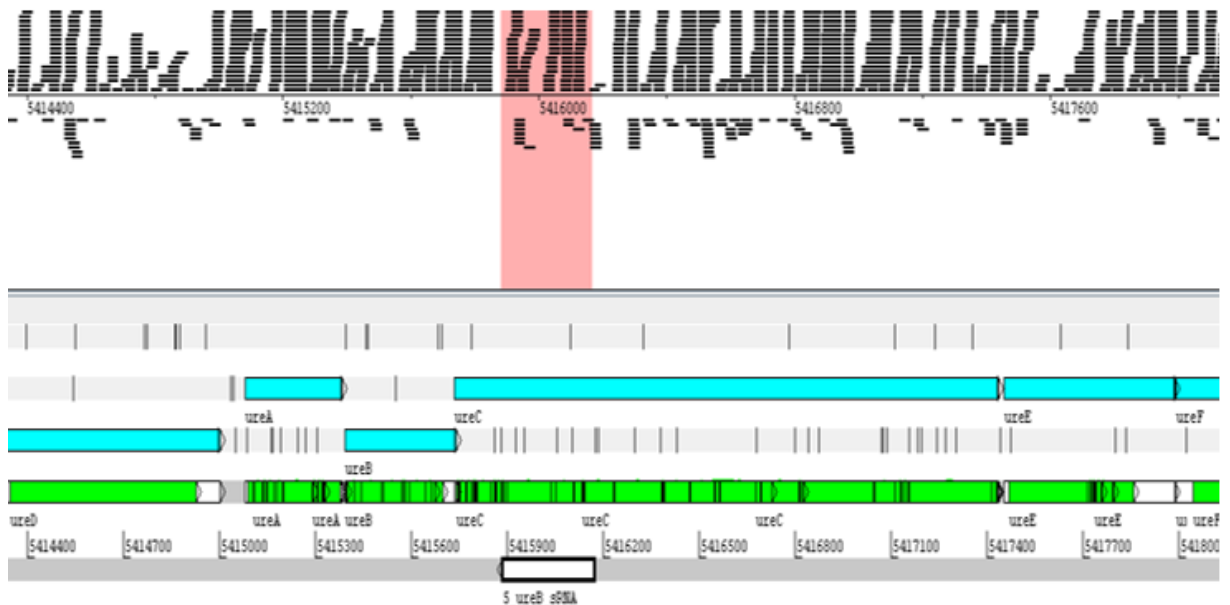


FIGURA 10 – Visualização da posição e do perfil de expressão do ncRNA 5_ureB_sRNA no genoma de *H. seropedicae* SmR1. A sequência foi localizada utilizando o programa Artemis. As leituras estão delimitadas pelo retângulo rosa. FONTE: AUTOR (2016).

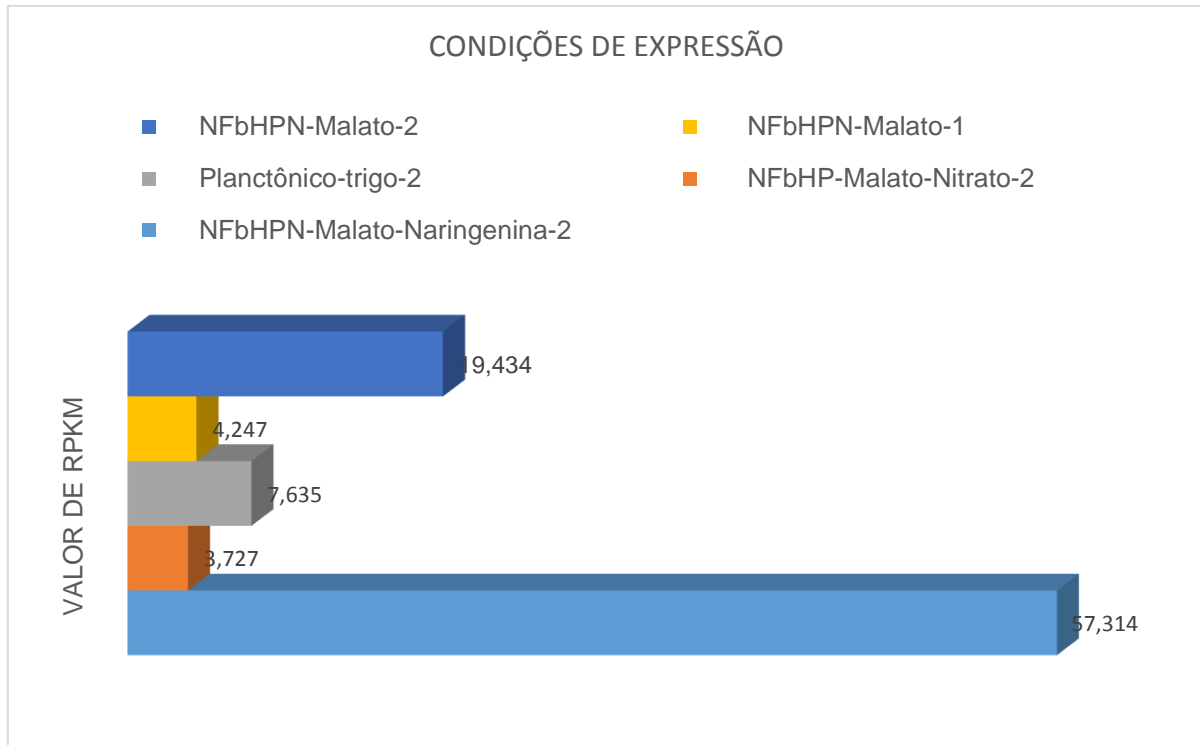


FIGURA 11 – Níveis de expressão do ncRNA 5_ureB_sRNA, em ensaios de RNA-seq, em *H. seropedicae* SmR1. Em cor, estão representadas as condições experimentais analisadas, onde 1 e 2 representam d.a.i (dias após inoculação). Os números representam o valor da expressão gênica em RPKM. FONTE: AUTOR (2016).

O ncRNA 5_ureB_sRNA apresenta maior o valor de expressão 57,314 RPKM na condição NFbHPN-Malato-Naringenina seguido da condição NFbHPN-Malato-2 com o valor de expressão de 19,434 RPKM (FIGURA 11).

H. seropedicae SmR1 utiliza naringenina como fonte de carbono, esta bactéria associa-se a plantas, onde ocorre a biossíntese de flavonoides, compostos envolvidos no mecanismo de defesa da planta. Nos ensaios de RNA-seq foi observado maior expressão do ncRNA 5_ureB_sRNA na condição NFbHPN-Malato-Naringenina, o que pode sugerir que este ncRNA 5_ureB_sRNA está envolvido na regulação da biossíntese de flavonóides, sendo que estes atuam na proteção contra agentes oxidantes na planta, e no uso de uréia gerada externamente e internamente como uma fonte de carbono.

A uréia é metabolizada pela urease, uma metaloenzima dependente do níquel, que catalisa a hidrólise da uréia para formar amoníaco e dióxido de carbono, que pode ser encontrada em bactérias, fungos e plantas, permitindo o uso de uréia gerada externamente e internamente como uma fonte de carbono (CARTER et al., 2011). Em *Klebsiella aerogenes*, o operon *ureABC* codifica as três subunidades da enzima urease (CARTER et al., 2011). Em *Helicobacter pylori*, o cluster da urease consiste de dois operons, o *ureAB* e *ureIEFGH*. Essa bactéria é a única descrita, até o momento, que contém o *5_ureB_sRNA*, um ncRNA *cis-encoded* antisense, com 290 nucleotídeos, que regula negativamente a expressão de *ureAB* ao parear na região 5' do *ureB* mRNA (WEN, FENG e SACHS, 2013). *H. seropedicae* SmR1 apresenta um provável operon *ureABC* e um ncRNA *5'_ureB*, com 286 nucleotídeos. De acordo a predição do Infernal 1.1.1, esse ncRNA é um *cis-encoded* que está posicionado antisense ao gene *ureC*.

Esta diferença pode ser justificada pelo fato do *H. pilory* não apresentar o gene *ureC*, as sequências gênicas de *H. seropedicae* serem maiores que aquela de *H. pilory* e o metabolismo da ureia envolver proteínas distintas.

O ncRNA *5'_ureB_sRNA* se repete em alguns gêneros de *Herbaspirillum* por apresentar 100% de identidade de sequência como, por exemplo, *H. seropedicae* Z67, *H. lusitanun*, *H. hiltineri* N3, *H. rubrisubalbicans* M1 e *H. seropedicae* AU14040. Realizou-se a análise da localização do locus do operon *ureABC* no genoma destas espécies, identificou-se a presença das subunidades ureaseC (*alpha*), ureaseA (*gamma*) e ureaseB (*beta*), e verificou-se que o ncRNAs *5'_ureB_sRNA* nessas espécies de *Herbaspirillum* faz interação *in cis* da subunidade ureaseC.

5.3.2 mRNA 5' UTR *cspA*

O mRNA 5' UTR *cspA* possui 373 nucleotídeos é classificado como um *cis-encoded* ncRNA de acordo com a localização no genoma. Em relação à função é considerado um termoregulator. Apresenta sequência nucleotídica abaixo:

mRNA 5' UTR *cspA* undefined product 1599563:1599935 reverse

```
5'ATCAGTTACATCTTGTTACTGTTGGCCGATAGAAGGGGGCACGGTTTAAGTGG
GGCTGCGTGACGCGCGATGTTTCTTGCTTTAACGGCTCGTGTGATAGTTTATAT
TTGAAAGATGCTTTTATGGCAACAGGTAAGTGTCAAGTGGTTCAACGATTCCAAAG
GCTTCGGCTTTATCACTCCGGATGACGGCGGCGAAGATCTGTTTCGCCCACTTCT
CCGCAATCCAGATGAACGGCTTCAAGACCCTGAAAGAAGGTCAAAAAGTCCAGT
TTGACGTCACGCAAGGCCCGAAGGGCAAGCAAGCATCGAACATCCAGAACGCC
GCCTAAACCCCGGTCTTCTCTGAAAAGCCCCGCCCGCGGGGCTTTTTTC3'
```

No genoma de *H. serpedicae* SmR1 o mRNA 5' UTR *cspA* aparece como uma sequência mRNA 5' UTR do gene anotado como *cspD* (*Hsero_1397*). Além disso, há um segundo gene *cspD* (*Hsero_3028*) que não contém a sequência mRNA 5'UTR. Segundo Yamanaka e Inouye (1997), em *E. coli* existem novos homólogos de *cspA*, denominados *cspA* a *cspI*, e o gene *cspD* não apresenta a sequência 5'UTR. Esta informação sugere que em *H. seropedicae* SmR1 o gene *cspD* (*Hsero_1397*) é na verdade o gene *cspA*.

Este gene está co-localizado com os genes *icd*, *clpS*, *clpA* (no sentido sense). A localização e o perfil de expressão observado no RNAseq estão mostrados na FIGURA 12.

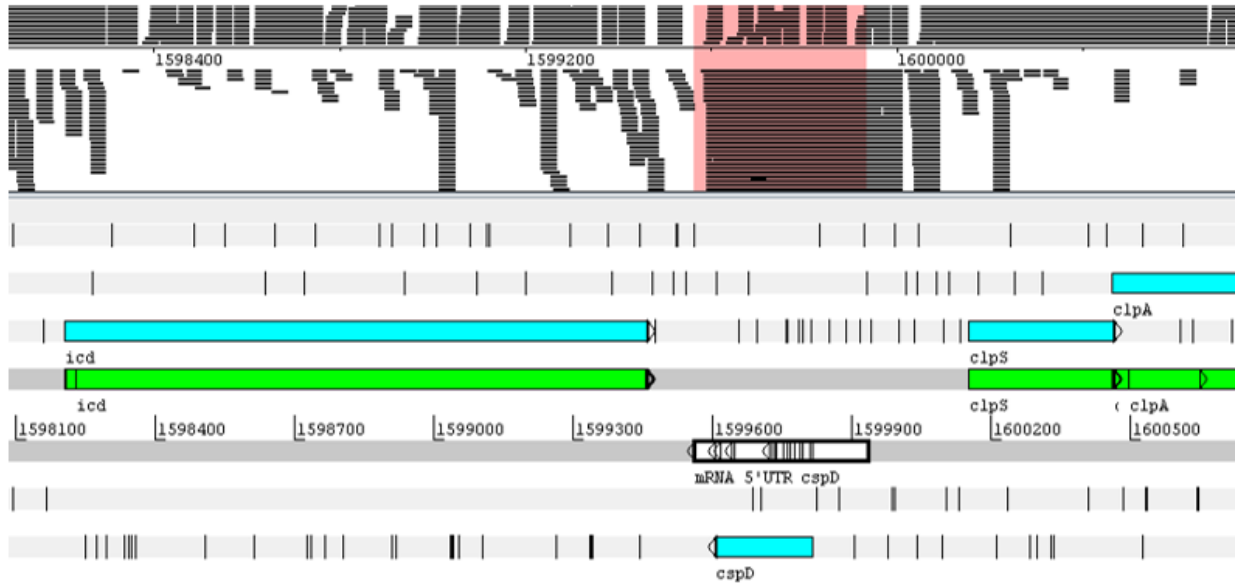


FIGURA 12 – Visualização da posição e do perfil de expressão do mRNA 5'UTR *cspA* no genoma de *H. seropedicae* SmR1. A sequência foi localizada utilizando o programa Artemis. As leituras estão delimitadas pelo retângulo rosa. FONTE: AUTOR (2016).

Jiang e colaboradores (1996) analisaram a função da sequência não traduzida 5' UTR mRNA *cspA* da principal proteína de adaptação ao choque frio a CspA, em *E. coli*. Observaram que a adaptação ao choque frio é bloqueada quando a região 5' UTR mRNA *cspA* é superexpressa, ocorrendo a síntese não somente da proteína CspA, como também, de proteínas da família CspA como CspB, CspC, CspD e CspE.

O gene *cspD* em *E. coli* codifica uma proteína de sequência homóloga com a proteína de choque frio a CspA, mas a expressão de *cspD* não é induzida pelo choque frio (YAMANAKA e INOUE, 1997). Yamanaka e Inouye (1997) utilizaram a fusão *cspD-lacZ* e verificaram que a expressão *cspD* induzida pelo crescimento em fase estacionária não depende do fator sigma σ^S , assim a expressão do gene *cspD* é inversamente dependente das taxas de crescimento induzida pela inanição de glucose.

O mRNA 5'UTR *cspA* está expresso em todas as condições ensaiadas, com destaque de 73552,195 RPKM em NFbHPN-Malato-Nitrato-1, 54949,258 e 34991,738 RPKM em NFbHP-Malato-Nitrato-2 e NFbHPN-Malato-Alto-Oxigênio-1, respectivamente (FIGURA 13). O maior valor obtido em ensaios de RNA-seq foi na condição NFbHPN-Malato-Nitrato-1, onde o *H. seropedicae* SmR1 é cultivado a 30°C em meio NFbHP, utilizando malato como fonte de carbono e nitrato 10mM como fonte de nitrogênio (BONATO, 2012).

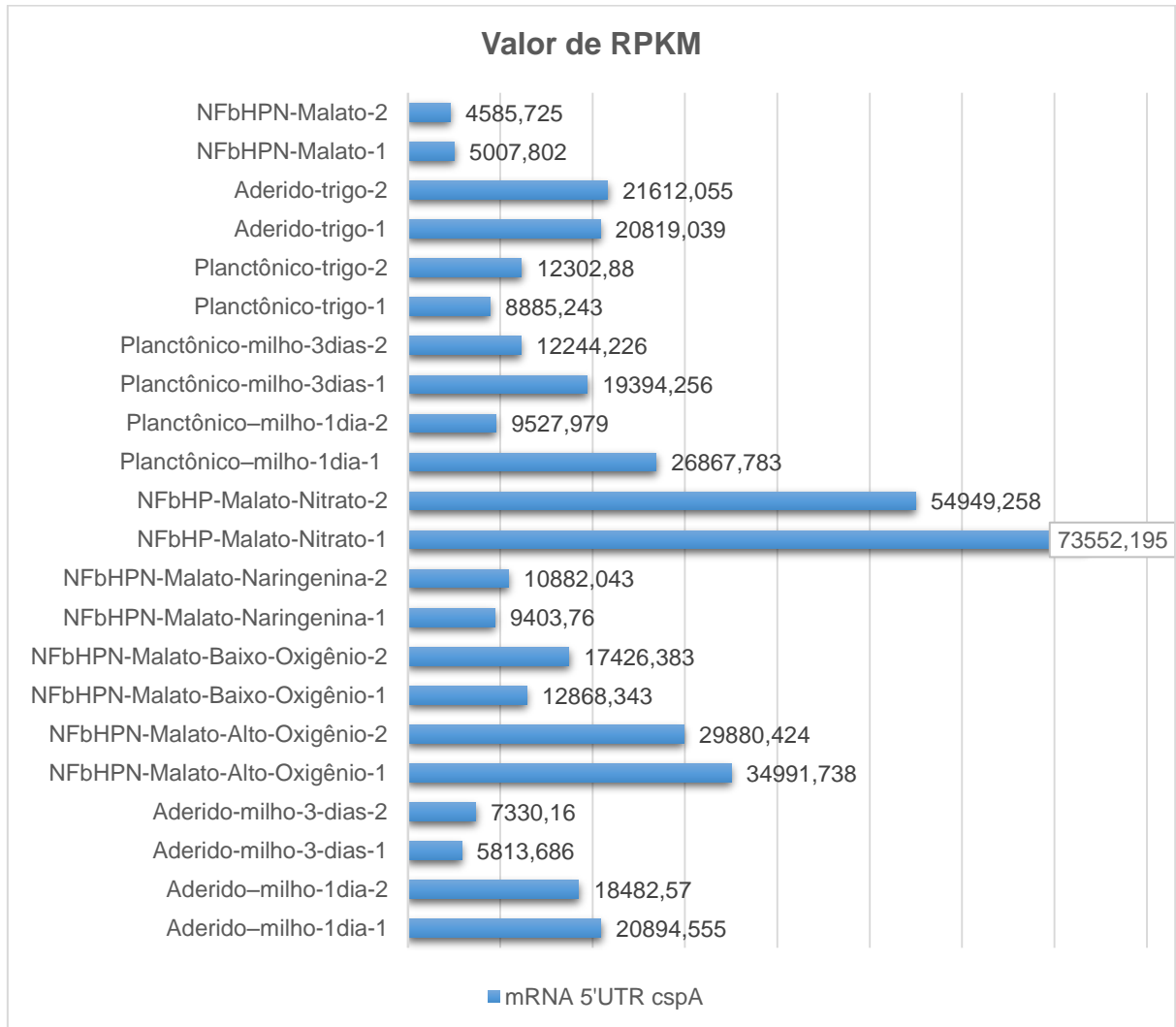


FIGURA 13 – Visualização dos níveis de expressão do mRNA 5'UTR *cspA*, em ensaios de RNA-seq, em *H. seropedicae* SmR1. Em cor, estão representadas as 11 condições experimentais analisadas, onde 1 e 2 representam d.a.i (dias após inoculação). Os números representam o valor da expressão gênica em RPKM. FONTE: AUTOR (2016).

5.3.3 ncRNA sX4

O ncRNA sX4 possui 113 nucleotídeos e é classificado *trans-encoded* antisense ncRNA. Apresenta sequência nucleotídica abaixo:

sX4 undefined product 109972:110084 forward

5'TGAGTTCCCTGCACGGAGAGGTGGCCGAGTGGTTAATGGCAGCAGACTGTAA
ATCTGCCCTCTTACGAGTACGCTGGTTCGAATCCAGCCCTCTCCACCATTAGGC
CGGGAGT3'

O ncRNA sX4 está em co-localização com os genes *ampG*, *tufB* e *trnG*. A localização e o perfil de expressão observado no RNAseq estão mostrados na FIGURA 14.

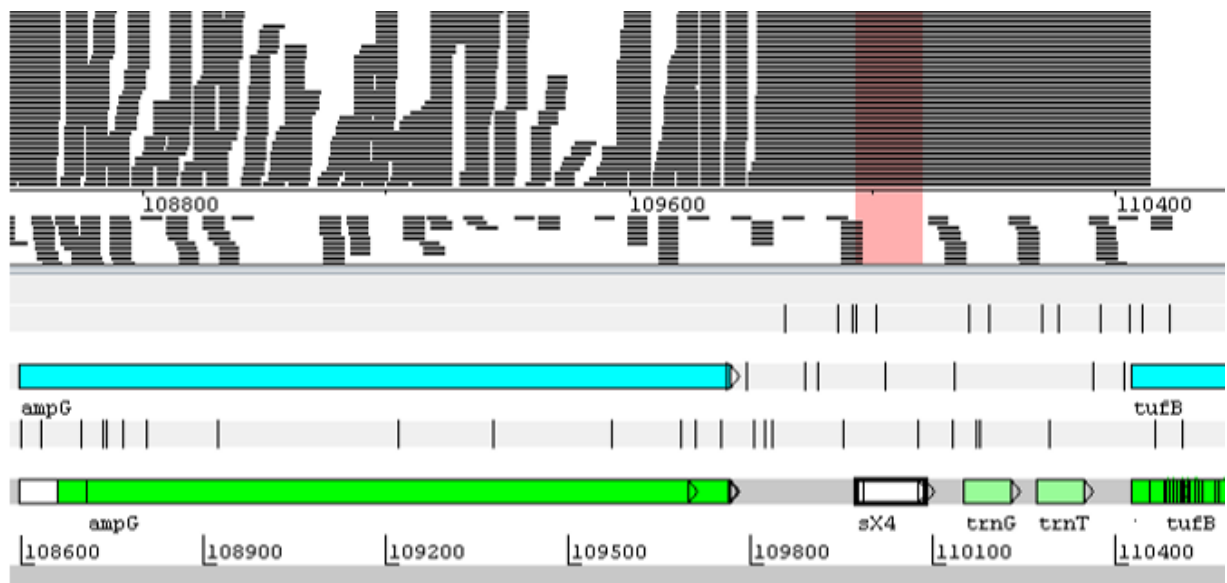


FIGURA 14 – Visualização da posição e do perfil de expressão do ncRNA sX4 no genoma de *H. seropedicae* SmR1. A sequência foi localizada utilizando o programa Artemis. As leituras estão delimitadas pelo retângulo rosa. FONTE: AUTOR (2016).

A denominação sX4 refere-se aos ncRNAs descritos em *Xanthomonas campestris* pv. *vesicatoria* (SCHMIDTKE et al., 2012). Esses microrganismos são γ -proteobactérias Gram-negativas e patógenos de muitas espécies de plantas economicamente importantes. Possuem o ncRNA sX4 cuja expressão é dependente das proteínas HrpG e HrpX, sugerindo que o(s) alvo(s) desse ncRNA esteja(m)

presentes na interação da bactéria com a planta (SCHMIDTKE et al., 2012, ABENDROTH, SCHMIDTKE e BONAS, 2014). Estudos com mutantes HrpG e HrpX em *X. campestris pv. vesicatoria* revelam que estes genes são essenciais para a patogenicidade e contribuem para a sobrevivência das bactérias epifíticas (BÜTTENER, BONAS; 2009).

Em *H. seropedicae* SmR1 o ncRNA sX4 está expresso em todas as condições ensaiadas (FIGURA 15) mas os maiores valores de expressão foram observados nas condições Aderido-milho-3dias com 6481,961 RPKM e 6239,265 RPKM respectivamente, seguido da condição Planctônico-milho-3dias-2 com 4605,488 RPKM. Nota-se então que a maior expressão se deu durante a colonização de milho.

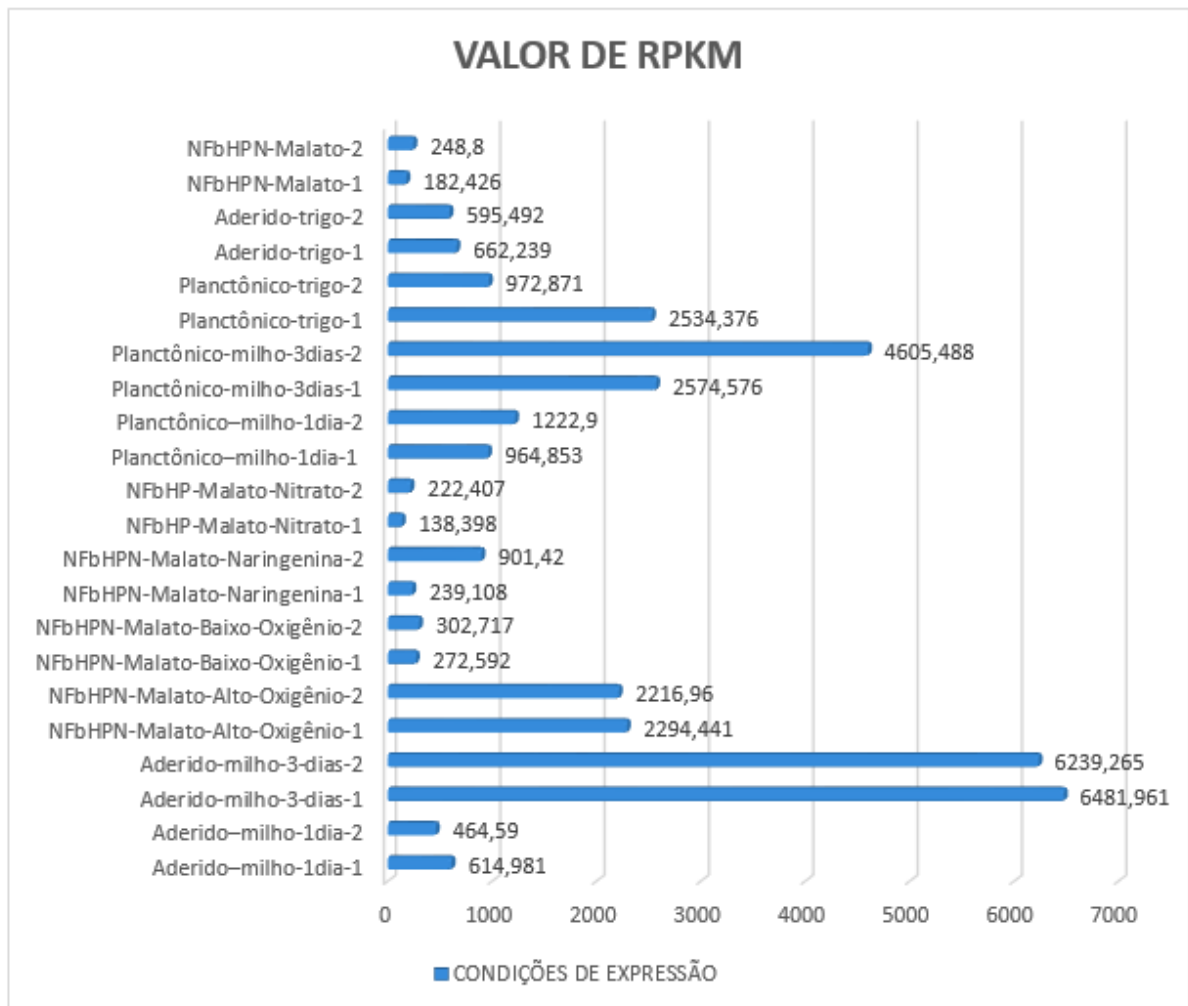


FIGURA 15 – Visualização dos níveis de expressão do ncRNA sX4, em ensaios de RNA-seq, em *H. seropedicae* SmR1. Em cor, estão representadas as 11 condições experimentais analisadas, onde 1 e 2 representam d.a.i (dias após inoculação). Os números representam o valor da expressão gênica em RPKM. FONTE: AUTOR (2016).

5.3.4 Riboswitch Cobalamina

O *Riboswitch Cobalamina* possui 113 nucleotídeos e é classificado *trans-encoded* antisense ncRNA. Apresenta sequência nucleotídica abaixo:

Cobalamin riboswitch undefined product 3038673:3038919 reverse

```
5'GAACTGGACAAGACCTTCGAGGAACGCGACCACATCAACACCGCCATCGTCA
GCGCCATCGACGAGTCGGCCGAGAACTGGGGCGTGAAGGTGCTGCGCTACGA
GATCAAGGACCTGACGCCGCCAAGGAAATCCTGCACGCCATGCAGGCCCAGA
TTACCGCCGAGCGTGAGAAGCGCGCCCTCATCGCCGCCTCCGAAGGCCGCAA
GCAGGAACAGATCAATATCGCCACCGGTGAGCGCGAAG3'
```

O *Riboswitch Cobalamina*, ou elemento Cobalamina, está co-localizado com o gene *Hsero_2659* (transportador de cobalto), como uma sequência 5'UTR, e com os genes *Hsero_2660* e *Hsero_2661* (sentido antisense). A localização e o perfil de expressão observado no RNAseq estão mostrados na FIGURA 16.

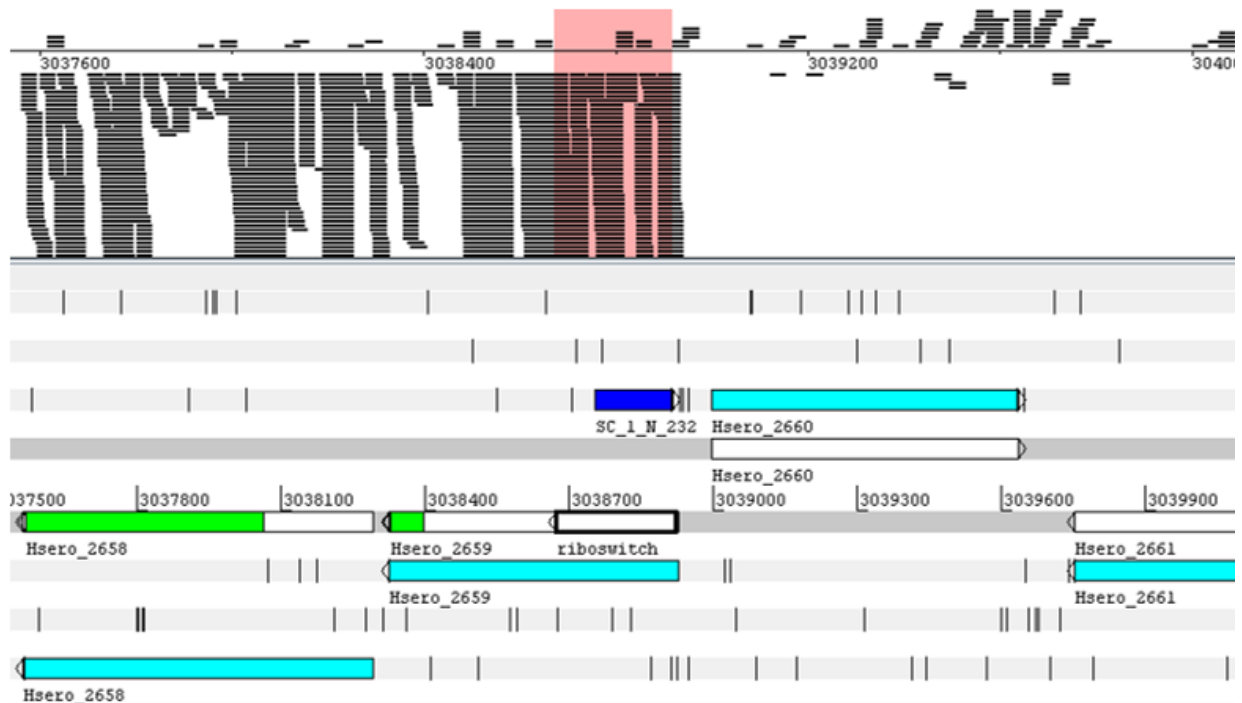


FIGURA 16 – Visualização da posição e do perfil de expressão do ncRNA Cobalamina no genoma de *H. seropedicae* SmR1. A sequência foi localizada utilizando o programa Artemis. As leituras estão delimitadas pelo retângulo rosa. FONTE: AUTOR (2016)

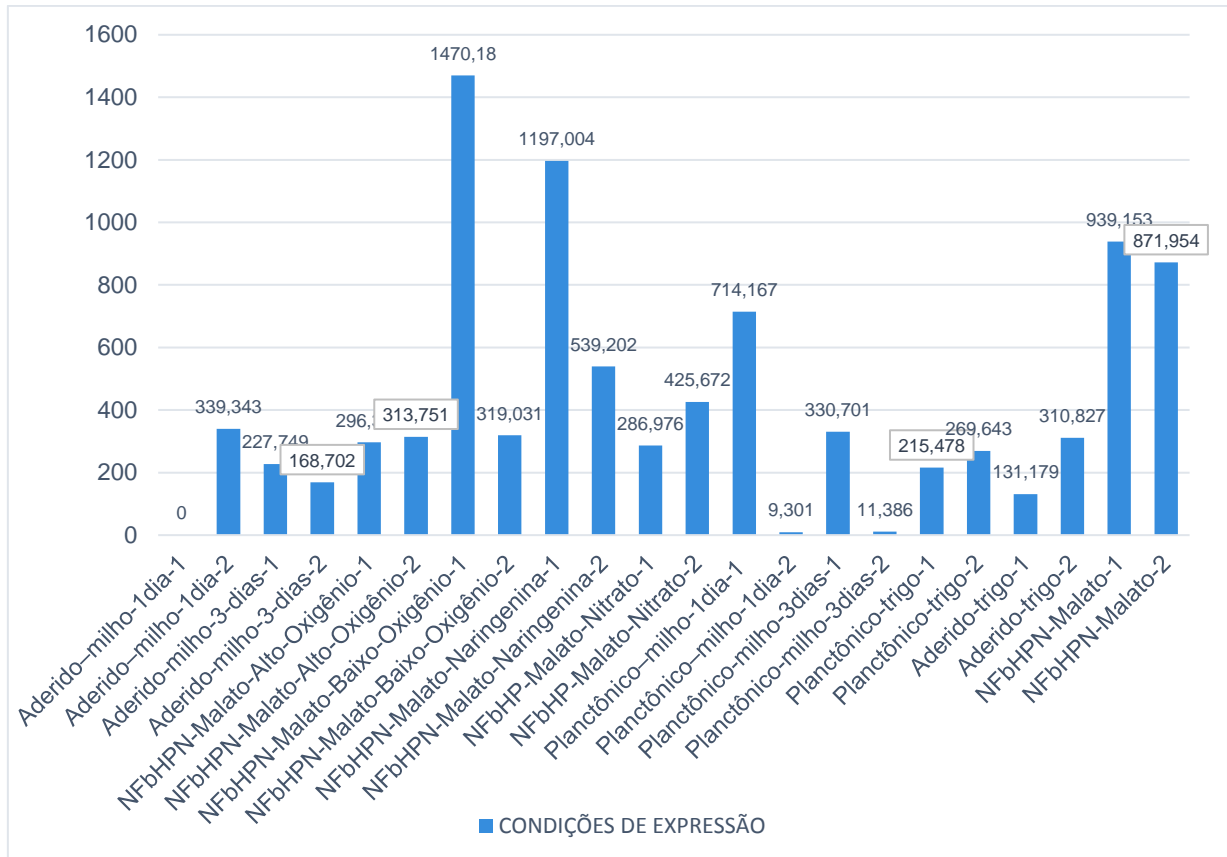


FIGURA 17 – Visualização dos níveis de expressão do ncRNA Cobalamina, em ensaios de RNA-seq, em *H. seropedicae* SmR1. Em cor, estão representadas as 11 condições experimentais analisadas, onde 1 e 2 representam d.a.i (dias após inoculação). Os números representam o valor da expressão gênica em RPKM. FONTE: AUTOR (2016).

Coenzima B12 ou *riboswitches* Cobalamina são elementos de controle amplamente distribuídos em procaríotos. Por exemplo, o gene *metE* contém um cobalamina *riboswitch* e a ligação da coenzima B12 a este elemento leva à formação de um terminador transcricional e à repressão da expressão do *metE*. Na ausência da coenzima o mRNA *metE* é sintetizado. Em *H. seropedicae* o ncRNA Cobalamina apresenta maior nível de expressão na condição NFbHPN-Malato-Baixo-Oxigênio-1 1470,18 RPKM e na condição NFbHPN-Malato-Naringenina-1 1197,004 RPKM (FIGURA 17). Esses resultados indicam que provavelmente não está ocorrendo ligação do ligante CbtB (subunidade transportador de cobalto) ao mRNA *hsero_2659* na região do aptâmero, onde ocorre a formação de um grampo alternativo.

5.3.5 ncRNA MicC

O ncRNA MicC possui 97 nucleotídeos e é classificado *trans-encoded* antisense. Apresenta sequência nucleotídica abaixo:

MicC undefined product 1775814:1775910 reverse
 5'CGGAAAGCGCCGCCCGCCGTGCGCCTTCCAGTATGATGACCACTGAATGGAC
 GGTCTGCGCGCCTGCGTCCGTCCACCTTGAGGAAACACCATGAAC3'

O gene para o ncRNA MicC está co-localizado com os genes *murI* e *fumA* e com o gene *gst* (sentido sense). A localização no genoma de *H. seropedicae* SmR1 e o perfil de expressão observado no RNAseq estão mostrados na FIGURA 18.

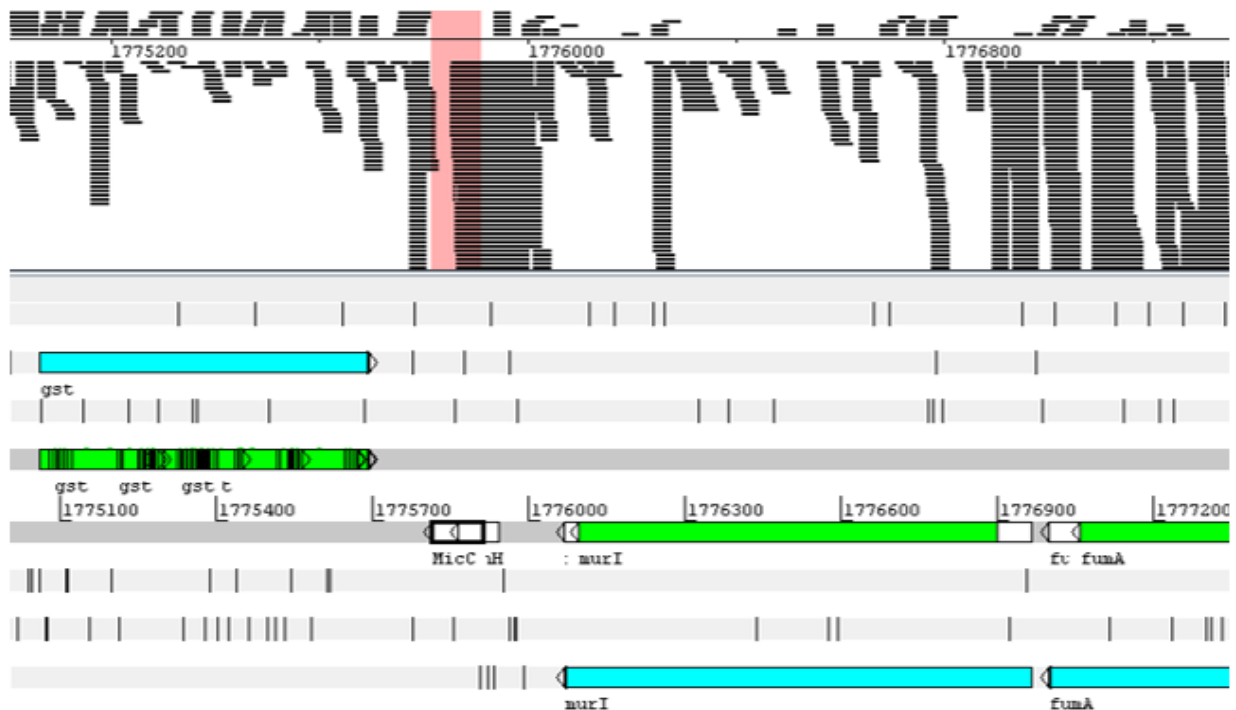


FIGURA 18 – Visualização da posição e do perfil de expressão do ncRNA MicC no genoma de *H. seropedicae* SmR1. A sequência foi localizada utilizando o programa Artemis. As leituras estão delimitadas pelo retângulo rosa. FONTE: AUTOR (2016).

Em *E. coli*, o ncRNA MicC está localizado entre os genes *ompN* e *ydbk* e regula a expressão da proteína OmpC ao se parear com a sequência líder do mRNA e inibir a ligação do *ompC* mRNA ao ribossomo (CHEN et al., 2004). No genoma de *H.*

seropedicae SmR1 o gene *ompC* está localizado no locus *Hsero_1282*, o gene *ompA* está localizado no locus *Hsero_1287* e o gene *ompR* está localizado no locus *Hsero_1518*. Utilizando a ferramenta TargetRNA2 (ITEM 4.5.5), observou-se que o ncRNA MicC em *H. seropedicae* SmR1 possui como alvo o gene *rscB* localizado no locus *Hsero_1538*, que atua como um regulador transcricional que tem um domínio de receptor de sinal para a proteína OmpR.

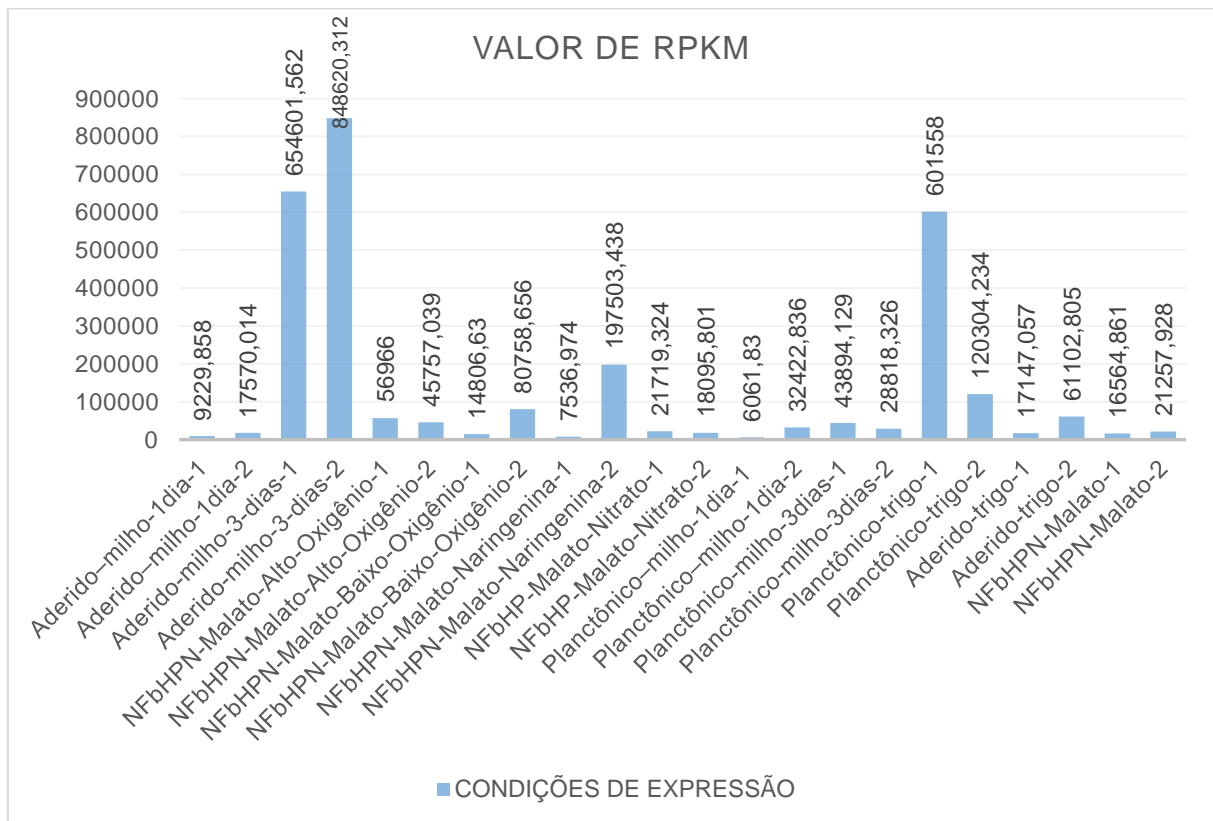


FIGURA 19 – Visualização dos níveis de expressão do ncRNA MicC, em ensaios de RNA-seq, em *H. seropedicae* SmR1. Em cor, estão representadas as 11 condições experimentais analisadas, onde 1 e 2 representam d.a.i (dias após inoculação). Os números representam o valor da expressão gênica em RPKM. FONTE: AUTOR (2016).

O ncRNA *MicC* está expresso em todas as condições ensaiadas e apresentou maiores valores de expressão observados nas condições, Aderido-milho-3-dias-2 e Aderido-milho-3-dias-1, foram observados valores, em RPKM, de 848620,312 e 654601,562 com maior nível de expressão em relação as outras condições (FIGURA 19). Esses resultados sugerem que a expressão ncRNA *MicC*, um regulador da expressão da proteína *Omp*, que em *H. seropedicae* SmR1 está envolvido na expressão transcricional do gene *rscB* quando está aderido ao milho.

5.3.6 ncRNAs IsrD e IsrG

O ncRNA IsrD possui 62 nucleotídeos e apresenta sequência nucleotídica abaixo:

IsrD undefined product 1093595:1093656 forward

5'ATTATTTATACGATTTCCGCTTCGGGAACAGCAAAAAGGCCAGCGCGAAGCTG
GCCTTTTCG3'

O ncRNA IsrD está co-localizado com os genes *mmr* e *Hsero_0991* e com o gene *Hsero_0992* (sentido antisense). A localização e o perfil de expressão observado no RNAseq estão mostrados na FIGURA 20.

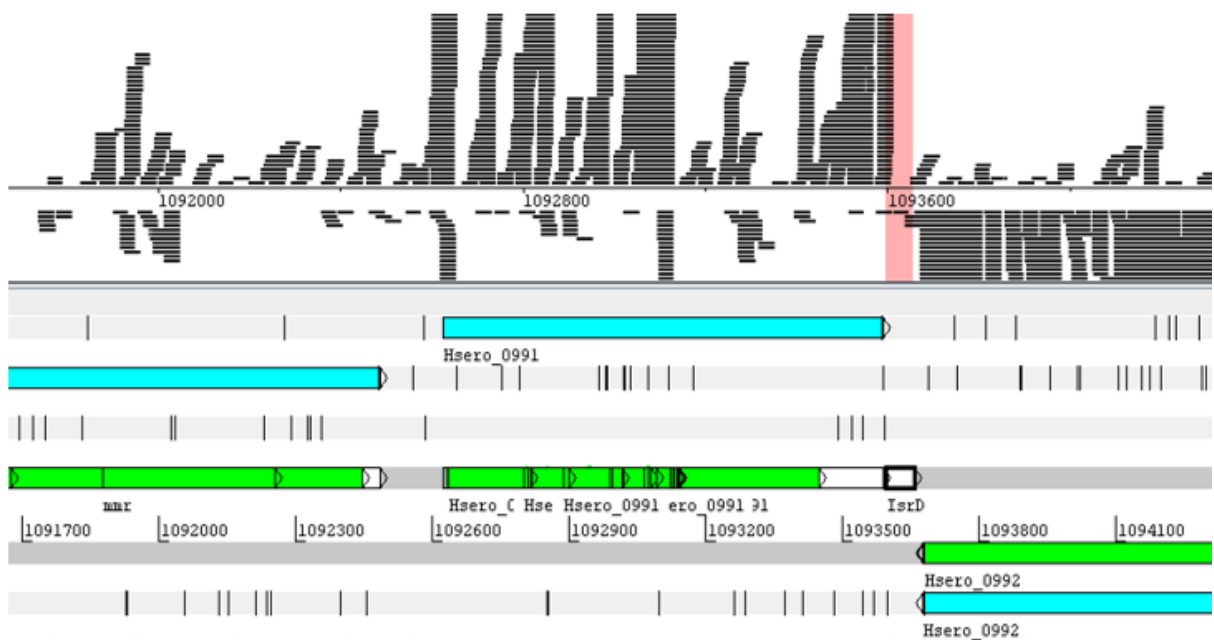


FIGURA 20 – Visualização da posição e do perfil de expressão do ncRNA IsrD no genoma de *H. seropedicae* SmR1. A sequência foi localizada utilizando o programa Artemis. As leituras estão delimitadas pelo retângulo rosa. FONTE: AUTOR (2016).

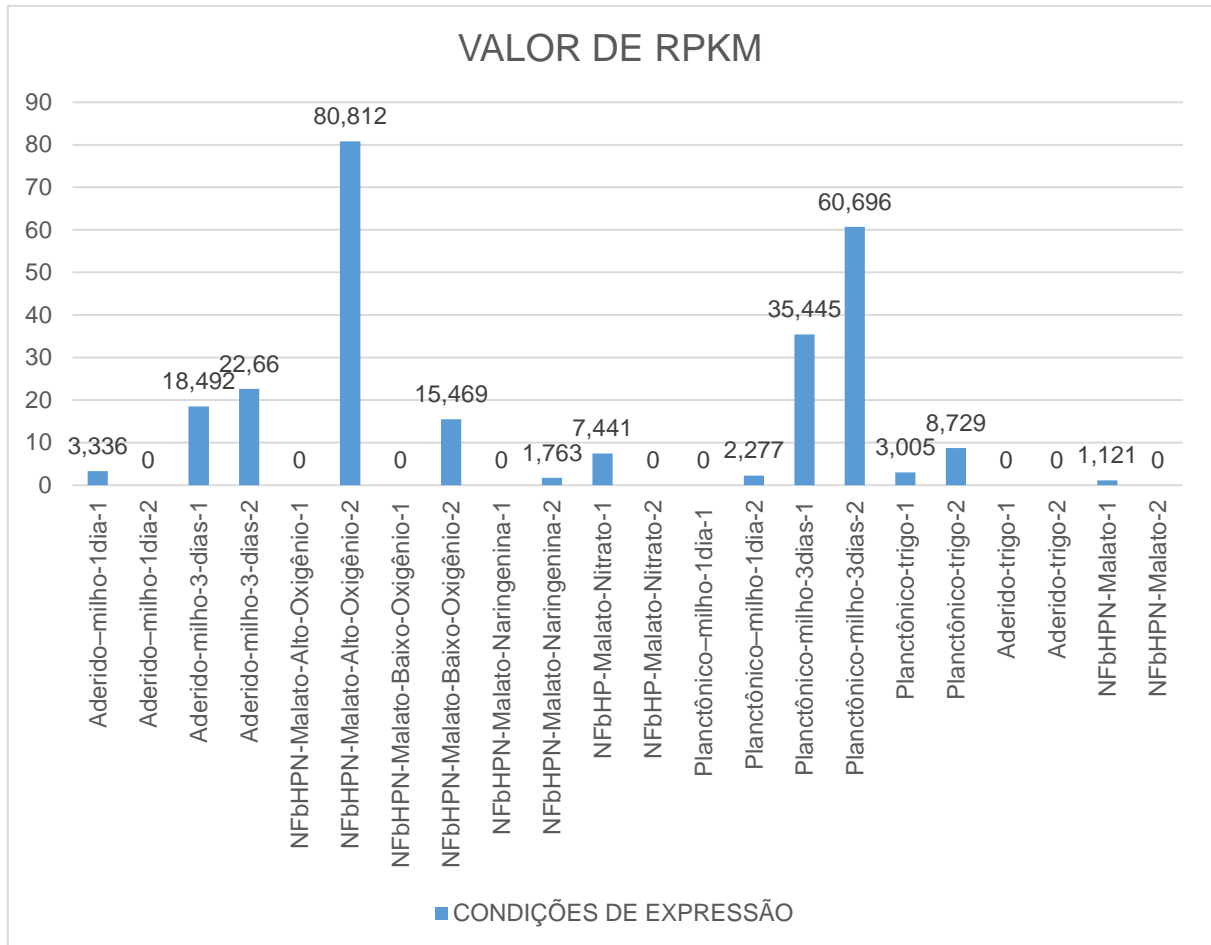


FIGURA 21 – Visualização dos níveis de expressão do ncRNA *IsrD*, em ensaios de RNA-seq, em *H. seropedicae* SmR1. Em cor, estão representadas as 11 condições experimentais analisadas, onde 1 e 2 representam d.a.i (dias após inoculação). Os números representam o valor da expressão gênica em RPKM. FONTE: AUTOR (2016).

O ncRNA *IsrD* está expresso em algumas condições ensaiadas. Nas condições, Planctônico-milho-3dias-2 e NFbHPN Malato-Alto-Oxigênio-2, foram observados maiores valores de expressão (60,696 e 80,812 RPKM respectivamente) (FIGURA 21).

O ncRNA *IsrG* possui 88 nucleotídeos e possui sequência abaixo:

IsrG undefined product 3775631:3775718 reverse

5'GGCTGCGGCAACCCAGGCCGCTGCTGCGGCAGCACCCGCACCTGCGGCGG
CGTCTGCCTCTACTACGGCAGCGCCTGCTGCTGCAGCC3'

O ncRNA *IsrG* está co-localizado com os genes *fimV* e *asd*, e com o *Hs_noco_697* (sentido sense). A localização e o perfil de expressão observado no RNAseq estão mostrados na FIGURA 22.

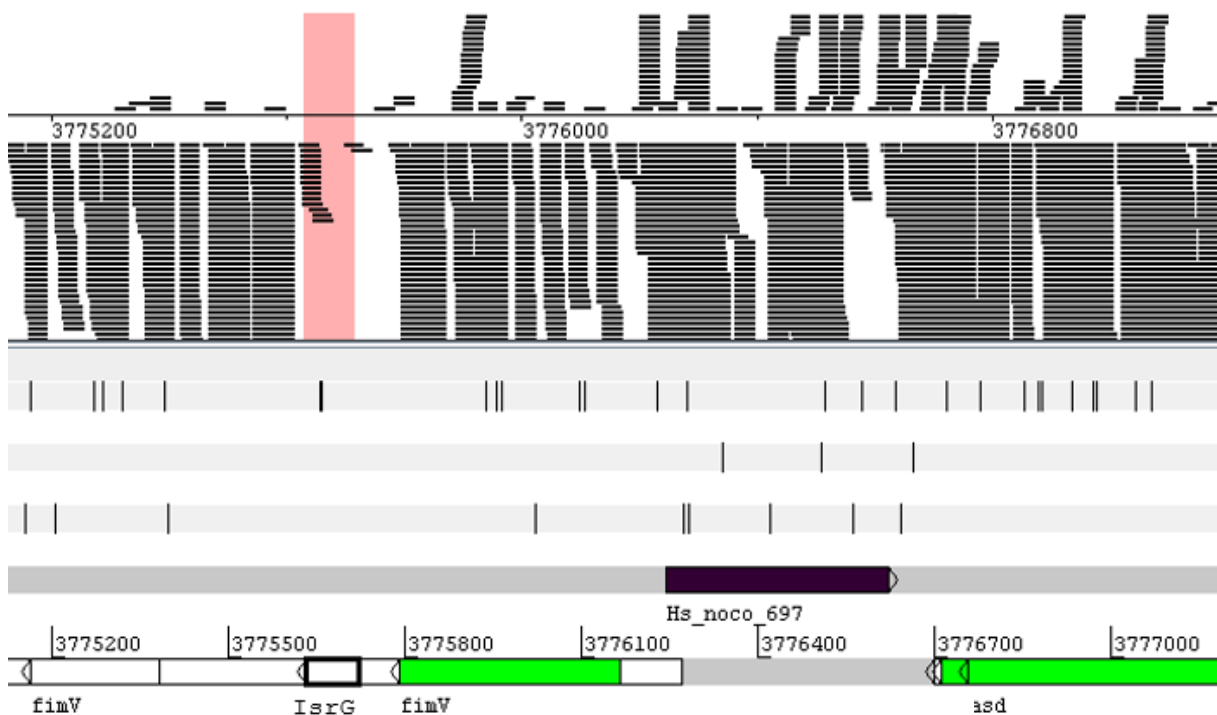


FIGURA 22 – Visualização da posição e do perfil de expressão do ncRNA *IsrG* no genoma de *H. seropedicae* SmR1. A sequência foi localizada utilizando o programa Artemis. As leituras estão delimitadas pelo retângulo rosa. FONTE: AUTOR (2016).

Em *Salmonella typhimurium* está envolvido na reposta ao choque frio e ao pH ácido, na fase estacionária do crescimento (ORTEGA, GONZALO-ASENSIO, PORTILLO, 2012). Em *H. seropedicae* SmR1 *IsrG* está expresso em apenas uma das condições ensaiadas. Na condição NFbHPN Malato-Alto Oxigênio foi observado o valor 38,541 RPKM (FIGURA 23).

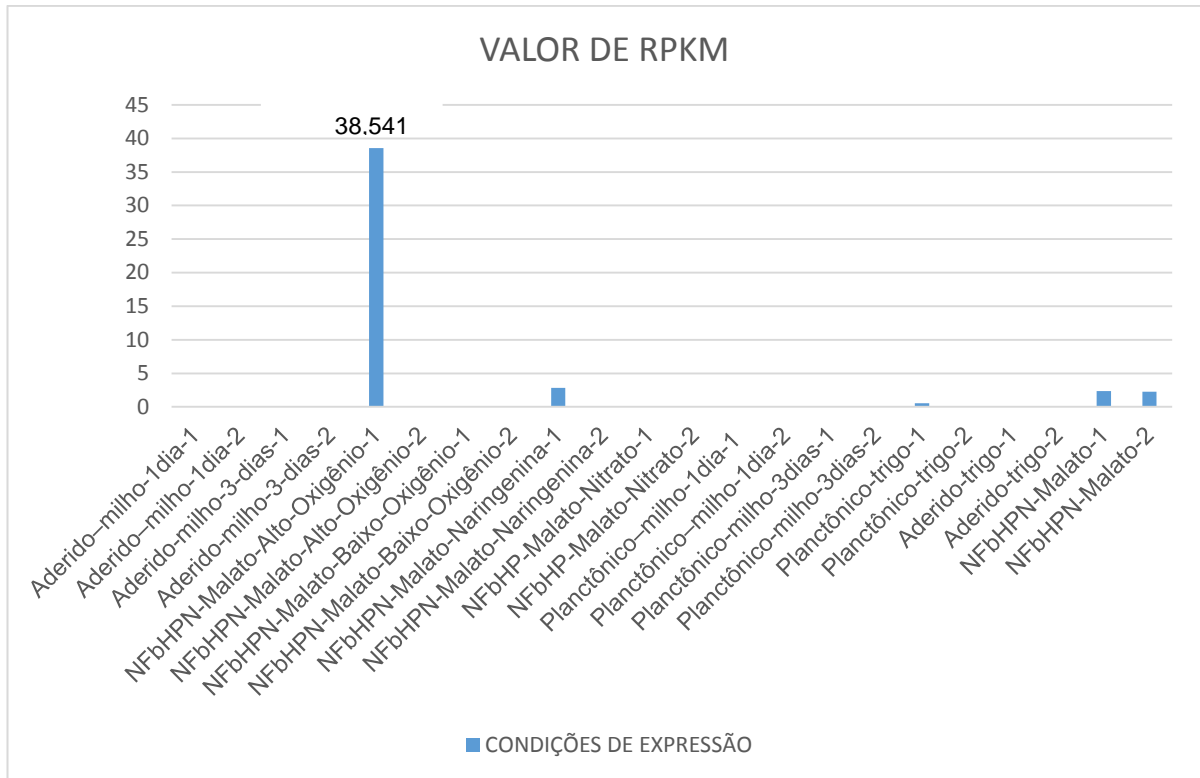


FIGURA 23 – Visualização dos níveis de expressão do ncRNA *isrG*, em ensaios de RNA-seq, em *H. seropedicae* SmR1. Em cor, estão representadas as 11 condições experimentais analisadas, onde 1 e 2 representam d.a.i (dias após inoculação). Os números apresentam o valor da expressão gênica em RPKM. FONTE: AUTOR (2016).

Segundo a literatura, genes da família ncRNA-*isr* apresentam sequência promotora sobreposta na extremidade 5' ou na extremidade 3' do gene vizinho (PADALON-BRAUCH et al., 2008). Em *H. seropedicae* SmR1 foi observado que o gene *IsrD* (FIGURA 21) está sobreposto à extremidade 3' do gene *Hsero_0991*. Já para o *IsrG*, expresso na condição NFbHPN Malato-Alto Oxigênio (FIGURA 23), foi observado que ele se sobrepõe ao gene *fimV*.

5.3.7 6S ncRNA (*housekeeping* ncRNA)

RNAs não codificadores considerados essenciais (*housekeeping* ncRNAs), como o 6S ncRNA, também foram identificados pela ferramenta *Infernal 1.1.1* (ITEM 4.5.4). Esse ncRNA já havia sido predito no trabalho de Moreno (2013) (QUADRO 2). A sequência do ncRNA 6S de *H. seropedicae* SmR1 está mostrada abaixo e sua localização no genoma, bem como a expressão observada no RNAseq, estão mostrados na FIGURA 24.

6S undefined product 1383620:1383786 reverse

```
5'CGTGTTTCGCGATTGCCATATATTCCTTGAACCATTTACGAGCATAGGTTGTGG
GATCGTTTGATGGGCGTGAGCGTCGCTAGTCCGATGAACCCGAAATTGAACTAC
TGCAACCAACTTGAACCGTCAGGTTTCAGGATGCCGGCACAGCGGCACAGGCGG
GGCTTGA3'
```

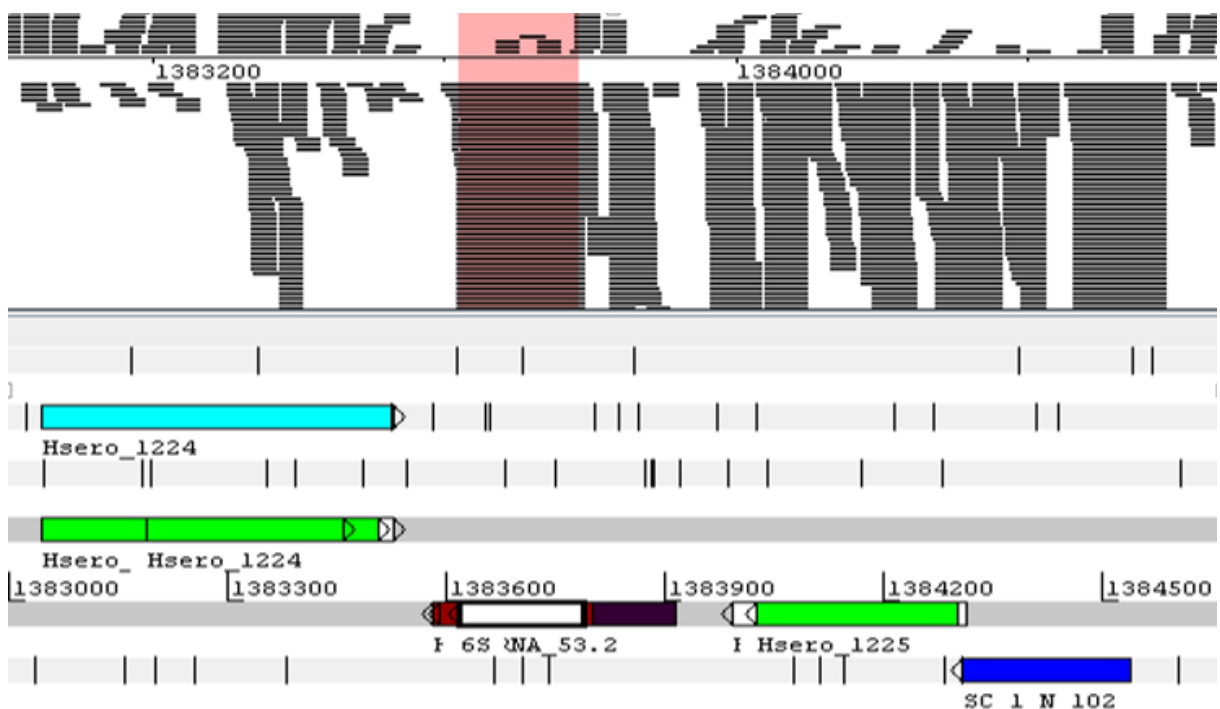


FIGURA 24 – Visualização da posição e do perfil de expressão do ncRNA 6S no genoma de *H. seropedicae* SmR1. A sequência foi localizada utilizando o programa Artemis. As leituras estão delimitadas pelo retângulo rosa. FONTE: AUTOR (2016).

A predição obtida por Moreno (2013) indicava uma sequência com 323 bases e a sequência predita pelo Infernal foi menor, com 177 bases. Como visto na FIGURA 24, o alinhamento das leituras de sequenciamento de RNAseq indicam que a predição atual é mais precisa que aquela apresentada anteriormente.

O ncRNA 6S em *E. coli* desempenha a função de regular a atividade da RNA polimerase (RNAP), que é dependente do fator sigma 70 necessário para o reconhecimento do promotor e iniciação da transcrição (WASSARMAN e STORZ, 2000).

Destaca-se que o nível de expressão observado é elevado, em relação aos seus vizinhos (*Hsero1224* e *Hsero1225*), nas condições ensaiadas, e isto concorda com o fato deste ncRNA ser essencial. Além disso, o ncRNA 6S está expresso em todas as condições ensaiadas, com destaque para a condição Aderido-trigo-1, em que foi observado o valor de 15.0374,531 RPKM (FIGURA 25).

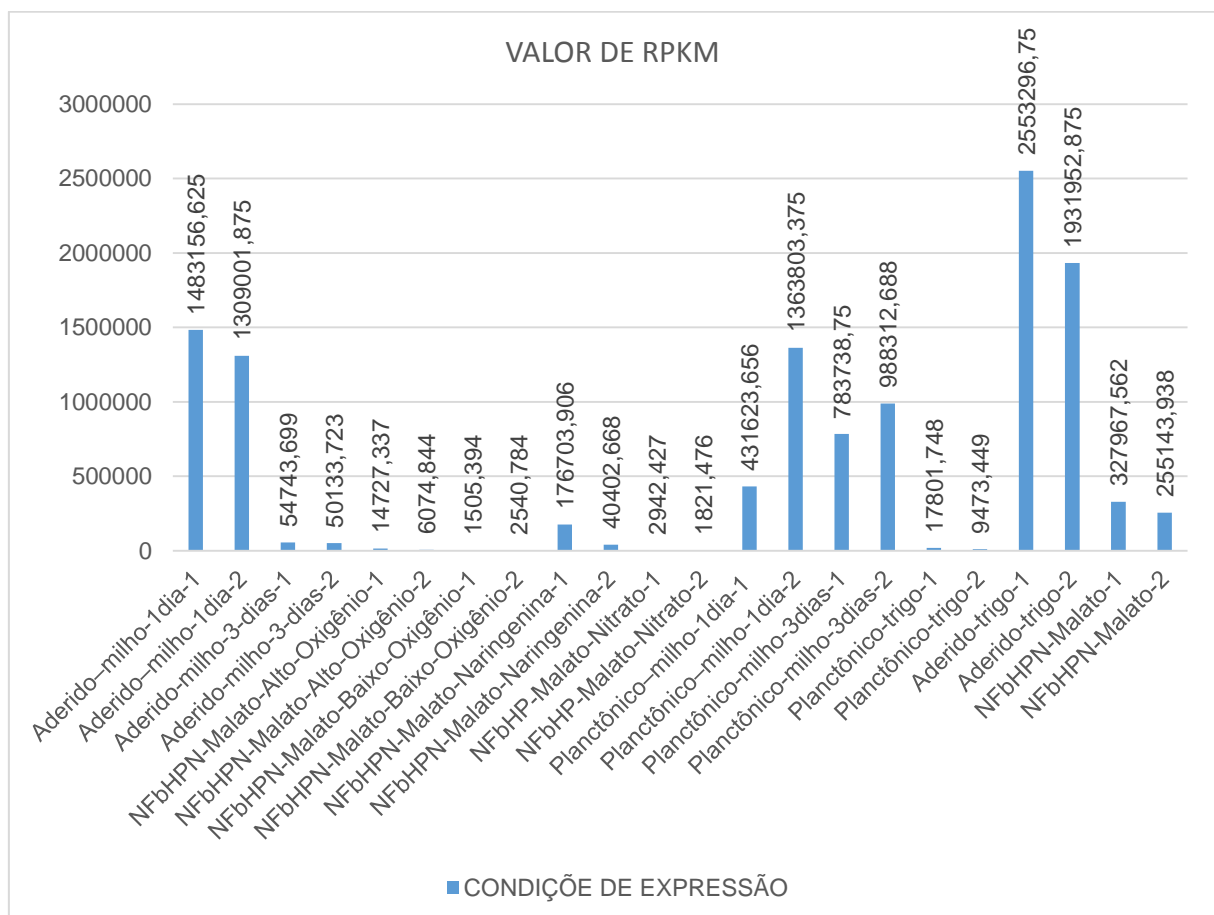


FIGURA 25 – Visualização dos níveis de expressão do ncRNA 6S, em ensaios de RNA-seq, em *H. seropedicae* SmR1. Em cor, estão representadas as 11 condições experimentais analisadas, onde 1 e 2 representam d.a.i (dias após inoculação). Os números representam o valor da expressão gênica em RPKM. FONTE: AUTOR (2016).

Dada a importância desse ncRNA, foi realizada a análise comparativa da estrutura secundária do 6S de *H. seropedicae* SmR1 com o 6S de *Janthinobacterium* spp. e *Burkholderia* CCGE1003. A análise foi realizada com a ferramenta RNAFold (ITEM 4.5.6). Com base nas interações termodinâmicas RNA-RNA foram gerados gráficos de pontos (*dot plot*) e as estruturas secundárias dos ncRNA 6S foram previstas (FIGURA 26). Foi realizado o alinhamento múltiplo de sequência pela ferramenta CLUSTALW (ITEM 4.5.3) (FIGURA 27). Observa-se que há diferença na estrutura secundária do ncRNA 6S entre essas espécies mesmo que elas apresentem um elevado grau de conservação na sequência de nucleotídeos alinhada (FIGURA 27).

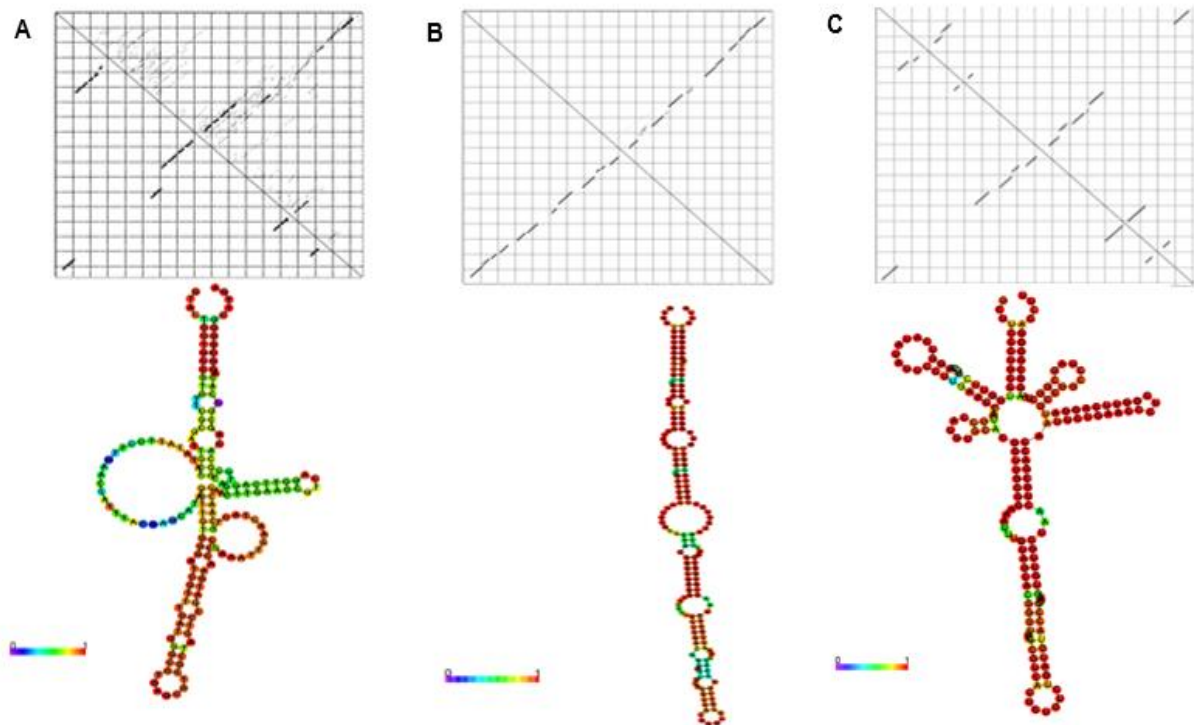


FIGURA 26 – Visualização do gráfico de pontos e da estrutura secundária do ncRNA 6S. A) *H. seropedicae* SmR1; B) *Janthinobacterium* sp; C) *Burkholderia* CCGE1003. Análise preditiva das estruturas secundárias de ncRNAs identificadas foi realizada utilizando o RNAFold. FONTE: AUTOR (2016).

```

Hsero          GUACUCCUGCCGUGUUCGCGAUUGCCAUAUAUCCUUGAACCAUU--UACGAGCAUAGGU
Janthinobacteri GCGCUCUCCUGCCGUGUUCGCCAUUGCCAUAUAAAUCCUUGAACCAUU--UAUGCGCAGCGGC
Burkholderia   GCUUCCCUGCCGUGUUCGCCAAGGUCAUUAUAUCCUUGAACCAUUGCCAUGUGCA-CGGU
                *      *      *      *      *      *      *      *      *      *      *      *      *
                *      *      *      *      *      *      *      *      *      *      *      *      *

Hsero          UGUGGGA--UCGUUUGAUGGGCGUGAGCGUCGCU-AGUCCGAUGAACCCGAAAUUGAACU
Janthinobacteri UGCGGAAUUUUUGUUCGACGGGAGUGCGCGUCGCU-AGUCCGAUGAACCCGAAAUUGAACU
Burkholderia   UGCGGAAUUUGUAGCACGGGUGUGCGCGUCACUCUGUCUGAUGUACCCGAAAGUGCUGCU
                ** **.*  * **  * **  * **  * **  * **  * **  * **  * **  * **
                ** **.*  * **  * **  * **  * **  * **  * **  * **  * **

Hsero          -ACUGCAACCAACUUGAACCGUCAGGUUCAGGAUGCCGGCACAGCGGCACAGGCGGGGCU
Janthinobacteri AACUGUGGCCAACUUGAACCUUUGGUUCAGGAUGCCGGCAAAGCGGCACUGGCGGGACU
Burkholderia   AACUGCGACCAAUUCUGAACCUUCGGUUCAGGAUGCCGGCCUAGCGGCUAAGGCGGGGAC
                ****  .* **** *  .* *  *****  *****  *****

Hsero          UGA
Janthinobacteri UG-
Burkholderia   CAU

```

FIGURA 27 – Alinhamento múltiplo de sequências do gene do ncRNA 6S pela ferramenta CLUSTALW. Hsero = *Herbapirillum seropedicae* SmR1; Janthinobacteri = *Janthinobacterium* sp. e Burkholderia = *Burkholderia* CCGE1003.

5.4 Predição dos alvos dos ncRNAs em *Herbaspirillum seropedicae* SmR1

A ferramenta computacional TargetRNA2 (ITEM 4.5.5) foi utilizada para predição de alvos dos ncRNAs em *H. seropedicae* SmR1 e como resultado observou-se uma grande variedade, independentemente do ncRNA analisado. Esse programa é específico para a predição de mRNA alvos de ncRNAs com atividade em *trans* (KERY et al., 2014).

O número de alvos para cada ncRNA foi variável. Por exemplo, Hs_noco_494 com 11 alvos, Hs_sRNA_Gs6 com 45 alvos e Hs_noco_749 com 99 alvos. E para alguns ncRNAs, como o Hs_noco_471, não houve predição de alvos (APÊNDICE II e Material Suplementar).

Os alvos mRNAs preditos em *H. seropedicae* SmR1 estão relacionados a diferentes mecanismos bioquímicos, por exemplo, alguns alvos do ncRNA 4.5S que estão anotados: *argB* (acetilglutamato quinase), *Hsero_3355* (lipoproteína), *glk* (glucoquinase), *htpX* (heat shock proteína HtpX), *Hsero_2759* (proteína reguladora de transcrição).

Dentre a relação de mRNA para os alvos preditos, foi realizada a análise do provável sítio de interação ncRNA-RNA alvo, com ênfase naqueles casos em que essa interação ocorre próxima ou sobre a região do RBS. Em *H. seropedicae* SmR1 foi considerada a sequência RBS consenso e suas variações, sendo o núcleo do RBS AGGA e a sequência CAAGGACA.

Tomando-se como exemplo a sequência nucleotídica do *Hs_noco_530* que possui 40 possíveis alvos, dentre eles o gene *serS* que codifica uma *seril-tRNA sintetase*:

Hs_noco_530	36	3'	AGCAAGAAAGAAGGGCUGCG	5'	17
			:		
serS	-15	5'	UCGUAUUUUCUCCCAUGA	3'	4

Observa-se que o pareamento de bases entre o Hs_noco_530 e o mRNA *serS* acontece na região do RBS desse último. Isto é um forte indicativo que o ncRNA Hs_noco_530 regula negativamente a tradução da *seril-tRNA sintetase*.

O 4.5S RNA é um ncRNA comum em bactérias e tem a função de direcionar proteínas contendo peptídeo sinal para o aparato de secreção, formando uma

partícula reconhecedora de sinal com a proteína Fhh. Além disso, o 4.5S RNA participa da tradução ao interagir com Fator de alongamento G (NAKAMURA et al., 2001).

Em *E. coli* o 4.5S RNA é codificado pelo gene (*ffs*) e contém 114 nucleotídeos. É a molécula bifuncional que está envolvida na tradução e secreção de proteína pela ligação ao fator de alongamento EFG e proteína Ffh, respectivamente (JAGATH et al., 2001; MIKULÍK, 2003). Aparentemente, o 4.5S RNA facilita a liberação do EF-G-GTP do ribossomo por competir com o 23S rRNA pela ligação com o EF-G (NAKAMURA et al., 2001; GU et al., 2005). Buskiewicz e colaboradores (2005) analisaram as funções da partícula reconhecedora de sinal SRP em *E. coli* com base na ligação da proteína Fth ao 4.5S RNA e do complexo Fth-4.5S, sugeriram que Fth livre os sítios de ligação 4.5S RNA e FtsY são ocluídos por interações domínio-domínio fortes que devem ser interrompidas para a formação SRP ou do complex Fth-FtsY.

O 4.5S RNA de *H. seropedicae* Smr1 contém 146 nucleotídeos a sequência extraída pelo programa Infernal.1.1.1 (ITEM 4.5.4) está descrita abaixo.

```
5'TGGCCCGGTCGGTGTGATTTGAGATAACAATTCACCCCGGCGGGCCCCTGCGC
ATTGTGGCATGGTCAACCTGGTCAGGTCGGGAACGAAGCAGCCACAGCCATTT
CCCGCAAGTGCCGCAGACAAGGCTCGCCTCCTTCCGTTTTT3'
```

Dentre os alvos do 4.5S, preditos pelo Target RNA2, está o mRNA *mutL* que codifica uma proteína envolvida no reparo do DNA (Figura 31). Novamente, o pareamento do ncRNA ocorre próximo à do RBS do RNA alvo iniciando, neste caso específico, na terceira base a montante da adenosina do códon AUG.

```
4.5S      85      3' AAGGGCUGGACUGGU 5'      71
          |:||||| | | | :
mutL     -17     5' CUUCCGACCCGACCG 3'     -3
```

A análise do ncRNA 4.5S realizada pela ferramenta TargetRNA2 (ITEM 4.5.5), resultou em 77 mRNAs alvos e dentre eles destacam-se aqueles que já foram anotados mostrados na (FIGURA 28).

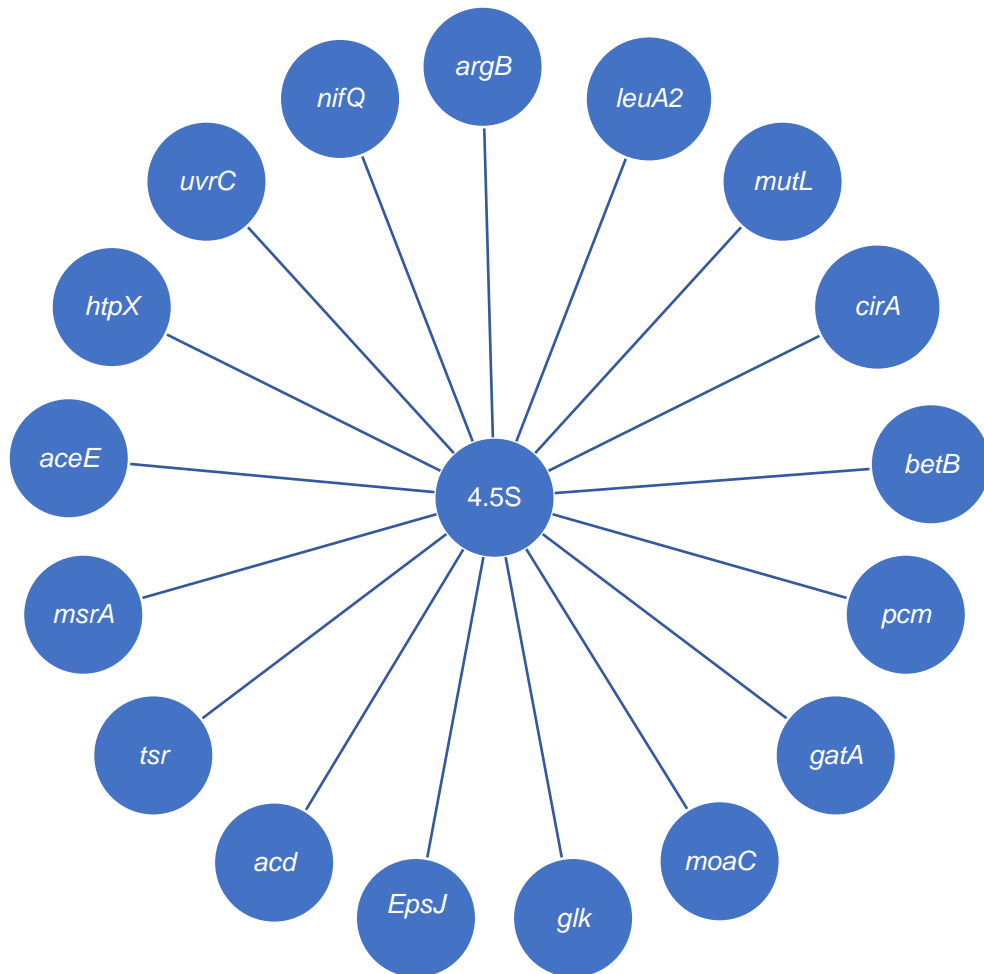


FIGURA 28 – Visualização dos alvos ncRNA 4.5S com o nome do gene em genoma de *H. seropedicae* SmR1. A predição dos alvos pela ferramenta TargetRNA2. FONTE: AUTOR (2016).

Nesse conjunto consta o gene *nifQ* cujo produto está envolvido no metabolismo de nitrogênio.

```

4.5S   95   3' CCG-ACGAAGCAAG-GGCUGG 5'   77
          || | :| | | | | | | | :| | |
nifQ  -41  5' AGCUUUUUUCGUUCUCUGACA 3'   -21

```

Pode-se observar que o pareamento predito inicia 21 bases a montante do códon AUG e por isso fora da região RBS. Comparado ao exemplo anterior tem-se, segundo o TargetRNA2, tem-se um valor de energia de interação para o pareamento de -16,21 para *mutL* e de -8,31 para *nifQ*. Esses dois aspectos, posição do pareamento e valor de energia de interação, fazem de *nifQ* um alvo improvável para 4.5S.

5.5 Predição de CRISPR em *Herbaspirillum seropedicae* SmR1

Uma nova categoria de RNA não codificador, o RNA CRISPR, foi descrita tanto para o domínio *Bacteria* quanto para o domínio *Archaea* (MOJICA et al., 2000; JANSEN et al., 2002). A transcrição de CRISPR resulta em pequenos fragmentos de RNA que reconhecem um DNA exógeno específico e que podem orientar a nuclease Cas, responsável pela clivagem desse DNA caso esse entre novamente em contato com o microrganismo. Desta forma, os procaríotos apresentam um mecanismo de defesa contra DNAs invasores (BARRANGOU e MARRAFFINI, 2014).

Neste trabalho, as ferramentas CRISPRFinder e CRISPRmap (ITEM 4.9) foram utilizadas para predição de CRISPR no genoma de *H. seropedicae* SmR1. Foi identificado um locus CRISPR com 164 nucleotídeos divididos em Repetições Palindrômicas DR (54 nucleotídeos) e Região Espaçadora (57 nucleotídeos) (FIGURA 29).

5'GTCCTTCTTCAGGCAGGACGACATGAAGGCCTTGCGTTCGTCCG CTTTCTTGCC
CGCGGCGTCAGCATTGCACTGCTTCATCTTGTCCTGCTGGGTGGCTGGCTTGGGCGC
 GTCCTTCTTCAGGCAGGACGACATGAAGGCCTTGCGTTCGTCCCCTTTCTTGCC3'

FIGURA 29 – Visualização do locus CRISPR no genoma de *H. seropedicae* SmR1. Sentido sense no genoma, com 164 nucleotídeos divididos em Repetições Palindrômicas DR (54 nucleotídeos) e Região Espaçadora (57 nucleotídeos) sublinhada. FONTE: AUTOR (2016).

O locus CRISPR está co-localizado com o gene *rhtB* (no sentido sense) e com os genes *psiF* e *Hsero_1878* (no sentido antisense) (FIGURA 30).

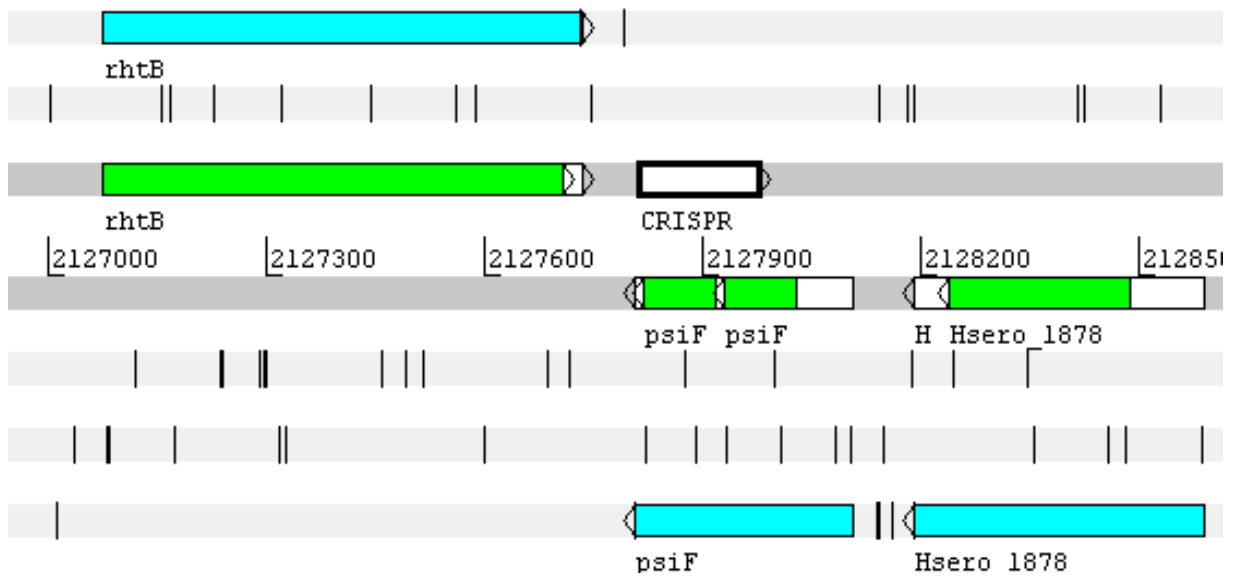


FIGURA 30 – Visualização da posição e do perfil de expressão do CRISPR no genoma de *H. seropedicae* SmR1. A sequência foi localizada utilizando o programa Artemis. As leituras estão delimitadas pelo retângulo rosa. FONTE: AUTOR (2016).

Em relação à expressão, o CRISPR está expresso em algumas condições, sendo observado o maior nível de expressão na condição NFbHPN-Malato-Alto-Oxigênio-1 e NFbHPN-Malato-Alto-Oxigênio-2, com 1429,72 RPKM e 941,335 RPKM, respectivamente (FIGURA 31).

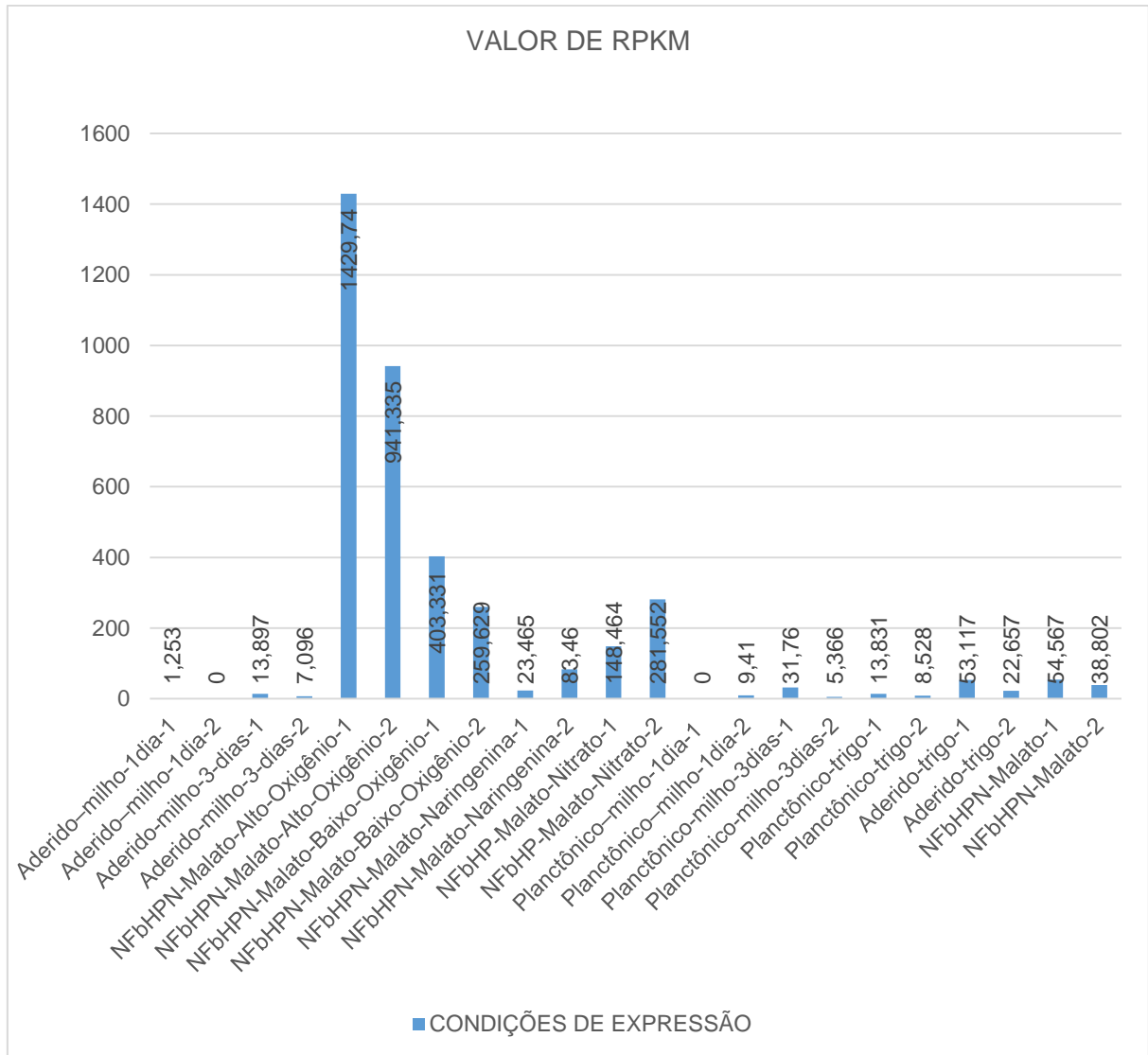


FIGURA 31 – Visualização dos níveis de expressão do CRISPR, em ensaios de RNA-seq, em *H. seropedicae* SmR1. Em cor, estão representadas as 11 condições experimentais analisadas, onde 1 e 2 representam d.a.i (dias após inoculação). Os números representam o valor da expressão gênica em RPKM. FONTE: AUTOR (2016).

Além disso, a ferramenta CRISPRFinder identificou a proteína associada ao CRISPR, a RAMP da Família Cmr4 em *H. seropedicae* SmR1, que está relacionada a cas_RAMP_Cmr4 que participa do processo de desenvolvimento da memória imunológica. O complexo Cas RAMP modula a clivagem de RNAs (HALE et al., 2012).

5.6 Predição ncRNAs em outras espécies de *Herbaspirillum*

Na predição de ncRNAs em outras espécies de *Herbaspirillum*, ou estirpes da espécie *H. seropedicae*, foi utilizada a ferramenta computacional Infernal 1.1.1 (ITEM 4.5.4). Os genomas completos de *H. seropedicae* Z67, *H. lusitanum* P6-12, *H. hiltineri* N3, *H. frisingense* GSF30, *Herbaspirillum*spp. estirpe B65, *H. rubrisubalbicans* M1, *H. autotrophicum* IAM 14942, *H. huttiense subsp puttei*, *Herbaspirillum*spp. estirpe B501, *Herbaspirillum*spp. estirpe GW103, *Herbaspirillum*spp. estirpe RV423, *H. rhizosphaerae* UMS-37, *Herbaspirillum*spp. estirpe Os34, *Herbaspirillum*spp. estirpe Os45 e *H. seropedicae* AU14040 foram analisados. Os resultados estão listados no Quadro 3. Por esta abordagem não foram detectados ncRNAs em *H. lusitanum* P6-12 e *Herbaspirillum* spp. estirpe B65. *H. seropedicae* Z67, *H. hiltineri* N3, *H. frisingense*, *H. rubrisubalbicans* M1 e *H. seropedicae* AU14040 foram os organismos que apresentaram maior número de ncRNAs comuns a *H. seropedicae* SmR1. Os ncRNAs ctRNA_p42d, Betaproteobacteria_toxic_sRNA e P10 são exemplos de ncRNAs presentes em *Herbaspirillum*spp. RV423, *H. rhizosphaerae* UMS-37, e *Herbaspirillum*spp. OS34 e O45 e comuns a *H. seropedicae* SmR1.

No QUADRO 3 estão apresentados os ncRNAs nas diferentes espécies de *Herbaspirillum*spp. e diferentes daqueles observados em *H. seropedicae* SmR1.

Esses resultados sugerem que os ncRNAs preditos no gênero *Herbaspirillum* spp. há uma conservação no alinhamento de sequência dos ncRNAs comuns, por exemplo,

Betaproteobacteria_toxic_sRNA, pfl, sX4 e alpha_tmRNA, ar35, tmRNA, dentre outros, que se repetem em todas as espécies com exceção em *Herbaspirillum*spp. B65, *Herbaspirillum*spp. B501 e *Herbaspirillum lusitanum* P6-12. A caracterização dos ncRNAs preditos no gênero *Herbaspirillum* spp. sugere que apesar da diferença dentro destas espécies, há preservação de sequência dentre os ncRNAs comuns.

Genêro <i>Herbaspirillum</i>	ncRNAs
<i>Herbaspirillum seropedicae</i> Z67	ohsC_RNA, Rhizobiales-2 e UPD-PKc
<i>Herbaspirillum lusitanum</i> P6-12	Ausência
<i>Herbaspirillum hiltneri</i> N3	ar14, Acido-1, Alpha_RBS, ar14, asX2, asX3, Bp2_287, csRNA, isrN, istR, Ms_AS-4, Ms_IGR-8 ncr1175, NsiR1, P26, P5, PrrB_RsmZ, RsmY, RydC, sau-5971, SpF41_sRNA, Spot_42, Xoo5, sraA, sX9, TB10Cs4H2, TB11Cs5H2, TB9Cs1H1
<i>Herbaspirillum frisingense</i> GSF30	Entero_5_CRE, P10 e STnc430
<i>Herbaspirillum</i> spp. estirpe B65	snoZ248
<i>Herbaspirillum rubrisubalbicans</i> M1	ALIL, Alpha_RBS, Archaea_SRP, asX2, b55, BjrC1505, C4, ceN84, istR, K_chan_RES, MAT2A_B, MEG3_2, NRF2_IRES, OrzO-P, P36, pRNA, PtaRNA1, rli60, rox2, rseX, sau-6072, sau-63, SpF25_sRNA, SpR10_sRNA, sroB, sX15, TB9Cs1H1 e TB9Cs1H2.
<i>Herbaspirillum huttiense</i> subsp. <i>putei</i> estirpe IAM 15032	asX2, Cardiovirus_CRE, Chloroflexi-1, MEG3_2, SBWMV2_UPD-PKk, sroB, veev_FSE, asX2, Cardiovirus_CRE, Chloroflexi-1, MEG3_2, SBWMV2_UPD-PKk, sroB e veev_FSE.
<i>Herbaspirillum autotrophicum</i> IAM 14942	MIR1444 e SpF10_sRNA.
<i>Herbaspirillum</i> spp. estirpe B501	Ausência
<i>Herbaspirillum</i> spp. estirpe GW103	sX2, Flavi_CRE, RyeB, SAM-IV, sau-6072 e sroB.
<i>Herbaspirillum</i> spp. estirpe RV1423	psRNA2, sbcD e STnc40 .
<i>Herbaspirillum rhizosphaerae</i> UMS-37	NRF2_IRES, NsiR1, Pseudomon-groES e sX9 .
<i>Herbaspirillum rubrisubalbicans</i> spp. estirpe Os34	ryfA, SpR20_sRNA e asX2.
<i>Herbaspirillum rubrisubalbicans</i> spp. estirpe Os45	asX2, ROSE_2 e ryfA.
<i>Herbaspirillum</i> AU14040	Alpha_RBS, Atu_C8, Atu_C9, eeev_FSE, GRIK4_3p_UTR, IRES_Bip, IRES_L-myc, Ms_AS-4, Ms_IGR-4, ncr1241, NRF2_IRES, ohsC_RNA, P37, radC, RF_site2, Rhizobiales-2, SpF36_sRNA, SpF41_sRNA, TB9Cs1H1, UPD-PKc e Xoo8.

QUADRO 3- ncRNAs preditos para bactérias do gênero *Herbaspirillum* e diferentes daqueles já conhecidos em *H. seropedicae* SmR1. Foi utilizada a ferramenta computacional Infernal 1.1.1. FONTE: AUTOR (2016).

5.7 Seleção de ncRNAs para validação e desenhos dos primers para PCR e qRT-PCR

O critério de seleção de ncRNAs de *H. seropedicae* SmR1 para validação baseou-se nos perfis de expressão dos ncRNAs nas 11 condições experimentais de RNA-seq (TABELA 3). Oligonucleotídeos específicos para cada ncRNA foram desenhados usando o programa Thermo Fisher. Os primers foram desenhados para a realização da reação em cadeia da amplificação da polimerase a PCR (do inglês, *Polymerase Chain Reaction Amplification*) e a reação quantitativa em tempo real (do inglês, *Quantitative Real-Time PCR*) (PFAFFL, 2001; BUSTIN et al., 2009; KOZERA e RAPACZ, 2013; LIU et al., 2013). Os primers universal e reverso para PCR alinham um na região codificante do ncRNA e outro na região intergênica, com objetivo de excluir a hipótese que o ncRNA pode ser maior do que o predito (QUADRO 4). Para o ensaio de qRT-PCR os primers universal e reverso foram desenhados considerando-se somente a sequência codificadora predita (QUADRO 5). Para descartar a eventualidade de alinhamento com outras sequências foi realizado o BlastN.

ncRNAs	Universal	Reverso
4.5S	CGGTCGGTGTGATTTGAGAT	AAATGGCTGTGGCTGCTT
CUFF.1109	AACTTGTGCAATCCGCTAA	TCACCTTTCTTCGCGCTTC
Hs_noco_732	TGCAAACGACGAATCTTACG	CGGACGCACCGCTTCGACA
small_49.2	AGCTTCGGCACTGACGTTTA	GGGAGGAACCTTGCACAGTA
CUFF.1011	GAGATGGCATTCTCCTCAA	CCAGGTCACTGTAGGCATCA
Hs_noco_530	GTCATCGACGCTCGGGAA	GATGCCGAATTCCTTGCT
Hs_noco_517	GTATGTAAGCATTAGTTATTA	CCGCATTTGCTGACTACT
Hs_noco_820	TTCTTGTCAGCCGCCAGCGC	AGGGAATCCGATGCAAAAAC
Hs_noco_824	TAATTTAGAAAGATAGTCCTG	CCCGGTTTTACTTTGCTGTT
Hs_noco_746	GAGGCTACCTGATCGCATTTC	TAAAGCAAGCCCGAAAACAG

QUADRO 4 – Oligonucleotídeos desenhados para PCR. FONTE: AUTOR (2016).

ncRNAs	Universal	Reverso
	AAAGGTACTGCGGGAGAA	TTTACCGACACCGACGAA
CUFF.1109	TCAAATAAAAAAGCGAATC	AACTTGTGCAATCCGCTAA
Hs_noco_732	TGCAAACGACGAATCTTACG	TAGGGAGTAGGTTATAATTG
small_49.2	CAGCAGGGTTTTTCGGGTTT	GGGAGGAACCTTGCACAGTA
CUFF.1011	TGACATTACGTTTCCTGCATC	CCAGGTCACTGTAGGCATCA
Hs_noco_530	AAGAACGAGCCAGTGACACC	AGTGTACTCACAAATCGGGA
Hs_noco_517	CCGCATTTGCTGACTACTGA	CCGCATTTGCTGACTACT
Hs_noco_820	TTCTGCTTGGCCATCTGGTC	GCGCTGGCGGCTGACAAGAA
Hs_noco_824	TTTTTCAACTTCGCTTCCTGT	CCCGGTTTTACTTTGCTGTT
Hs_noco_746	TAAAGCAAGCCCGAAAACAG	TAAAGCAAGCCCGAAAACAG

QUADRO 5 – Oligonucleotídeos desenhados para qRT-PCR. FONTE: AUTOR (2016).

6 CONCLUSÕES

Neste trabalho foram identificados 55 novos ncRNAs em *H. seropedicae* SmR1 com expressão observada em material de RNA-seq, ampliando assim a lista obtida em trabalhos anteriores. Os resultados indicam a existência de ncRNA *cis-encoded*, ncRNA *trans-encoded*, riboswitch e CRISPR em *H. seropedicae* SmR1.

A seleção de alguns ncRNAs e o desenho de oligonucleotídeos para qRT-PCR é o primeiro passo em direção à validação experimental de ncRNAs em *H. seropedicae* SmR1.

Os resultados obtidos neste trabalho permitem concluir que:

- O genoma de *H. seropedicae* SmR1 codifica ncRNAs que são expressos em diferentes condições experimentais e apenas 55 deles apresentam identidade de acordo com o banco de dados Rfam 12.0;
- Um *CRISPR* RNA foi predito em *H. seropedicae* SmR1 e expresso;
- *H. seropedicae* SmR1 possui os ncRNAs *MicC*, *IsrK*, *IsrD* e *IsrG* Hfq-dependentes;
- O número de mRNAs alvos preditos pela ferramenta TargetRNA2 para os ncRNAs de *H. seropedicae* SmR1 foi bastante variável e há também variação na posição da região de pareamento em relação ao códon AUG;
- A ferramenta Infernal 1.1.1 não detectou ncRNAs em *H. lusitanun* P6-12 e *Herbaspirillum* sp. B65.

REFERÊNCIAS

ABENDROTH, U.; SCHMIDTKE, C.; BONAS, U. Small non-coding RNAs in plant-pathogenic *Xanthomonas* RNA. *Biol*, v. 11, n. 5, p. 457-463, 2014.

ABREU-GOODGER, C; MERINO, E. RibEx: a web server for locating riboswitches and other conserved bacterial regulatory elements. **Nucleic. Acids. Res**, (Web Server issue): W690-2, 2005.

ALTUVIA, S. Identification of bacterial small non-coding RNAs; experimental approaches. **Curr Opin Microbiol**, v. 10, p. 257-261, 2007.

ALTSCHUL, S. F.; GISH, W.; MILLER, W.; MYERS, E. W.; LIPMANL, D. J. Basic local alignment search tool. **J. Mol. Biol**, p. 403–410, 1990.

AMITAI, G.; SOREK, R. CRISPR-Cas adaptation: insights into the mechanism of action. **Nat. Rev. Microbiol**, v. 14, n. 2, p. 67-76, 2016.

ANTHONY, P. C.; PEREZ, C. F.; GARCÍA-GARCÍA, C.; BLOCK, S. M. Folding energy landscape of the thiamine pyrophosphate riboswitch aptamer. **PNAS**, v. 109, n. 5, p. 1485-1489, 2012.

ARNVIG, K. B., CORTES, T., YOUNG, D. B. Noncoding RNA in *Mycobacteria*. **Microbiol Spectrum**, 2014.

ARAÚJO, N. D. D; FARIAS, R. P. D.; PEREIRA, P. B.; FIGUEIRÊDO, F. M.; MORAIS, A. M. B. D.; SALDANHA, L. C.; GABRIEL, J. E. A era da bioinformática: seu potencial e suas implicações para as ciências da saúde. *The effects and applications of bioinformatics on the biomedical area*. **Estud. Biol**, p. 143-148, 2008.

ARGAMAN, L., HERSHBERG, R., VOGEL, J., BEJERANO, G., WAGNER, E. G. H., MARGALIT, H., ALTUVIA, S. Novel small RNA-encoding genes in the intergenic region of *Escherichia coli*. **Current Biology**, v.11, n. 12, p. 941–950, 2001.

AZHIKINA, T. L., IGNATOV D. V., SALINA, E. G., FURSOV, M. V., KAPRELYANTS, A. S. Role of Small Noncoding RNAs in Bacterial Metabolism. **Biochemistry**, Moscow, v. 80, n. 13, p. 1633-1646, 2015.

BABSKI, J.; MAIER, L. K.; HEYER, R.; JASCHINSKI, K.; PRASSE, D.; JÄGER, D.; RANDAU, L.; SCHMITZ, R. A.; MARCHFELDER, A.; SOPPA, J. Small regulatory RNAs in Archaea. **RNA. Biol**, v. 11, n. 5, p. 484-493, 2014.

BABU, M.; BELOGLAZOVA, N.; FLICK, R.; GRAHAM, C.; SKARINA, T.; NOCEK, B.; GAGARINOVA, A.; POGOUTSE, O.; BROWN, G.; BINKOWSKI, A.; PHANSE, S.; JOACHIMIAK, A.; KOONIN, E. V.; SAVCHENKO, A.; EMILI, A.; GREENBLATT, J.; EDWARDS, A. M.; YAKUNIN, A. F. A dual function of the CRISPR-Cas system in bacterial antiviral immunity and DNA repair. **Mol. Microbiol**, v. 79, n. 2, p. 484-502, 2011.

BACKOFEN, R; HESS, W. R. COMPUTATIONAL PREDICTION OF SRNAS AND THEIR TARGETS IN BACTERIA. **RNA. BIOL**, v. 7, n. 1, p. 33-42, 2010.

BALDANI, J. I., BALDANI, V. L. D., SAMPAIO, M. J. A. M., DÖBEREINER, J. A fourth azospirillum species from cereal roots. *Anais da Academia Brasileira de Ciência*, v. 56, p. 365, 1984.

BALDANI, J. I.; BALDANI, V. L. D.; SELDIN, L.; DÖBEREINER, J. Characterization of *Herbaspirillum seropedicae* nov. sp. a root associated nitrogen fixing bacterium. **International. Journal. of. Systematic. Bacteriology**, v. 36, p. 86-93, 1986.

BALDANI, V. L. D.; BALDANI, J. I.; OLIVARES, F. L.; DÖBEREINER, J. Identification and ecology of *Herbaspirillum seropedicae* and the closely related *Pseudomonas rubusubalbica*. **Symbiosis**, v. 13, p. 65-73, 1992.

BALDANI, J. I.; POT, B.; KIRCHHOF, G.; FALSEN, E.; BALDANI, V. L. D. OLIVARES, F. L.; HOSTE, B. KERSTERS, K.; HARTMANN, A.G.; DOBEREINER, J. Emended description of *Herbaspirillum*; inclusion of *Pseudomonas rubrisubalbicans*, a mild plant pathogen, as *Herbaspirillum rubrisubalbicans* comb nov; and classification of a group of clinical isolates (EF group 1) as *Herbaspirillum* species 3. *International Journal of Systematic Bacteriology*, v. 46, p. 802-810, 1996.

BALSANELLI, E.; **Caracterização de fatores moleculares envolvidos na interação de *Herbaspirillum seropedicae* com gramíneas.** Tese de doutorado, Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Paraná. Curitiba, 2013.

BARRANGOU, R.; FREMAUX, C.; DEVEAU, H.; RICHARDS, M.; BOYAVAL, P.; MOINEAU, S.; ROMERO, D. A.; HORVATH, P. CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, v. 315, n. 5819, p. 1709-12, 2007.

BARRANGOU R; MARRAFFINI L. A. CRISPR-Cas systems: Prokaryotes upgrade to adaptive immunity. *Mol. Cell*, v. 54, n. 2, p. 234-244, 2014.

BARRICK, J. E.; CORBINO, K. A.; WINKLER, W. C.; NAHVI, A.; MANDAL, M.; COLLINS, J.; LEE, M.; ROTH, A.; SUDARSAN, N.; JONA, I.; WICKISER, J. K.; BREAKER, R. R. New RNA motifs suggest an expanded scope for riboswitches in bacterial genetic control. *PNAS*, v. 101, n. 17, p. 6421-6426, 2004.

BATISTA, B. M., WASSEM, R., PEDROSA, O. F., SOUZA, E. M., DIXON, R., MONTEIRO, R.A. Enhanced oxygen consumption in *Herbaspirillum seropedicae* fnr mutants leads to increased NifA mediated transcriptional activation. *BMC Microbiology*, 15:95, 2015.

BINNECK, E. As ômicas: integrando a bioinformação. *Revista. Biotecnologia*, n. 32, 2004.

BONATO, P. **Genes do metabolismo de Nitrato em *Herbaspirillum seropedicae*: regulação transcricional e análise funcional.** Dissertação (Mestrado em Bioquímica) - Setor de Ciências Biológicas, Universidade Federal do Paraná, Curitiba, 2012.

BRANTL, S. Regulatory mechanisms employed by *cis-encoded* antisense RNAs. **Current Opinion in Microbiology**, v. 10, p. 102–109, 2007.

BREAKER, R. R. Natural and engineered nucleic acids as tools to explore biology. **Nature**, v. 432, n. 7019, p. 838-845, 2004.

BREAKER, R. R. Ancient, giant riboswitches at atomic resolution. **Nat. Struct. Mol. Biol**, v. 19, n. 12, p. 1208-1209, 2012.

BRENNAN, R. G.; LINK, T. M. Hfq structure, function and ligand binding. **Curr. Opin. Microbiol**, v. 10, n. 2, p.125-33, 2007.

BRENES-ÁLVAREZ, M.; OLMEDO-VERD, E.; VIOQUE, A.; MURO-PASTOR, A. M. Identification of Conserved And Potentially Regulatory Small RNAs In Heterocystous Cyanobacteria. **Front. Microbiol**, v. 7, p. 48, 2016.

BROSIUS J, TIEDGE H. NOMENCLATURE. **RNA BIOL**, v. 1, n. 2, p. 81-83, 2004.

BROSIUS, J., RAABE, C. A. WHAT IS AN RNA? A TOP LAYER FOR RNA CLASSIFICATION. **RNA. BIOLOGY**, v. 13, n. 2, p. 140-144, 2016.

BUSKIEWICZ, I.; KUBARENKO, A.; PESKE, F.; RODNINA, M. V.; WINTERMEYER, W. Domain rearrangement of SRP protein Ffh upon binding 4.5S RNA and the SRP receptor FtsY. **RNA**, v. 11, n. 6, p. 947-957, 2005.

BUSCH, A.; RICHTER, A. S.; BACKOFEN, R. IntaRNA: efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. **Bioinformatics**, v. 24, n. 24, p. 2849-2856, 2008.

CAMPS, M. Modulation of ColE1-like plasmid replication for recombinant gene expression. **Recent. Pat. DNA, Gene. Seq.**, v. 4, n. 1, p. 58-73, 2010.

CAO, S.; XU, X.; CHEN, S. J. Predicting structure and stability for RNA complexes with intermolecular loop-loop base-pairing. **RNA**, v. 20, n. 6, p. 835-845, 2014.

CARRO, C.; RIVAS, R.; LEÓN-BARRIOS, M.; GONZÁLEZ-TIRANTE, M.; VELÁZQUEZ, M.; VALVERDE, A. *Herbaspirillum canariense* sp. nov., *Herbaspirillum aurantiacum* sp. nov., and *Herbaspirillum soli* sp. nov., isolated from volcanic mountain soil, and emended description of the genus. **International Journal of Systematic and Evolutionary Microbiology**, v. 62, p. 1300-1306.

CARVER, T.; BÖHME, U.; OTTO, T. D.; PARKHILL, J.; BERRIMAN, M. BamView: viewing mapped read alignment data in the context of the reference sequence. **Bioinformatics**, v. 26, n. 5, p. 676-677, 2010.

CARTER, E. L.; BOER, J. L.; FARRUGIA, M. A.; FLUGGA, N.; TOWNS, C. L.; HAUSINGER, R. P. Function of UreB in *Klebsiella aerogenes* urease. **Biochemistry**, v. 50, n. 43, p. 9296-9308, 2011.

CASWELL, C.C; GAINES, J. M; CIBOROWSKI, P; SMITH, D; BORCHERS, C. H; ROUX C. M; SAYOOD, K; DUNMAN, P. M; ROOP, II. R. M. Identification of two small regulatory RNAs linked to virulence in *Brucella abortus* 2308. **Mol. Microbiol**, v. 85, n. 2, p. 345-360, 2012.

CECH, G. M.; SZALEWSKA-PAŁASZ, A.; KUBIAK, K.; MALABIRADE, A.; GRANGE, W.; ARLUISON, V.; WĘGRZYN, G. The *Escherichia Coli* Hfq Protein: An Unattended DNA-Transactions Regulator. **Front. Mol. Biosci**, v. 3, p. 36, 2016.

CIAMPI, M. S. Rho-dependent terminators and transcription termination. **Microbiology**, v. 152, p. 2515–2528, 2006.

CIRINO, H. S. B. **Re-Anotação do Genoma de *Herbaspirillum seropedicae* SmR1 com Dados de Transcriptoma por RNA-Seq**. Dissertação (Mestrado em Bioinformática) – Programa de Pós-Graduação em Bioinformática, Setor de Educação Profissional e Tecnológica da Universidade Federal do Paraná, Curitiba, 2014.

CHEN S., LESNIK E. A., HALL T. A., SAMPATH R., GRIFFEY R. H. A bioinformatics based approach for the identification of small RNAs genes in the *Escherichia coli* genome. **Biosystems**, 65: 157–177, 2002.

CHEN, S.; ZHANG, A.; BLYN, L. B.; STORZ, G. MicC, a second small-RNA regulator of Omp protein expression in *Escherichia coli*. **J. Bacteriol**, v. 186, n. 20, p. 6689-6697, 2004.

CLARK, P. M.; LOHER, P.; QUANN, K.; BRODY, J.; LONDIN, E. R.; RIGOUTSOS, I. Argonaute CLIP-Seq reveals miRNA targetome diversity across tissue types. **Sci. Rep**, v. 4, p. 5947, 2014.

CORBINO, K. A.; BARRICK, J. E.; LIM, J.; WELZ, R.; TUCKER, B. J.; PUSKARZ, I.; MANDAL, M.; RUDNICK, N. D.; BREAKER, R. R. Evidence for a second class of S-adenosylmethionine riboswitches and other regulatory RNA motifs in alpha-proteobacteria. **Genome. Biol**, v. 6, n. 8, p. R70, 2005.

COVENTRY, A.; KLEITMAN, D. J.; BERGER, B. MSARI: multiple sequence alignments for statistical detection of RNA secondary structure. **PNAS**, v. 101, n. 33, p. 12102-12107, 2004.

COX, M. M., DOUDNA, J. A., O'DONNELL, M. **Biologia Molecular – Princípios e Técnicas**, 1ªed. Artmed, 2012.

CROMIE, M. J.; SHI, Y.; LATIFI, T.; GROISMAN, E. A. An RNA sensor for intracellular Mg(2+). **Cell**, v. 125, n. 1, p. 71-84, 2006.

DADZIE, I.; XU, S.; NI, B.; ZHANG, X.; ZHANG, H.; SHENG, X.; XU, H.; HUANG, X. Identification and characterization of a *cis-encoded* antisense RNA associated with the replication process of *Salmonella enterica* serovar Typhi. **PLoS. One**, v. 8, n. 4, p. e61308, 2013.

DANN, C. E. R. D.; WAKEMAN, C. A.; SIELING, C. L.; BAKER, S. C.; IRNOV, I.; WINKLER, W. C. Structure and mechanism of a metal-sensing regulatory RNA. **Cell**, v. 130, n. 5, p. 878-892, 2007.

DING, L.; YAKOTA, A. Proposals of *Curvibacter gracilis* gen. nov., sp. nov. and *Herbaspirillum putei* sp. nov. for bacterial strains isolated from well water and reclassification of [*Pseudomonas*] *huttiensis*, [*Pseudomonas*] *lanceolata*, [*Aquaspirillum*] *delicatum* and [*Aquaspirillum*] *autotrophicum* as *Herbaspirillum huttiense* comb. nov., *Curvibacter lanceolatus* comb. nov., *Curvibacter delicatus* comb. nov. and *Herbaspirillum autotrophicum*. comb. nov. International. **Journal. of. Systematic. and Evolutionary. Microbiology**, v. 54, p. 2223–2230, 2004.

DÖBEREINER, J.; BALDANI, V. L. D.; BALDANI, J. I. Como isolar e identificar bactérias diazotróficas de plantas não-leguminosas. Brasília, EMBRAPA-SPI; Seropédica, EMBRAPA-CNPAB. p. 60,1995.

DOBRIŤSA, A. P.; REDDY, M. C.; SAMADPOUR, M. Reclassification of *Herbaspirillum putei* as a later heterotypic synonym of *Herbaspirillum huttiense*, with the description of *H. huttiense* subsp. *huttiense* subsp. nov. and *H. huttiense* subsp. *putei* subsp. nov., comb. nov., and description of *Herbaspirillum aquaticum* sp. nov. **Int. J. Syst. Evol. Microbiol**, v. 60, (Pt 6), p.1418-1426, 2009.

EPSHTEIN, V.; MIRONOV, A. S.; NUDLER, E. The riboswitch-mediated control of sulfur metabolism in bacteria. **Proc. Natl. Acad. Sci**, v. 100, n. 9, p. 5052-5056, 2003.

ERMOLAEVA, M. D.; KHALAK, H. G.; WHITE, O.; SMITH, H. O.; SALZBERG, S. L. Prediction of Transcription Terminators in Bacterial Genomes. **Mol. Biol**, v. 301, p. 27-33, 2000.

EVGUENIEVA-HACKENBERG, E.; KLUG, G. New aspects of RNA processing in prokaryotes. **Curr. Opin. Microbiol**, 2011, v. 14, n. 5, p. 587-92, 2011.

FANER, M. A., FEIG A. L. Identifying and characterizing Hfq–RNA interactions. **Methods**, n. 63, p. 144–159, 2013.

FUCHS, R. T.; GRUNDY, F. J.; HENKIN, T. M. The S(MK) box is a new SAM-binding RNA for translational regulation of SAM synthetase. **Nat. Struct. Mol. Biol**, v. 13, n. 3, p. 226-233, 2006.

GAUTHERET, D.; LAMBERT, A. Direct RNA motif definition and identification from multiple sequence alignments using secondary structure profiles. **J. Mol. Biol**, v. 313, n. 5, p. 1003-1011, 2001.

GONG, H.; VU, G. P.; BAI, Y.; CHAN, E.; WU, R.; YANG, E.; LIU, F.; LU, S. A Salmonella small non-coding RNA facilitates bacterial invasion and intracellular replication by modulating the expression of virulence factors. **PLoS. Pathog**, v. 7, n. 9, p. e1002120, 2011.

GONZALO-ASENSIO, J.; ORTEGA, A. D.; RICO-PÉREZ, G.; PUCCIARELLI, M. G.; GARCÍA-DEL PORTILLO, F. A novel antisense RNA from the Salmonella virulence plasmid pSLT expressed by non-growing bacteria inside eukaryotic cells. **PLoS One**. v. 8, n. 10, p. e77939, 2013.

GOTTESMAN, S., STORZ, G. Bacterial Small RNA Regulators: Versatile Roles and Rapidly Evolving Variations. **Cold Spring Harb Perspect Biol**. v. 3, p. a003798, 2011.
Gordon, D., Abajian, C., Green, P. Consed: a graphical tool for sequence finishing. *Genome Res*, v. 8, n. 3, p. 195-202, Mar 1998.

GRIPENLAND, J.; NETTERLING, S.; LOH, E.; TIENSUU, T.; TOLEDO-ARANA, A.; JOHANSSON, J. RNAs: regulators of bacterial virulence. **Nat. Rev. Microbiol**, v. 8, n. 12, p. 857-866, 2010.

GRISSA, I.; VERGNAUD, G.; POURCEL, C. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. **Nucleic Acids Res**, (Web Server issue): W52-7, 2007.

GRUBER, A. R.; FINDEISS, S.; WASHIETL, S.; HOFACKER, I. L.; STADLER, P. F. RNAz 2.0: improved noncoding RNA detection. **Pac. Symp. Biocomput**, p. 69–79, 2010.

GRUNDY, F. J.; LEHMAN, S. C.; HENKIN, T. M. The L box regulon: lysine sensing by leader RNAs of bacterial lysine biosynthesis genes. **Proc. Natl. Acad. Sci**, v. 100, n. 21, p. 12057-12062, 2003.

GU, S.Q.; JÖCKEL, J.; BEINKER, P.; WARNECKE, J.; SEMENKOV, Y. P.; RODNINA, M. V.; WINTERMEYER, W. Conformation of 4.5S RNA in the signal recognition particle and on the 30S ribosomal subunit. **RNA**, v. 11, n. 9, p. 1374-1384, 2005.

GUEIROS-FILHO F. J, LOSICK, R. A widely conserved bacterial cell division protei that promotes assembly of the tubulin-like protein ftsz. **Genes Dev**, v. 16, n. 19, p. 2544-2556.

GUNDERSON, F. F.; CIANCIOTTO, N. P. The CRISPR-associated gene Cas2 of Legionella pneumophila is required for intracellular infection of amoebae. **M. Bio**, v. 4, n. 2, p. e00074-13, 2013.

GUTIÉRREZ, P.; LI, Y.; OSBORNE, M. J.; POMERANTSEVA, E.; LIU, Q.; GEHRING, K. Solution structure of the carbon storage regulator protein CsrA from Escherichia coli. **J. Bacteriol**, v. 187, n. 10, p. 3496-3501, 2005.

HALE, C. R.; MAJUMDAR, S; ELMORE, J; PFISTER, N.; COMPTON, M.; OLSON, S.; RESCH, A. M.; GLOVER, C. V. 3.R.D.; GRAVELEY, B. R.; TERNS, R. M.; TERNS, M. P. Essential features and rational design of CRISPR RNAs that function with the Cas RAMP module complex to cleave RNAs. **Mol. Cell**, v. 45, n. 3, p. 292-302, 2012.

HAN, Y.; LIU, L.; FANG, N.; YANG, R.; ZHOU, D. Regulation of pathogenicity by noncoding RNAs in bacteria. **Future. Microbiol**, v. 8, n. 5, p. 579-591, 2013.

HENKIN, T. M. Riboswitch RNAs: using RNA to sense cellular metabolism. **Genes Dev**, v. 22, n. 24, p. 3383-3390, 2008.

HERBIG, A.; NIESELT, K. nocoRNAc: characterization of non-coding RNAs in prokaryotes. **BMC. Bioinformatics**, v. 12, p. 40, 2011.

HIGGINS, D. G.; SHARP, P. M. CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. **Gene**, v. 73, n. 1, p. 237-244, 1988.

HOE, C. H.; RAABE, C. A.; ROZHDESTVENSKY, T. S.; TANG, T. H. Bacterial sRNAs: Regulation in stress. **Int. J. Med. Microbiol**, 2013.

HOLMQVIST, E.; WRIGHT, P. R.; LI, L.; BISCHLER, T.; BARQUIST, L.; REINHARDT, R.; BACKOFEN, R.; VOGEL, J. Global RNA recognition patterns of post-transcriptional regulators Hfq and CsrA revealed by UV crosslinking in vivo. **EMBO. J**, v. 35, n. 9, p. 991-1011, 2016.

HOOTON, S. P.; BRATHWAITE, K. J.; CONNERTON, I. F. 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**, v. 7, p. 355, 2016.

HÜTTENHOFER, A., SCHATTNER, P., POLACEK, P. Non-coding RNAs: hope or hype? **Trends genet**, v. 21, n. 5, p. 289-97, 2005.

IZAR, B.; MRAHEIL, M. A.; HAIN, T. Identification and Role of Regulatory Non-Coding RNAs in *Listeria monocytogenes*. **Int. J. Mol. Sci**, v. 12, p. 5070-5079, 2011.

IM, W. T.; BAE, H. S., YOKOTA, A.; LEE, S. T. *Herbaspirillum chlorophenicum* sp. nov., a 4-chlorophenol-degrading bacterium. **International. Journal of Systematic and Evolutionary Microbiology**, v. 54, p. 851–855, 2004.

IGNATOV, D.; MALAKHO, S.; MAJOROV, K.; SKVORTSOV, T.; APT, A.; AZHIKINA, T. RNA-seq analysis of *mycobacterium avium* non-coding transcriptome. **Plos. One**, v. 8, n. 9, p. e74209, 2013.

JAMES, E. K.; OLIVARES, F. L.; BALDANI, J. I.; DOBEREINER, J. *Herbaspirillum*, An endophytic diazotroph colonizing vascular tissue in leaves of Sorghum Bicolor L. Moench. **Journal Of Experimental Botany**, v. 48, n. 308, p. 785-797, 1997.

JAGATH, J. R.; MATASSOVA, N. B.; LEEUW, D. E.; WARNECKE, J. M.; LENTZEN, G.; RODNINA, M. V.; LUIRINK, J.; WINTERMEYER, W. Important role of the tetraloop region of 4.5S RNA in SRP binding to its receptor FtsY. **RNA**, v. 2, n. 7, p. 293-301, 2001.

JUNG, S.Y.; LEE, M. H.; OH, T. K.; YOON, J. H. *Herbaspirillum rhizosphaerae* sp. nov., isolated from rhizosphere soil of *Allium victorialis* var. *platyphyllum*. **International Journal of Systematic and Evolutionary Microbiology**, v. 57, p. 2284–2288, 2007.

KEILER, K. C.; SHAPIRO, L.; WILLIAMS, K. P. tmRNAs that encode proteolysis-inducing tags are found in all known bacterial genomes: a two-piece tmrna functions in caulobacter. **PNAS**, v. 97, n. 14, p. 7778-7783, 2000.

KERY, M. B.; FELDMAN, M.; LIVNY, J.; TJADEN, B. TargetRNA2: identifying targets of small regulatory RNAs in bacteria. **Nucleic. Acids. Res**, v. 42, p. 124-129, 2014.

KIRCHHOF, G.; ECKERT, B.; STOFFELS, M.; BALDANI, J. I.; REIS, V. M.; HARTMANN, A. *Herbaspirillum frisingense* sp. Nov., a new nitrogen-fixing bacterial species that occurs in c4-fibre plants. **Int. J. Syst. Evol. Microbiol**, v. 51, (pt 1), p. 157-68, 2001.

KINGSFORD, C. L., AYANBULE, K., SALZBERG, S. L. Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake. **Genome Biol**, v. 8, n. 2, p. R22, 2007.

KIMELMAN, A.; LEVY, A.; SBERRO, H.; KIDRON, S.; LEAVITT, A.; AMITAI, G.; YODER-HIMES, D. R.; WURTZEL, O.; ZHU, Y.; RUBIN, E. M.; SOREK, R. A vast collection of microbial genes that are toxic to bacteria. **Genome. Res**, v. 22, n. 4, p. 802-809, 2012.

KIM, J. Y.; MUKHERJEE, S.; NGO, L.C.; CHRISTIANI, D. C. Urinary 8-hydroxy-2'-deoxyguanosine as a biomarker of oxidative DNA damage in workers exposed to fine particulates. **Environ. Health. Perspect**, v. 112, n. 6, p. 666–671, 2004.

KOZERA, B.; RAPACZ, M. Reference genes in real-time PCR. **J. Appl. Genet**, v. 54, n. 4, p. 391-406, 2013.

KHOO, J. S.; CHAI, S. F.; MOHAMED, R.; NATHAN, S.; FIRDAUS-RAIH, M.; Computational discovery and RT-PCR validation of novel Burkholderia conserved and Burkholderia pseudomallei unique sRNAs. **BMC. Genomics**, v. 13 Suppl 7:S13, 2012.

LALAOUNA, D.; SIMONEAU-ROY, M.; LAFONTAINE, D.; MASSÉ, E. Regulatory RNAs and target mRNA decay in prokaryotes. **Biochim. Biophys. Acta**, v. 1829, n. 6-7, p. 742-747, 2013.

LAGIER, J. C.; GIMENEZ, G.; ROBERT, C.; RAOULT, D.; FOURNIER, P. E. 2012. Non-contiguous finished genome sequence and description of *Herbaspirillum massiliense* sp. **Standards in Genomic Sciences**, 7:200-209, 2012.

LANGE, S J.; ALKHNABASHI, O. S.; ROSE, D.; WILL, S.; BACKOFEN, R. CRISPRmap: an automated classification of repeat conservation in prokaryotic adaptive immune systems. **Nucleic. Acids. Res**, v. 41, n. 17, p. 8034-8044, 2013.

LAYTON, D. M.; BUNDSCHUH, R. A statistical analysis of RNA folding algorithms through thermodynamic parameter perturbation. **Nucleic. Acids. Res**, v. 33, n. 2, p. 519-524, 2005.

LECLERC, J. M.; DOZOIS, C. M.; DAIGLE, F. Role of the Salmonella enterica serovar Typhi Fur regulator and small RNAs RfrA and RfrB in iron homeostasis and interaction with host cells. **Microbiology**, v. 159, n. Pt 3, p. 591-602, 2013.

LEE, E. J.; GROISMAN, E. A. An antisense RNA that governs the expression kinetics of a multifunctional virulence gene. **Mol. Microbiol**, v. 76, n. 4, p. 1020-1033, 2010.

LENEVE, P. A. The structure of yeast nucleic acid. Iv. Ammonia hydrolysis. **Institute for Medical Research**, 1919.

LI, H.; HANDSAKER, B.; WYSOKER, A.; FENNELL, T.; RUAN, J.; HOMER, N.; MARTH, G.; ABECASIS, G.; DURBIN, R. The Sequence Alignment/Map Format and SAMtools. **Bioinformatics**, v. 25, n. 16, p. 2078-2079, 2009.

LI, S. K.; NG, P. K.; QIN, H.; LAU, J. K.; LAU, J. P.; TSUI, S. K.; CHAN, T. F.; LAU, T. C. Identification of small RNAs in Mycobacterium smegmatis using heterologous Hfq. **RNA**, v. 19, n. 1, p. 74-84, 2013.

LI, W.; YING, X.; LU, Q.; CHEN, L. Predicting sRNAs and their targets in bacteria. **Genomics Proteomics Bioinformatics**. v.10, n. 5, p. 276-284, 2012.

LINK, T. M.; VALENTIN-HANSEN, P.; BRENNAN, R. G. Structure of Escherichia coli Hfq bound to polyriboadenylate RNA. **PNAS**, v. 106, n. 46, p. 19292-19297, 2009.

LIU, J. M.; CAMILLI, A. A broadening world of bacterial small RNAs. **Curr. Opin. Microbiol**, v. 13, p. 18-23, 2010.

LIU, M. Y.; YANG, H.; ROMEO, T. The product of the pleiotropic Escherichia coli gene csrA modulates glycogen biosynthesis via effects on mRNA stability. **J. Bacteriol**, v. 177, n. 10, p. 2663-2672, 1995.

LIU, W. B.; SHI, Y.; YAO, L. L.; ZHOU, Y.; YE, BC. Prediction and characterization of small non-coding RNAs related to secondary metabolites in *saccharopolyspora erythraea*. **Plos one**. v. 8, n. 11, p. e80676, 2013.

LIV, S.Y., YOUNG, C.C., HUPFER, H., SIERING, C., ARUN, A.B., CHEN, W.M., LAI, W.A., SHEN, F.T., REKHA, P.D., YASSIN, A.F. *Azospirillum picis* sp. nov., isolated from discarded tar. **International Journal of Systematic and Evolutionary Microbiology** v. 59, p. 761-765, 2009.

LIVNY, J., BRENCIC, A., LORY, S., WALDOR, M. K. Identification of *Pseudomonas aeruginosa* sRNAs and prediction of sRNA-encoding genes in 10 diverse pathogens using the bioinformatic tool sRNAPredict2. **Nucleic Acids Res.** v. 34, p. 3484–3493, 2006.

LIVNY, J., FOGEL, M. A., DAVIS, B. M., WALDOR, M. K. sRNAPredict: an integrative computational approach to identify sRNAs in bacterial genomes. **Nucleic Acids Res.** v. 33, p. 4096–4105, 2005.

LIVNY, J., TEONADI, H., LIVNY, M., WALDOR, M. K. High-Throughput, Kingdom wide prediction and annotation of bacterial non-coding RNAs. **PLoS One** v. 3, n. 9, p. e3197, 2008.

LIVNY, J., WALDOR, M. K. Identification of small RNAs in diverse bacterial species. **Curr. Opin. Microbiol.** v. 10, p. 1096-1101, 2007.

LOH, E.; DUSSURGET, O.; GRIPENLAND, J.; VAITKEVICIUS, K.; TIENSUU, T.; MANDIN, P.; REPOILA, F.; BUCHRIESER, C.; COSSART, P.; JOHANSSON, J. A trans-acting riboswitch controls expression of the virulence regulator PrfA in *Listeria monocytogenes*. **Cell**, v. 139, n. 4, p. 770-9, 2009.

LORENZ, R.; BERNHART, S. H.; HÖNER, ZU.; SIEDERDISSEN, C.; TAFER, H.; FLAMM, C.; STADLER, P. F.; HOFACKER, I. L. ViennaRNA Package 2.0. **Algorithms. Mol. Biol.**, v. 6, n. 26, 2011.

MACKE, T. J.; ECKER, D. J.; GUTELL, RR, GAUTHERET, D.; CASE, D. A.; SAMPATH, R. RNAMotif, an RNA secondary structure definition and search algorithm. **Nucleic. Acids. Res**, v. 29, n. 22, p. 4724-4735, 2001.

MANDAL, M.; BREAKER, R. R. Adenine riboswitches and gene activation by disruption of a transcription terminator. **Nat. Struct. Mol. Biol**, v. 11, n. 1, p. 29-35, 2003.

MANDAL, M.; BOESE, B.; BARRICK, J. E.; WINKLER, W. C.; BREAKER, R. R. Riboswitches control fundamental biochemical pathways in *Bacillus subtilis* and other bacteria. **Cell**, v. 113, n. 5, p. 577-586, 2003.

MANDIN, P.; GUILLIER, M. Expanding control in bacteria: interplay between small RNAs and transcriptional regulators to control gene expression. **Curr. Opin. Microbiol**, v. 16, n. 2, p. 125-132, 2013.

MARRAFFINI, A. L. The CRISPR-Cas system of *Streptococcus pyogenes*. Function and applications. In: FERRETTI, J. J.; STEVENS, D. L.; FISCHETTI, V. A. **Streptococcus pyogenes: Basic Biology to Clinical Manifestations**. Oklahoma City (OK): University of Oklahoma Health Sciences Center, p. 1-17, 2016.

MARRAFFINI, L. A; SONTHEIMER, E. J. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. **Science**, v. 322, n. 5909, p. 1843-1845, 2008.

MARQUES, A. C.; PALUDO, K. S.; DALLAGASSA, C. B.; SUREK, M.; PEDROSA, F. O.; SOUZA, E. M.; CRUZ, L. M.; LIPUMA, J. J.; ZANATA, S. M.; REGO, F. G.; FADEL-PICHETH, C. M. Biochemical characteristics, adhesion, and cytotoxicity of environmental and clinical isolates of *Herbaspirillum* **J. Clin. Microbiol**, v. 53, n. 1, p. 302-308, 2015.

MASSÉ, E.; GOTTESMAN, S. A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*. **Proc. Natl. Acad. Sci**, v. 7, p. 4620-4625, 2002.

MASSÉ, E., MAJDALANI, N., GOTTESMAN, S. Regulatory roles for small RNAs in bacteria. **Curr. Opin. Microbiol.** v. 6, p. 120-124, 2003.

MCDANIEL, B. A.; GRUNDY, F.J.; ARTSIMOVITCH, I.; HENKIN, T. M. Transcription termination control of the S box system: direct measurement of S-adenosylmethionine by the leader RNA. **Proc. Natl. Acad. Sci.**, v. 100, n. 6, p. 3083-3038, 2003.

MELLIN, J. R.; GOSWAMI, S.; GROGAN, S.; TJADEN, B.; GENCO, C. A. 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.**, v. 189, n. 10, p. 3686-94, 2007.

MELLIN, J. R.; MCCLURE, R.; LOPEZ, D.; GREEN, O.; REINHARD, B.; GENCO, C. Role of Hfq in iron-dependent and -independent gene regulation in *Neisseria meningitidis*. **Microbiology**, v. 156, n. Pt 8, p. 2316-2326, 2010.

MEYER, M. M., HAMMOND, M. C., SALINAS, Y., ROTH, A., SUDARSAN, N., BREAKER, R. R. Challenges of ligand identification for riboswitch candidates. **RNA Biology**, v. 8, n. 1, p. 5-10, 2011.

MIKULÍK, K. Structure and functional properties of prokaryotic small noncoding RNAs. **Folia. Microbiol. (Praha)**, v. 48, n. 4, p. 443-468, 2003.

MIRONOV, A. S.; GUSAROV, I.; RAFIKOV, R.; LOPEZ, L. E.; SHATALIN, K.; KRENEVA, R. A.; PERUMOV, D. A.; NUDLER, E. Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria. **Cell**, v. 11, n. 5, p. 747-756, 2002.

MODI, S. R.; CAMACHO, D. M. KOHANSKI, M. A.; WALKER, G. C.; COLLINS, J. J. Functional characterization of bacterial sRNAs using a network biology approach. **PNAS**, v. 108, p. 15522-15527, 2011.

MOJICA, F. J.; DÍEZ-VILLASEÑOR, C.; SORIA, E.; JUEZ, G. Biological significance of a family of regularly spaced repeats in the genomes of archaea, bacteria and mitochondria. **Mol. Microbiol**, v. 36, n. 1, p. 244-246, 2000.

MORENO, L. F. **Predição e análise de expressão de RNAs não codificadores com função regulatória presentes na bactéria *Herbaspirillum seropedicae* Smr1.** Dissertação (Mestrado em Bioquímica) – Setor de Ciências Biológicas, Universidade Federal do Paraná, Curitiba, 2013.

MORTAZAVI, A.; WILLIAMS, B. A.; MCCUE, K.; SCHAEFFER, L.; WOLD, B. Mapping and Quantifying Mammalian Transcriptomes by RNA-Seq. **Nat. Methods**, v. 5, n. 7, p. 621-628, 2008.

MURINA, V. N., NIKULIN, A. D. Bacterial Small Regulatory RNAs and Hfq Protein. **BIOCHEMISTRY**, v. 80, n.13, p. 1647-1654, 2015

NAHVI, A.; SUDARSAN, N.; EBERT, M. S.; ZOU, X.; BROWN, K. L.; BREAKER, R. R. Genetic control by a metabolite binding mRNA. **Chem. Biol**, v. 9, n. 9, p. 1043.

NAKAMURA, K.; MIYAMOTO, H.; SUZUMA, S.; SAKAMOTO, T.; KAWAI, G.; YAMANE, K. Minimal functional structure of *Escherichia coli* 4.5 S RNA required for binding to elongation factor G. **J. Biol. Chem**, v. 276, n. 25, p. 22844-22849, 2001.

NAWROCKI, E. P.; KOLBE, D. L.; EDDY, S. R. Infernal 1.0: inference of RNA alignments. **Bioinformatics**, v. 25, n. 10, p. 1335-1337, 2009.

NAWROCKI, E. P.; EDDY, S. R. Infernal 1.1: 100-fold faster RNA homology searches. **Bioinformatics**, v. 29, n. 22, p. 2933-2935, 2013.

NOINDORF, L., BONATTO, A.C., MONTEIRO, R.A., SOUZA, E.M., RIGO, L.U., PEDROSA, F.O., STEFFENS, M.B.R., CHUBATSU, L.S. Role of PII proteins in nitrogen fixation control of *Herbaspirillum seropedicae* strain SmR1. **BMC Microbiology**, v. 11, n. 8, 2011.

OLIVA, G.; SAHR, T.; BUCHRIESER, C. Small RNAs, 5' UTR elements and RNA-binding proteins in intracellular bacteria: impact on metabolism and virulence. **FEMS. Microbiol. Rev**, v. 39, n. 3, p. 331-49, 2015.

ORTEGA, A. D.; GONZALO-ASENSIO, J.; GARCÍA-DEL PORTILLO, F. Dynamics of Salmonella small RNA expression in non-growing bacteria located inside eukaryotic cells. **RNA. Biol**, v. 4, p. 469-488, 2012.

PADALON-BRAUCH, G.; HERSHBERG, R.; ELGRABLY-WEISS, M.; BARUCH, K.; ROSENSHINE, I.; MARGALIT, H.; ALTUVIA, S. Small RNAs encoded within genetic islands of *Salmonella typhimurium* show host-induced expression and role in virulence. **Nucleic. Acids. Res**, v. 36, n. 6, p. 1913-27, 2008.

PANKIEVICZ, V. C. S. **Investigação de mecanismos envolvidos na interação entre *Herbaspirillum seropedicae* e gramíneas.** Projeto de Tese de doutorado, Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Paraná. Curitiba, 2013.

PÁNEK, J.; BOBEK, J.; MIKULÍK, K.; BASLER, M.; VOHRADSKÝ, J. Biocomputational prediction of small non-coding RNAs in streptomyces. **BMC. Genomics**, v. 9, p. 217, 2008.

PAPENFORT, K.; VANDERPOOL, C. K. Target activation by regulatory RNAs in bacteria. **FEMS. Microbiol. Rev**, v. 39, n. 3, p. 362-378, 2015.

PEDROSA, F. O.; MONTEIRO, R. A.; WASSEM, R.; CRUZ, L. M.; AYUB, R. A.; COLAUTO, N. B.; FERNANDEZ, M. A.; FUNGARO, M. H.; GRISAR, E. C.; HUNGRIA, M.; MADEIRA, H. M. F.; NODARI, R. O.; OSAKU, C. A.; PETZL-ERLER, M. L.; TEREZI, H.; VIERIA, L. G. E.; STEFFENS, M. B. R.; WEISS, V. A.; PEREIRA, L. F. P., ALMEIDA, I. M., ALVES, L. R., MARIN, A., ARAUJO, L. M., BALSANELLI, E., BAURA, V. A., CHUBATSU, L. S., FAORO, H., FAVETTI, A., FRIEDERMANN, G., GLIENKE, C., KARP, S., KAVA-CORDEIRO, V., RAITTZ R., RAMOS, H. J. O., RIBEIRO, S. E. M. F., RIGO, L.U., ROCHA, S. N., SCHWAB, S., SILVA, A. G., SOUZA, E.M., TADRA-SFEIR, M. Z., TORRES, R. A., DABUL, A. N. G., SOARES, M. A., GASQUES, L. S., GIMENES, C. C. T., VALLE, J. S., CIFERRI, R. R., CORREA, L.

C., MURACE, N. K., PAMPHILE, J. A., PATUSSI, E. V., PRIOLI, A. J., PRIOLI, S. M. A., ROCHA, C. L. M. S. C., ARANTES, O. M. N.; FURNALETO, M. C.; GODOY, L.P.; OLIVEIRA, C. E. C.; SATORI, D.; VILAS-BOAS, L. A.; WATANABE, M. A. E.; DAMBROS, B. P.; GUERRA, M. P.; MATHIONI, S. M.; SANTOS, K. L.; STEINDEL, M.; VERNAL, J.; BARCELLOS, F. G.; CAMPO, R. J.; CHUEIRE, L. M. O.; NICOLÁS, M. F.; PEREIRA-FERRARI, L.; SILVA, J. L.C.; GIOPPO, N. M. R.; MARGARIDO, V. P.; MENCK-SOARES, M. A.; PINTO, F. G. S.; SIMÃO, R. C. G.; TAKAHASHI, E. K.; YATES, M. G.; SOUZA, E. M. Genome of *Herbaspirillum seropedicae* strain SmR1, a specialized diazotrophic endophyte of tropical grass. **PLoS Geneti**, cs. In press, 2011.

PENG, X.; DONG, H.; WU, Q. A new *cis-encoded* sRNA, BsrH, regulating the expression of hemH gene in *Brucella abortus* 2308. *FEMS. Microbiol, Lett*, v. 362, n. 2, p. 1-7, 2015.

PEDERSEN, J. S.; BEJERANO, G.; SIEPEL, A.; ROSENBLOOM, K.; LINDBLADTOH, K.; LANDER, E. S.; KENT, J.; MILLER, W.; HAUSSLER, D. Identification and classification of conserved RNA secondary structures in the human genome. **PLoS. Comput. Biol**, v. 4, p. e33, 2006.

PÉREZ, A. A.; RODIONOV, D. A.; BRYANT, D. A. Identification and Regulation of Genes for Cobalamin Transport in the Cyanobacterium *Synechococcus* sp. Strain PCC 7002. **J. Bacteriol**, v. 198, n. 19, p. 2753-2761, 2016.

PICHON, C.; FELDEN, B. Intergenic sequence inspector: searching and identifying bacterial RNAs. **Bioinformatics**, v. 19, p. 1707–1709, 2003.

PICHON, C.; FELDEN, B. Proteins that interact with bacterial small RNA regulators. **Fems. Microbiol. Rev**, v. 31, p. 614-625, 2007.

PICHON, C.; FELDEN, B. Small RNA gene identification and mRNA target predictions in bacteria. **Bioinformatics**, v. 24, n. 24, p. 2807-2813, 2008.

PICHON, C.; MERLE, L. D.; CALIOT, M. E.; TRIEU-CUOT, P.; BOUGUÉNEC, C. L. 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. Research**, v. 40, n. 7, 2012.

PFAFFL, M. W. A new mathematical model for relative quantification in real-time RT-PCR. **Nucleic. Acids. Res**, v. 29, n. 9, p. e45, 2001.

PFEIFFER, V.; PAPENFORT, K.; LUCCHINI, S.; HINTON, J. C.; VOGEL, J. Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. **Nat. Struct. Mol. Biol**, v. 16, n. 8, p. 840-846, 2009.

RAY, B. K.; APIRION, D. Characterization of 10S RNA: a new stable rna molecule from *Escherichia coli*. **Mol. Gen. Genet**, v. 174, n. 1, p. 25-32, 1979.

REPOILA, F.; DARFEUILLE, F. Small regulatory non-coding RNAs in bacteria: **physiology** and mechanistic aspects, **Biol Cell**, v. 101, n. 2, p. 117-131, 2009.

RICHARDS, J.; BELASCO, J. G. A new window onto translational repression by bacterial sRNAs. **Mol. Cell**, v. 32, n. 6, p. 751-753, 2008.

RICHARDS, G. R.; VANDERPOOL, C. K. Molecular call and response: the physiology of bacterial small RNAs. **Biochim. Biophys. Acta**, v. 1809, n. 10, p. 525-31, 2011.

RICHTER, C.; CHANG, J. T.; FINERAN, P. C. Function and regulation of clustered regularly interspaced short palindromic repeats (CRISPR) / CRISPR associated (Cas) systems. **Viruses**, v. 4, n. 10, p. 2291-311, 2012.

RIVAS E., KLEIN R. J., JONES T. A., EDDY S. R. Computational identification of noncoding RNAs in *E. coli* by comparative genomics. **Curr. Biol**, v. 11 n. 17 p. 1369, 2001.

ROTH, A.; WINKLER, W. C.; REGULSKI, E. E.; LEE, B. W.; LIM, J.; JONA, I.; BARRICK, J. E.; RITWIK, A.; KIM, J. N.; WELZ, R.; IWATA-REUYL, D.; BREAKER, R. R. A riboswitch selective for the queuosine precursor preQ1 contains an unusually small aptamer domain. **Nat. Struct. Mol. Biol**, v. 14, n. 4, p. 308-317, 2007.

ROTHBALLER, M.; SCHMID, M.; KLEIN, I.; GATTINGER, A.; GRUNDMANN, S.; HARTMANN, A. 2006. *Herbaspirillum hiltneri* sp. nov., isolated from surface-sterilized wheat roots. **International Journal of Systematic and Evolutionary Microbiology**, v. 56, p. 1341–1348.

ROSSI, C. C.; BOSSÉ, J. T.; LI, Y.; WITNEY, A. A.; GOULD, K A.; LANGFORD, P. R.; BAZZOLLI, D. M. A computational strategy for the search of regulatory small RNAs in *Actinobacillus pleuropneumoniae*. **RNA**, v. 22, n. 9, p. 1373-1385, 2016.

RUTHERFORD, K., PARKHILL, J., CROOK, J., HORSNELL, T., RICE, P., RAJANDREAM, M. A., BARRELL, B. Artemis: sequence visualization and annotation **Bioinformatics**, v.16, n.10, p. 944-945, 2000.

SAMPSON, T. R.; SAROJ, S. D.; LLEWELLYN, A. C.; TZENG, Y. L.; WEISS, D. S. A CRISPR/Cas system mediates bacterial innate immune evasion and virulence. **Nature**, v. 497, n. 7448, p. 254-257, 2013.

SAHR, T.; BRÜGGEMANN, H.; JULES, M.; LOMMA, M.; ALBERT-WEISSENBERGER, C.; CAZALET, C.; BUCHRIESER, C. Two small ncRNAs jointly govern virulence and transmission in *Legionella pneumophila*. **Mol. Microbiol**, v. 72, n. 3, p. 741-762, 2009.

SASS, AM.; VAN ACKER, H.; FÖRSTNER, K. U.; VAN NIEUWERBURGH, F.; DEFORCE, D.; VOGEL, J.; COENYE, T. Genome-wide transcription start site profiling in biofilm-grown *Burkholderia cenocepacia* J2315. **BMC. Genomics**, v. 16, p. 775, 2015.

SESTO, N.; TOUCHON, M.; ANDRADE, J. M.; KONDO, J.; ROCHA, E. P.; ARRAIANO, C. M.; ARCHAMBAUD, C.; WESTHOF, É.; ROMBY, P.; COSSART, P. A PNPase dependent CRISPR System in *Listeria*. **PLoS. Genet**, v. 10, n. 1, p. e1004065, 2014.

SRIDHAR, J.; SAMBATURU, N.; SABARINATHAN, R.; OU, H. Y.; DENG, Z.; SEKAR, K.; RAFI, Z. A.; RAJAKUMAR, K. sRNAscanner: a computational tool for intergenic small RNA detection in bacterial genomes. **PLoS. One**, v.5, n. 8, p. e11970, 2010.

SCHMID, M.; BALDANI, J.I., HARTMANN, A. The Genus *Herbaspirillum*. **The Prokaryotes**. New York: Springer: v. 141-150, 2006.

SCHMIDTKE, C.; FINDEISS, S.; SHARMA, C. M.; KUHFUSS, J.; HOFFMANN, S.; VOGEL, J.; STADLER, P. F.; BONAS, U. Genome-wide transcriptome analysis of the plant pathogen *Xanthomonas* identifies sRNAs with putative virulence functions. *Nucleic. Acids. Res.*, v. 40, n. 5, p. 2020-2031, 2012.

SCHROEDER, C. L. C., NARRA, H. P., ROJAS, M., SAHNI, A., PATEL, J., KHANIPOV, K., WOOD, T. G., FOFANOV, Y., SAHNI, S. K. Bacterial small RNAs in the Genus *Rickettsia*. **BMC Genomics**, v. 16, p. 1075, 2015.

SCHROEDER, C. L.; NARRA, H. P.; SAHNI, A.; ROJAS, M.; KHANIPOV, K.; PATEL, J.; SHAH, R.; FOFANOV, Y.; SAHNI, S. K. Identification and Characterization of Novel Small RNAs in *Rickettsia prowazekii*. **Front. Microbiol**, v. 8, n. 7, p. 859, 2016.

SHARMA, C. M., VOGEL, J. Experimental approaches for the discovery and characterization of regulatory small RNA. **Curr. Opin. Microbiol**, v. 12, n. 5, p. 536-546, 2009.

SHINHARA, A., MATSUI, M., HIRAOKA, K., NOMURA, W., HIRANO, R., NAKAHIGASHI, K., TOMITA, M., MORI, H., KANAI, A. Deep sequencing reveals as-yet-undiscovered small RNAs in *Escherichia coli*. **BMC. Genomics**, v. 12, n. 428, p. 1471-2164, 2011.

SMITH, N. A; EAMENS, A. L; WANG, M.B. The presence of high-molecular-weight viral RNAs interferes with the detection of viral small RNAs. **RNA**, v.16, n. 5, p. 1062-1067, 2010.

SOLANS, L.; GONZALO-ASENSIO, J.; SALA, C.; BENJAK, A.; UPLEKAR, S.; ROUGEMONT, J.; GUILHOT, C.; MALAGA, W.; MARTÍN, C.; COLE, S. T. The PhoP-dependent ncRNA Mcr7 modulates the TAT secretion system in *Mycobacterium tuberculosis*. **PLoS. Pathog**, v. 10, n. 5, p. e1004183, 2014.

SOPER, T.; MANDIN, P.; MAJDALANI, N.; GOTTESMAN, S.; WOODSON, S. A. Positive regulation by small RNAs and the role of Hfq. **Proc. Natl. Acad. Sci**, v. 107, n. 21, p. 9602-9607, 2010.

SRIDHAR, J.; GUNASEKARAN, P. Computational small RNA prediction in bacteria. **Bioinform. Biol. Insights**, v. 7, p. 83-95, 2013.

SRIDHAR, J.; SAMBATURU, N.; SABARINATHAN, R.; OU, H. Y.; DENG, Z.; SEKAR, K.; RAFI, Z.A.; RAJAKUMAR, K. sRNAscanner: a computational tool for intergenic small RNA detection in bacterial genomes. **Plos. One**, v. 5, n. 8, p. e11970, 2010.

SUDARSAN, N.; LEE, E. R.; WEINBERG, Z.; MOY, R. H.; KIM, J. N.; LINK, K. H.; BREAKER, R. R. Riboswitches in eubacteria sense the second messenger cyclic di-GMP. **Science**, v. 321, n. 5887, p. 411-413, 2008.

SURESH, G.; SRINIVASAN, H.; NANDA, S.; PRIYAKUMAR, U. D. Ligand-Induced Stabilization of a Duplex-like Architecture Is Crucial for the Switching Mechanism of the SAM-III Riboswitch. **Biochemistry**, v. 55, n. 24, p. 3349-3360, 2016.

STEUTEN, B.; HOCH, P. G.; DAMM, K.; SCHNEIDER, S.; KÖHLER, K.; WAGNER, R.; HARTMANN, R. K. Regulation of transcription by 6S RNAs: insights from the *Escherichia coli* and *Bacillus subtilis* model systems. **RNA. Biol**, v. 11, n. 5, p. 508-521, 2014.

STORZ G.; ALTUVIA, S.; WASSARMAN, K. M. An abundance of RNA regulators. **Annu. Rev. Biochem**, v. 74, p. 199-217, 2005.

STORZ, G.; VOGEL, J.; WASSARMAN, K. M. Regulation by small RNAs in bacteria: expanding frontiers. **Mol. Cell**, v. 43, n. 6, p. 880-891, 2011.

TADRA-SFEIR, M. Z. **Caracterização funcional de genes de *Herbaspirillum seropedicae* regulados pelo flavonóide naringenina.** Tese de doutorado, Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Paraná. Curitiba, 2011.

TATTERSALL, J.; RAO, G. V.; RUNAC, J.; HACKSTADT, T.; GRIESHABER, S. S.; GRIESHABER, N. A. Translation inhibition of the developmental cycle protein HctA by the small RNA lhtA is conserved across Chlamydia. **PLoS. One**, v. 7, n. 10, p. e47439. 2012.

TALES, A. A. F. **Biologia Celular e Molecular**, 2^a ed. Átomo, 2013.

TOLEDO-ARANA, A.; DUSSURGET, O.; NIKITAS, G.; SESTO, N.; GUET-REVILLET, H.; BALESTRINO, D.; LOH, E.; GRIPENLAND, J.; TIENSUU, T.; VAITKEVICIUS, K.; BARTHELEMY, M.; VERGASSOLA, M.; NAHORI, M. A.; SOUBIGOU, G.; RÉGNAULT, B.; COPPÉE, J. Y.; LECUIT, M.; JOHANSSON, J.; COSSART, P. The *Listeria* transcriptional landscape from saprophytism to virulence. **Nature**, v. 459, n. 7249, p. 950-956, 2009.

TRAPNELL, C.; WILLIAMS, B. A.; PERTEA, G.; MORTAZAVI, A.; KWAN, G.; VAN BAREN, M. J.; SALZBERG, S. L.; WOLD, B. J.; PACHTER, L. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. **Nat. Biotechnol**, v. 28, n. 5, p. 511-5, 2010.

TJADEN, B. TargetRNA: a tool for predicting targets of small RNA action in bacteria. **Nucleic. Acids. Res**, v. 36, p. 109-113, 2008.

TJADEN, B.; GOODWIN, S. S.; OPDYKE, J. A.; GUILLIER, M.; FU, D. X.; GOTTESMAN, S.; STORZ, G. Target prediction for small, noncoding RNAs in bacteria. **Nucleic. Acids. Res**, v. 34, n. 9, p. 2791-2802, 2006.

TRAN, T. T.; ZHOU, F.; MARSHBURN, S.; STEAD, M.; KUSHNER, S. R.; XU, Y. De novo computational prediction of non-coding RNA genes in prokaryotic genomes. **Bioinformatics**, v. 25, n. 22, p. 2897-2905, 2009.

TROTOCHAUD, A.E.; WASSARMAN, K. M. 6S RNA function enhances long-term cell survival. **J. Bacteriol**, v. 186, n. 15, p. 4978-4985, 2004.

TYAGI, J. S.; KINGER, A. K. Identification of the 10Sa RNA structural gene of *Mycobacterium tuberculosis*. **Nucleic. Acids. Res.** v. 20, n. 1, p. 138, 1992.

ULVÉ, V. M., SEVIN, E. W., CHÉRON, A., BARLOY-HUBLER, F. Identification of chromosomal alpha-proteobacterial small RNAs by comparative genome analysis and detection in *Sinorhizobium meliloti* strain 1021. **BMC Genomics**. v. 8, n. 467, 2007.

VAKULSKAS, C. A.; LENG, Y.; ABE, H.; AMAKI, T.; OKAYAMA, A.; BABITZKE, P.; SUZUKI, K.; ROMEO, T. Antagonistic control of the turnover pathway for the global regulatory sRNA CsrB by the CsrA and CsrD proteins. **Nucleic. Acids. Res.** v. 44, n. 16, p. 7896-7910, 2016.

VALVERDE, V. B.; A.; VELÁZQUEZ, E.; GUTIÉRREZ, C.; CERVANTES, E.; VENTOSA, A.; IGUAL, J. M. *Herbaspirillum lusitanum* sp. nov., a novel nitrogen-fixing bacterium associated with root nodules of *Phaseolus vulgaris*. **Int. J. Syst. Evol. Microbiol**, v. 53, (Pt 6), p. 1979-1983, 2003.

VOGEL, J.; WAGNER, E. G. Target identification of small noncoding RNAs in bacteria. **Curr. Opin. Microbiol**, v. 10, n. 3, p. 262-270, 2007.

VOSS, B.; GEORG, J.; SCHON, V.; UDE, S.; HESS, W. R. Biocomputational prediction of non-coding RNAs in model *cyanobacteria*. **BMC. Genomics**, v. 10, p. 123, 2009.
ZAHA, A., FERREIRA, H. B., PASSAGLIA, L. M. P. (Org.). **Biologia molecular básica**, 5. ed. Porto Alegre: Artmed, 2014.

ZHANG, A. WASSARMAN, K. M.; ROSENOW, C.; TJADEN, B. C.; STORZ, G.; GOTTESMAN, S. Global analysis of small RNA and mRNA targets of Hfq. **Mol. Microbiol**, v. 50, n. 4, p. 1111-1124, 2003.

ZAHA, A; BUNSELMEYER, H. F; PASSAGLIA, L. M. P. *Biologia Molecular Básica*, 5ª ed. Artemis, 2014.

ZWIEB, C.; WOWER, I.; WOWER, J. Comparative sequence analysis of tmRNA. **Nucleic. Acids. Res**, v. 27, n. 10, p. 2063-2071, 1999.

WAGNER, E. G.; SIMONS, R. W. Antisense RNA control in bacteria, phages, and plasmids. **Annu. Rev. Microbiol**, v. 48, p. 713-42, 1994.

WARRIER, I.; HICKS, L. D.; BATTISTI, J. M.; RAGHAVAN, R.; MINNICK, M. F. Identification of novel small RNAs and characterization of the 6S RNA of *Coxiella burnetii*. **PLoS One**, v. 9, n. 6, p. e100147, 2014.

WASSARMAN, K.M; KILEY, P. J. Global approaches for finding small RNA and small open reading frame functions. **J Bacteriol**, v. 192, n. 1, p. 26-8, 2010.

WASSARMAN, K. M., REPOILA, F., ROSENOW, C., STORZ, G., GOTTESMAN, S. Identification of novel small RNAs using comparative genomics and microarrays. **Genes Dev**, v. 15, p. 1637–1651, 2001.

WATERS, L. S., STORZ, G. Regulatory RNAs in bacteria. **Cell**, v. 136, n. 4, p. 615-628, 2009.

WEHNER, S.; DAMM, K.; HARTMANN, R. K.; MARZ, M. Dissemination of 6S RNA among bacteria. **RNA. Biol**, v. 11, n. 11, p. 1467-1478, 2014.

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**, v. 35, n. 14, p. 4809-4819, 2007.

WEINBERG, Z.; WANG, J. X.; BOGUE, J.; YANG, J.; CORBINO, K.; MOY, R. H.; BREAKER, R. R. Comparative genomics reveals 104 candidate structured RNAs from bacteria, archaea, and their metagenomes. **Genome. Biol**, v. 11, n. 3, p. R31, 2010.

WEN, Y.; FENG, J.; SACHS, G. Helicobacter pylori 5'ureB-sRNA, a *cis-encoded* antisense small RNA, negatively regulates ureAB. **J Bacteriol**, v. 195, n. 3, p. 444-452, 2013.

WILLEMANN, G. M. **Predição de RNAs curtos em *Herbaspirillum seropedicae* SmR1**. Dissertação (Mestrado em Bioinformática), Programa de Pós-Graduação em Bioinformática, Setor de Educação Profissional e Tecnológica, Universidade Federal do Paraná, Curitiba, 2010.

WILLIAMS, K. P.; BARTEL, D. P. Phylogenetic analysis of tmRNA secondary structure. **RNA**, v. 2, n. 12, p. 1306-1310, 1996.

WINGENDER, E.; CHEN, X.; HEHL, R.; KARAS, H.; LIEBICH, I.; MATYS, V.; MEINHARDT, T.; PRUSS, M.; REUTER, I.; SCHACHERER, F. TRANSFAC: an integrated system for gene expression regulation. **Nucleic. Acids. Res**, v. 28, n. 1, p. 316–319, 2000.

WINKLER, W. C. *Riboswitches* and the role of noncoding RNAs in bacterial metabolic control. **Current Opinion in Chemical Biology**, v. 9, p. 594–602, 2005.

WINKLER, W. C.; COHEN-CHALAMISH, S.; BREAKER, R. R. An mRNA structure that controls gene expression by binding FMN. **Proc Natl. Acad**, v. 99, n. 25, p. 15908-15913, 2002.

WINKLER, W. C.; NAHVI, A.; ROTH, A.; COLLINS, J. A.; BREAKER, R. R. Control of gene expression by a natural metabolite-responsive ribozyme. **Nature**, v. 428, n. 6980, p. 281-286, 2004.

WINKLER, W. C.; NAHVI, A.; SUDARSAN, N.; BARRICK, J. E.; BREAKER, R. R. An mRNA structure that controls gene expression by binding S-adenosylmethionine. **Nat. Struct. Biol.**, v. 10, n. 9, p. 701-707, 2003.

WOWER, I. K.; ZWIEB, C.; WOWER, J. Transfer-messenger RNA unfolds as it transits the ribosome. **RNA**, v. 11, n. 5, p. 668-673, 2005.

WRIGHT, P. R.; GEORG, J.; MANN, M.; SORESCU, D. A.; RICHTER, A. S.; LOTT, S.; KLEINKAUF, R.; HESS, W. R.; BACKOFEN, R. CopraRNA and IntaRNA: predicting small RNA targets, networks and interaction domains. **Nucleic. Acids. Res.**, v. 42, p. 119-123, 2014.

WROBLEWSKA, Z.; OLEJNICZAK, M. Hfq assists small RNAs in binding to the coding sequence of ompD mRNA and in rearranging its structure. **RNA**, v. 22, n. 7, p. 979-994, 2016.

ANEXO I – CONDIÇÕES DE CULTIVO EMPREGADAS NOS EXPERIMENTOS DE RNaseq COM *H. seropedicae* SmR1.

CONDIÇÕES DE CULTIVO	DESCRIÇÃO
1. NFbHPN-Malato (Controle1, Controle2)	A condição controle consiste nas bactérias cultivadas a 30°C em meio NFbHPN, utilizando malato como fonte de carbono até D.O 0,8 (TADRA-SFEIR, 2011).
2. NFbHPN-Malato-Naringenina (Nar_1, Nar_2)	Bactérias cultivadas a 30°C em meio NFbHPN, utilizando malato como fonte de carbono na presença de 100µM do flavonóide naringenina, até D.O. 0,8. (TADRA-SFEIR, 2011).
3. NFbHPN-Malato-Nitrato (Nitrato_1, Nitrato_2)	Bactérias cultivadas a 30°C em meio NFbHPN, utilizando malato como fonte de Carbono e nitrato 10mM como fonte de nitrogênio. (BONATO, 2012).
4. Planctônico–milho-1dia (PS1_1, PS1_2)	Bactérias recuperadas do meio líquido do sistema hidropônico (denominadas plantônicas) após um dia de inoculação em raiz de milho (BALSANELLI, 2013).
4. Planctônico-milho-3dias (PS3_1, PS3_2)	Bactérias recuperadas do meio líquido do sistema hidropônico (denominadas plantônicas) após um dia de inoculação em raiz de milho (BALSANELLI, 2013).
5. Aderido–milho- 1dia (AD1_1, AD1_2)	Bactérias recuperadas da superfície radicular de milho (denominadas aderidas) após um dia de inoculação (BALSANELLI, 2013).
6. Aderido-milho3dias (AD3_1, AD3_2)	Bactérias recuperadas da superfície radicular de milho (denominadas aderidas) após três dias de inoculação (BALSANELLI, 2013).
7. Planctônico- trigo (PT1, PT2)	Bactérias recuperadas do meio líquido do sistema hidropônico (denominadas plantônicas) após três dias de inoculação em raiz de trigo (PANKIEVICZ, 2013).
8. Aderido-trigo (WR1, WR2)	Bactérias recuperadas da superfície radicular de trigo (denominadas aderidas) após três dias de inoculação (PANKIEVICZ, 2013).
9. NFbHPN-Malato-Alto Oxigênio (AltoO2_1, AltoO2_2)	Bactérias cultivadas a 30°C em meio NFbHPN, utilizando malato como fonte de carbono até D.O. de 0,4. Depois de atingida a D.O. 0,4 as bactérias continuaram sendo cultivadas nesta condição por mais 1,5 horas. (BATISTA, 2013).
11. NFbHPN-Malato-Baixo Oxigênio (BaixoO2_1, BaixoO2_2)	Bactérias cultivadas a 30°C em meio NFbHPN, utilizando malato como fonte de carbono até D.O. de 0,4. Nesse momento as bactérias passaram a ser cultivadas em condiçõesmicroanaerobicas (concentração inicial de oxigênio 2%) por 1,5 horas. (BATISTA, 2013).

ANEXO II – ESPÉCIES DE *Herbaspirillum* COM CÓDIGO E DATA DE ACESSO

Genêro <i>Herbaspirillum</i>	Código de Acesso ao NCBI	Data de Acesso
<i>Herbaspirillum seropedicae</i> Z67	NZ_CP011930.1	17/02/2016
<i>Herbaspirillum lusitanum</i> P6-12	NZ_AJHH01000397.1	19/03/2016
<i>Herbaspirillum hiltneri</i> N3	NZ_CP011409.1	21/03/2016
<i>Herbaspirillum frisingense</i> GSF30	NZ_AEEC02000001.1	22/03/2016
<i>Herbaspirillum</i> spp. estirpe B65	NZ_BADIO1000775.1	22/03/2016
<i>Herbaspirillum rubrisubalbicans</i> M1	NZ_CP013737.1	22/03/2016
<i>Herbaspirillum huttiense</i> subsp. <i>putei</i> estirpe IAM 15032	NZ_ANJR01000025.1	22/03/2016
<i>Herbaspirillum autotrophicum</i> IAM 14942	NZ_LFLS01000001.1	22/03/2016
<i>Herbaspirillum</i> spp. estirpe B501	NZ_BADJ01000320.1	22/03/2016
<i>Herbaspirillum</i> spp. estirpe GW103	NZ_AJVC01000003.1	22/03/2016
<i>Herbaspirillum</i> spp. estirpe RV1423	NZ_CBXX010000004. 1	22/03/2016
<i>Herbaspirillum rhizosphaerae</i> UMS-37	NZ_LFLU01000001.1	22/03/2016
<i>Herbaspirillum rubrisubalbicans</i> spp. estirpe Os34	NZ_AMSB01000001. 1	22/03/2016
<i>Herbaspirillum rubrisubalbicans</i> spp. estirpe Os45	NZ_AMSA01000003. 1	22/03/2016
<i>Herbaspirillum</i> AU14040	NZ_CP011930.1	22/03/2016

ANEXO III – ncRNAs PREDITOS POR MORENO (2013) E CIRINO (2014) em *H. seropedicae* SmR1

ncRNA	Posição Inicial	Posição Final	Sentido	Rfam 12.0	Acesso Rfam 12.0
Hs_noco_471	219058	219229	+		
Hs_noco_494	457262	457311	+		
Hs_noco_525	932018	932172	-		
Hs_noco_530	957469	957626	+		
Hs_noco_534	1067480	1067693	-		
Hs_noco_560	1495976	1496389	-		
Hs_noco_637	2815035	2815266	-		
Hs_noco_651	3173522	3173721	-		
Hs_noco_682	3633077	3633378	-		
Hs_noco_698	3779834	3780049	-		
Hs_noco_700	3791718	3791849	-		
Hs_noco_704	3879771	3880105	+		
Hs_noco_709	3928269	3928556	+		
Hs_noco_719	4074941	4075134	-		
Hs_noco_722	4105880	4106065	-		
Hs_noco_723	4127435	4127634	-		
Hs_noco_726	4148338	4148890	-		
Hs_noco_732	4239065	4239615	+	tmRN A	RF00023
Hs_noco_742	4396086	4396487	+		
Hs_noco_744	4417512	4417740	-		
Hs_noco_746	4428208	4428652	-		
Hs_noco_748	4467236	4467750	+		
Hs_noco_749	4473739	4474140	-		
Hs_noco_750	4474754	4474942	+		
Hs_noco_761	4618606	4618928	+		
Hs_noco_788	5026868	5027073	-		
Hs_noco_802	5239833	5240191	+		
Hs_noco_820	5397562	5398115	+		
Hs_sRNA_4.2	2234033	2234214	+		
Hs_sRNA_19.2	612014	612436	-		
Hs_sRNA_2.1 ou 4.5S	4108897	4109042	-	4.5S	RF00169
TPP	462479	462706	+	TPP	RF00059
Hs_sRNA_Gs6	1915816	1915944	+		
Hs_sRNA_Gs10	808782	808981	-		
Hs_sRNA_Gs12	2233545	2233793	-		
Hs_sRNA_Gs36	1788675	1788860	-		
Hs_noco_811	5335529	5335812	-		
Hs_noco_498	551015	551235	+		
Hs_noco_780	4896320	4896704	+		
Hs_noco_786	5019358	5019702	+		
Hs_noco_757	4564996	4565251	-		
Hs_noco_725	4136562	4136795	-		
Hs_noco_714	3998445	3998747	-		

Hs_noco_640	2939610	2939824	+		
Hs_noco_644	2972948	2973103	+		
Hs_noco_602	2140731	2141104	-		
Hs_noco_604	2177189	2177339	+		
Hs_noco_564	1587873	1588410	+		
Hs_noco_565	1596987	1597381	-		
Hs_noco_568	1636498	1636822	+		
Hs_noco_593	2032308	2032540	+		
Hs_noco_551	1383591	1383913	-	6S	RF00013
Hs_noco_555	1433080	1433239	+		
Hs_noco_526	935305	935540	+		
Hs_noco_490	417253	417332	+		
Hs_noco_514	808928	809191	+		
Hs_noco_724	4127694	4127848	-		
"CUFF.55"	205713	205798	+		
"CUFF.111"	359683	359916	+	crcB	RF01734
"CUFF.124"	412442	412676	+		
"CUFF.223"	664741	664801	+		
"CUFF.251"	741007	741139	+		
"CUFF.273"	792569	792756	+		
"CUFF.395"	1155828	1155922	+		
"CUFF.404"	1189098	1189333	+		
"CUFF.504"	1336439	1336564	+		
"CUFF.522"	1381282	1381477	+		
"CUFF.524"	1382857	1382981	+		
"CUFF.689"	1867577	1867679	+		
"CUFF.690"	1880603	1880826	+		
"CUFF.738"	2011544	2011601	+		
"CUFF.829"	2225907	2226140	+		
"CUFF.852"	2287397	2287464	+		

**APÊNDICE I – SEQUÊNCIA DOS ncRNAs PREDITOS EM *H. seropedicae* SmR1
PELA FERRAMENTA INFERNAL 1.1.1**

>5_ureB_sRNA undefined product 5415884:5416169 reverse

5'CTTCCTCGATCTGCTGCGGACAGATGAAGTGGATGTGGCTGTTCGATGCCGCC
GGCGGTGACGATCATGCCTTCACCCGCGATGATTTTCGGTAGCGCCACCGATGG
CCATGGTCACGCCAGGCTGGATGTCGGGATTGCCGGCCTTGCCGATGCCGGC
GATCTTGCCCCACTTGATGCCGATGTCGGCCTTGACGATGCCCCAGTGATCCA
CGATCAGGGCATTGGTGTGACGGTATCCATCACGTCCGCATGGGCGCGTTGC
GACTGGCCCATGCCGTCGCGGAT3'

>6S undefined product 1383620:1383796 reverse

5'GTACTCCTGCCGTGTTTCGCGATTGCCATATATTCCTTGAACCATTTACGAGCAT
AGGTTGTGGGATCGTTTGATGGGCGTGAGCGTCGCTAGTCCGATGAACCCGAA
ATTGAACTACTGCAACCAACTTGAACCGTCAGGTTTCAGGATGCCGGCACAGCG
GCACAGGCGGGGCTTGA3'

>Afu_254 undefined product 4233427:4233498 forward

5'GCTCTTGCTGGCGCTGCTGGACATGTCCGAAGCCGTGGCGCTGGACGCGAT
GGCGAGATGCTGGGCTTTGC3'

>alpha_tmRNA undefined product 2687581:2687684 forward

5'GCGAACCCCTTTGCGGTGAGCAACACCCGGCAAGGTCGGCCATCACATGGTGCC
CATGCCGCACCTGGGCGTGATCTTGCCGATGCCGCAATGGCAGGCCCTGGCC3
,

>AS1726 undefined product 1536374:1536442 reverse

5'GGAAGGGCAACGACAGCACCCAGCGCGCAGCCCCAGATGATGGCCTTCTGCA
CATCCGGGTCCGGGGTCC3'

>ar35 undefined product 3551853:3552014 forward

5'GTACGGCCCGTTTCGGCCGAGGGTGAGCTGGTGGTGGTGGGCTTGTCC
TCCTGCTCGATGGCTGTTCGCGACATGGAGCGTACGGCAGTGGCGTAAGCTGCG
GGCAATATGGGCAAGAAAGCAAAACGGCCGATCTTCTACGATCCGGCCGTTTT
GTTT3'

>ASdes undefined product 788645:788709 forward

5'GATCGAAGTGGCCGCTGACGCGCAGACCCCGATGATCTCGCTGGCCGCTGC
CGCGCGCATCGTCG3'

>Bacteria_large_SRP undefined product 4108910:4109005 reverse

5'GGCGGGCCCCTGCGCATTGTGGCATGGTCAACCTGGTCAGGTCGGGAACGA
AGCAGCCACAGCCATTTCCCGCAAGTGCCGCAGACAAGGCTCGCC3'

>Bacteria_small_SRP undefined product 4108907:4109005 reverse

5'GGCGGGCCCCTGCGCATTGTGGCATGGTCAACCTGGTCAGGTCGGGAACGA
AGCAGCCACAGCCATTTCCCGCAAGTGCCGCAGACAAGGCTCGCCTCC3'

>Betaproteobacteria_toxic_sRNA undefined product 4408979:4409043 reverse
 5'TGTCTCCTCCACCTCCTCAATGGTGTGGATTAAACCCGGAACCACAGTGTCC
 GGGTTTTTTTTG3'

>beta_tmRNA undefined product 2687527:2687848 forward
 5'CCGATACCTGGGTTCGATATCGACTTCTTCGGTCATCAGCTCTCGCTGCATCTG
 GGCGAACCTTTGCGGTGAGCAACACCGGCAAGGTCGGCCATCACATGGTGCC
 CATGCCGCACCTGGGCGTGATCTTGCCGATGCCGCAATGGCAGGCCCTGGCC
 GACCGCCTCAGCGCCAGCGCACAGGTGCAGTTCGTGCTGGCGCCGCAGATCC
 GCTTCGCCGCGAACC GGCGAGCAGGCCACCATGTTCTTTTTCGATCCCTCG
 GGCAATCCCATCGAGGTCAAGGGCTTTGCCGATCTCCAGCAGGTCTACGCCCG
 ATGAGC3'

>BjrC68 undefined product 2559916:2560015 reverse
 5'GCATGCCGCGTTTCGTATCAGCTGGATTACGACCAGGGTGGTGAAGTCGAGGG
 CTATTGGGGAACCTGCGGCGAGACCAGCGTGGCCAACGTGACCTTGAT3'

>Bp1_162 undefined product 4404895:4404959 reverse
 5'ACTCGATGGTGCATTACAAGCTGGGTGACAGCGTCCGTGCGCTTGATGCGCA
 TGGAATGCCTTG3'

>cspA undefined product 1599563:1599935 reverse
 5'ATCAGTTACATCTTGTTACTGTTGGCCGATAGAAGGGGGCACGGTTTAAGTGG
 GGCTGCGTGACGCGCGATGTTTCTTGCTTTAACGGCTCGTGTGATAGTTTATAT
 TTGAAAGATGCTTTTATGGCAACAGGTAAGTGTCAAGTGGTTCAACGATTCCAAAG
 GCTTCGGCTTTATCACTCCGGATGACGGCGGGCAAGATCTGTTCCGCCACTTCT
 CCGCAATCCAGATGAACGGCTTCAAGACCCTGAAAGAAGGTCAAAAAGTCCAGT
 TTGACGTCACGCAAGGCCCGAAGGGCAAGCAAGCATCGAACATCCAGAACGCC
 GCCTAAACCCCGGTCTTCTCTGAAAAGCCCCGCCCGCGGGGCTTTTTTTC3'

>crcB undefined product 2800602:2800675 forward
 5'ATGTCTCCGATGACCTTCTGGCGCAACCCTCGATCTGGCTGGATCAGTTGTC
 TCGTCATCGCGGCACGATTTTC3'

>CC3513 undefined product 5090400:5090549 forward
 5'CCGCATGCAGATCGCGCTGTCGAAATGGTGGTTCGAGGGCATTTCGACCAAC
 ATCCCGCTGCACCGCGAACTGATGGTGGATGCGCGCTTCTTCGAAGGTGGAAC
 CAATATCCATTATCTGGAACATAAGTTGTCGGAACGTCCGGCTTC3'

>CC1840 undefined product 3382967:3383058 forward
 5'GAACCGCCGTTGTTCGGCCAGACAAGTCTGGCCGACAACGGCGGGCGACCGAT
 CAGCGCTTGTTCGATCGGCGGCACGTTCGCGACGCGGGCGAGCC3'

>Chlorobi-RRM undefined product 3830675:3830742 reverse
 5'ACCCACTTGCGCGGGGTTCGATTGCAGGAAATCCATGCCCTGGGTTTCCTGCT
 GGTGCTTTTTTCTT3'

>Cobalamin undefined product 3038673:3038919 reverse

5'GAACTGGACAAGACCTTCGAGGAACGCGACCACATCAACACCGCCATCGTCA
GCGCCATCGACGAGTCGGCCGAGAAGTGGGGCGTGAAGGTGCTGCGCTACGA
GATCAAGGACCTGACGCCGCCAAGGAAATCCTGCACGCCATGCAGGCCAGCA
TTACCGCCGAGCGTGAGAAGCGCGCCCTCATCGCCGCCTCCGAAGGCCGCAA
GCAGGAACAGATCAATATCGCCACCGGTGAGCGCGAAG3'

>ctRNA_pGA1 undefined product 3317375:3317443 forward

5'GCTTGGTAATCGACGGATCAAAAACAATCGCCCGACAGGTTTGCGCCTGTC
GGGCGATTTTTTCAGCT3'

>cyano_tmRNA undefined product 3533778:3533837 forward

5'CCAATTCACCCACATATCCATGCGGGTCGCTGGTGGAGGTAAGCGGGATC
GAACCGCT3'

>drz-agam-1 undefined product 1577836:1577922 reverse

5'GGTGTGACCACCAGGGCGATCTCGTTGTTCTTGACCATGTCCACGATGTGC
GGACGGCCTTCCACCACCTTGTTGACGGTGGCCAC3'

>lsrD undefined product 1093595:1093656 forward

5'ATTATTTATACGATTTCCGCTTCGGGAACAGCAAAAAGGCCAGCGCGAAGCTG
GCCTTTTCG3'

>lsrG undefined product 3775631:3775718 reverse

5'GGCTGCGGCAACCCAGGCCGCTGCTGCGGCAGCACCCGCACCTGCGGGCGG
CGTCTGCCTCTACTACGGCAGCGCCTGCTGCTGCAGCC3'

>lsrK undefined product 3533630:3533695 reverse

5'GTGAGTACGGGGTTCGAATCCCCAGGGGACGCCAAATAAAAAAGCCCGCG
CAAGCGGGCTTTTT3'

>MicC undefined product 2332590:2332667 forward

5'GGCGGCACATCGGCTCACAGGGAGGGAACGCATTCTGCTAGGATCAGTCCTA
TCAGCTCGTTTCGTCCCGTGCGCCGA3'

>ncr1241 undefined product 3807488:3807558 forward

5'GCAAAGGCTTTCTGATTCTCCCCGTTTTCCCTTGCAAGAGGCCGGCTGGATG
CCGGCCTCTTGCAATTTG3'

>P13 undefined product 3578845:3578907 reverse

5'TGCCGCCGGGGAGGGCGGCGCCTGCCCCAGCTCGGGCGCGACGCGCTGGA
TCAGCGCCGGCGT3'

>P18 undefined product 4331255:4331365 reverse

5'GTGCACATGACGCGCAGCACGCTCTATGTGCTGCTGCCGCTGTCCGGTGGTGT
TGGCGCTGGTGGTGCAGCAGGGCAGCATCCAGAACTTCCGCGCCTACCAG
GACGTG3'

>P31 undefined product 5496836:5496908 forward

5'AAATTACCGCCGGGAAATCACCTGAAAAGGTGATTTCCGGCGGTAAAGGGAT
CATCGTCGCCGACGATCCCGA3'

>pfl undefined product 1381288:1381387 forward

5'TTCGTCTCACGTGACTGGCGAAACGCCGGCAAAGCCGGCCAAGATGGGGAT
CCATCGGGAAGCGTGAGATTTCAACAGCCGTGCGCCTGGGCAGCCGAA3'

>PyrR undefined product 752778:752894 reverse

5'ATAACCTTTTCGGTCGCCC GCACAGCATGCGGGCAAGCGATGCCAACGGTAGC
AAACATCCTGCCATCTAGGCTTTACAATAAGAACCTCTTTCGCTTGGCGAAAGAG
GTTTTTTTTT3'

>rimP undefined product 2691772:2691837 reverse

AAAGCCCCGGTTTTCCGGGGCTTTTCATTTATCGTAAAGTTTGAATAATGGCTC
TTAACATCCTT

>RNAI undefined product 1330649:1330725 reverse

5'AGAACTTCGCCTGATCTTCAGCATCAGCGACGGCCTTCCCGGCAAGAAAAAC
CACGCGCAAGCGTGTTTTTTTTT3'

>rsmX undefined product 4409813:4409948 reverse

5'ACCTGGGTGCAGCCGAGATGCCGGCGCCGGTCAAGGGCTTGATGAAGGCCA
TGGCCAAGGTCATGACCACGGTGGCTTACCGCATCTGAGGCAGGGCCTGGCG
GGACATGAAAAAGCGGCCTCGGCCGCTTTTTT3'

>sau-50 undefined product 2940369:2940517 forward

5'TAAATTTTCGGAACCTTTGCTGCCTTGATCTTGATCGCTGCACCAGTCTGCGGA
TTGCGACCGGTACGGGCTGCGCGCTTGCCACGGCGAAGGTGCCGAAACCTA
CCAGCGTCACCGTGCCATTCTTCTTCAGCGTGGTCTTTACTGCA3'

>SAH_riboswitch undefined product 4908117:4908206 reverse

5'CGCTGCAAGGAGCGTTGCGACAGGGACATCCCCTGCCAGGCTTGAAGCGGT
GAACAGGCGCACGTGCGCTGGCAACCGCGCTCACGTTAC3'

>SpF25_sRNA undefined product 1792781:1792848 reverse

5'TGACAGCGCCGGTATGACGGCGCTGCGAATGTTGAAAAGCCACGAGCGATCC
GTGGCTTTTTTTCAT3'

>SpF36_sRNA undefined product 4450248:4450298 forward

5'CAAAAAGCCTGATCCCGGATGAGATCAGGCTTTTTTATTTCCAAAAGAA3'

>sR45 undefined product 5447390:5447440 reverse

5'AGCTTTGGCGAGGTCGAGGCCATGAACCGCGACGCCTTCGCCTCGCTGACC3
,

>STnc240 undefined product 3961785:3961863 forward

5'ACTGCGTCAAGCAGTAAGCCTCAGCGCTGCCAGTAAACAAACAGCCGATGC
CTGCGAGGGCATCGGCTGTTTGTITTTT3'

>STnc310 STnc310 undefined product 5371582:5371642 forward

5'AGCGAGATGCACCGGCGCAAACGCCGCCAGGTGGATGCCTGGCCGGCGTT
TTTCATGGGC3'

>STnc320 undefined product 5252462:5252517 reverse

5'TTGATGCCTCAACAAGAAAGCGGGCGCACCCACCGGGGTGCGCCGCTTTCTTA
TGTG3'

>STnc350 undefined product 4257834:4257901 reverse

5'TGGCGCTGAAGTGAAGTAGCAAGGAAGGAAGCAAAGAACGGCCGCATCAGC
GGCCGTTCTGTTTTT3'

>STnc370 undefined product 3751391:3751443 forward

5'AATCTTCGGATTATTTTCGAAAAAACCGGCCGCAAGGCCGTTTTTTTTTCGT3
,

>STnc430 undefined product 245863:245950 reverse

5'AAGAACTAAAACTTAAAACTAAAGAGCTAAAAACCAAGAAAAATAAAAAACGC
C
GGCTCGCAAGCCGGCGTTTCCTCATTGC3'

>STnc50 undefined product 3365471:3365564 forward

5'TCTGGGCATCGTAGGCGGCAACCGCGACTGAACCTTGCCTGAACTGCAATAAACCGC
ATAAAAAAGCCGGCGCAAGCCGGCTTTTTTACATCG3'

>suhB undefined product 2556514:2556590 reverse

5'AGATTCCCCCGATCTTCGGCGTCTTATATCCCTCCCCTCAGGCCGCCTCCCC
AGGCGGCCATTTTTTTTATCCTTG3'

>sX11 undefined product 2548906:2549016 forward

5'GGGCAGGCCGCGCCGCAGGCTGACCAGTTCCAGCGTGCGGGCGCAGGGCGG
TGAGCATGCCAGCGCCAGCACATCGACCTTGAGCAGCTTGAGCGACTCCAGG
TCATCCTT3'

>sX4 undefined product 109972:110084 forward

5'TGAGTTCCCTGCACGGAGAGGTGGCCGAGTGGTTAATGGCAGCAGACTGTAA
ATCTGCCCTCTTACGAGTACGCTGGTTCGAATCCAGCCCTCTCCACCATTAGGC
CGGGAGT3'

>sX6 undefined product 1411810:1412066 reverse

5'TCAGCGAAGAAGACGTGATCAAGAAGATCATGCTCAAGGACACCGTCGATGG
CGAATTCACCACCACCCAGGACGTGGCCCAAACCGCCGTCTTCCTGGCGGCCT
TCCCGACCAATGCTCTGACCGGCCAGTCGGTGGTGGTCAGCCACGGCTGGCAT

ATGCAGTAAGCCGCGCCCGCTGCGACAACCATAGTCGCAACGCGTACAGCCGC
CCCGGTGCAGACCGGGGCGGCTTTTTTTGTGCTCTCACCTGCGCTT3'

>symR undefined product 3045259:3045340 reverse

5'CGCCATCCGGTCTGCTGCTGCAGAGGCAGGCCGATGCAGTGATGGCCAATG
CGCCCGGGTTCGCGCCCGGGCGCATTGCGCT3'

>TarA undefined product 655476:655577 forward

5'GTGTTGGAAGTCGAAATATTTCTATTCAATATCGAATAATTTTCGATTGACAAT
TTTTTTGGGTCTTTTGCCCGAAATTGGCAGGCCTTGGGGCTTGCTTG3'

>Xoo2 undefined product 3686760:3686833 forward

5'ACGTAAGCAATACGTAGGCAACTCGCTGCACAAACAAAAACAGCAGGCCGCA
TCACCGGGTCTGCTGTTTTTTC3'

>ykkC-III undefined product 4503620:4503684 forward

5'ACATCATGATTCTCCATCATCGCCGCGCGCCGTACCCGGTGTGGCGTGCGCT
GGCACGGCGCGGC3'

>yybP-ykoY undefined product 576481:576664 forward

5'AATGCGGCCTTCTCTTTGGGGAGTAGCCAGCTTCCGGATCATCCGGAAGAGC
CGTGTCAACATTCTCGGCAGCAGCAGCTGCCGTGGCGCGGGTAGCCAATCGGT
AGGCGAGACCATAGACGTTTCCTGCGGCCAGGCTGGGCGCGCAGGCCGCGTCC
ATGGACTTTTCGCCCCGGCCAAGGAAT3'

APÊNDICE II – NÚMERO DE mRNAs ALVOS PREDITOS DOS 173 ncRNAs DE *H. seropedicae* SmR1.

ncRNA	Posição Inicial	Posição Final	Tamanho em nucleotídeos	Número de alvos
				p-value: 0,05
Hs_noco_471	219058	219229	180	0 Alvos
Hs_noco_494	457262	457311	50	11 Alvos
Hs_noco_525	932018	932172	70	0 Alvos
Hs_noco_530	957469	957626	78	40 Alvos
Hs_noco_534	1067480	1067693	69	13 Alvos
Hs_noco_560	1495976	1496389	366	5 Alvos
Hs_noco_637	2815035	2815266	145	0 Alvos
Hs_noco_651	3173522	3173721	158	7 Alvos
Hs_noco_682	3633077	3633378	300	64 Alvos
Hs_noco_698	3779834	3780049	158	13 Alvos
Hs_noco_700	3791718	3791849	82	0 Alvos
Hs_noco_704	3879771	3880105	303	32 Alvos
Hs_noco_709	3928269	3928556	182	11 Alvos
Hs_noco_719	4074941	4075134	69	5 Alvos
Hs_noco_722	4105880	4106065	135	0 Alvos
Hs_noco_723	4127435	4127634	118	0 Alvos
Hs_noco_726	4148338	4148890	553	32 Alvos
Hs_noco_732	4239065	4239615	510	12 Alvos
Hs_noco_742	4396086	4396487	402	53 Alvos
Hs_noco_744	4417512	4417740	104	0 Alvos
Hs_noco_746	4428208	4428652	392	6 Alvos
Hs_noco_748	4467236	4467750	510	43 Alvos
Hs_noco_749	4473739	4474140	340	99 Alvos
Hs_noco_750	4474754	4474942	141	7 Alvos
Hs_noco_761	4618606	4618928	308	77 Alvos
Hs_noco_788	5026868	5027073	119	52 Alvos
Hs_noco_802	5239833	5240191	323	20 Alvos
Hs_noco_820	5397562	5398115	525	12 Alvos
Hs_sRNA_4.2	2234033	2234214	182	16 Alvos
Hs_sRNA_19.2	612014	612436	423	0 Alvos
Hs_sRNA_2.1	4108897	4109042	146	77 Alvos
TPP	462479	462706	293	33 Alvos
Hs_sRNA_Gs6	1915816	1915944	129	45 Alvos
Hs_sRNA_Gs10	808782	808981	200	13 Alvos
Hs_sRNA_Gs12	2233545	2233793	249	15 Alvos
Hs_sRNA_Gs36	1788675	1788860	186	19 Alvos
Hs_sRNA_8.2	5335529	5335812	284	0 Alvos
Hs_noco_498	551015	551235	221	0 Alvos
Hs_noco_498	551015	551235	221	0 Alvos

smallRNA_49.2	4896565	4896712	148	0 Alvos
Hs_noco_780	4896320	4896704	385	18 Alvos
Hs_noco_786	5019358	5019702	345	0 Alvos
Hs_noco_757	4564996	4565251	256	8 Alvos
Hs_noco_725	4136562	4136795	234	0 Alvos
Hs_noco_714	3998445	3998747	60	0 Alvos
Hs_noco_640	2939610	2939824	141	3 Alvos
Hs_noco_644	2972948	2973103	156	5 Alvos
Hs_noco_602	2140731	2141104	307	4 Alvos
Hs_noco_604	2177189	2177339	84	14 Alvos
Hs_noco_564	1587873	1588410	468	31 Alvos
Hs_noco_565	1596987	1597381	289	22 Alvos
Hs_noco_568	1636498	1636822	325	23 Alvos
Hs_noco_593	2032308	2032540	136	14 Alvos
Hs_noco_551	1383591	1383913	323	72 Alvos
Hs_noco_555	1433080	1433239	74	16 Alvos
Hs_noco_526	935305	935540	185	11 Alvos
Hs_noco_490	417253	417332	80	1 Alvo
Hs_noco_514	808928	809191	169	24 Alvos
Hs_noco_724	4127694	4127848	41	9 Alvos
"CUFF.55"	205713	205798	86	14 Alvos
"CUFF.111"	359683	359916	234	12 Alvos
"CUFF.124"	412442	412676	235	33 Alvos
"CUFF.223"	664741	664801	61	17 Alvos
"CUFF.251"	741007	741139	131	10 Alvos
"CUFF.273"	792569	792756	102	15 Alvos
"CUFF.395"	1155828	1155922	95	85 Alvos
"CUFF.404"	1189098	1189333	236	53 Alvos
"CUFF.504"	1336439	1336564	126	0 Alvos
"CUFF.522"	1381282	1381477	54	30 Alvos
"CUFF.524"	1382857	1382981	125	0 Alvos
"CUFF.689"	1867577	1867679	103	0 Alvos
"CUFF.690"	1880603	1880826	224	10 Alvos
"CUFF.738"	2011544	2011601	58	33 Alvos
"CUFF.829"	2225907	2226140	142	15 Alvos
"CUFF.852"	2287397	2287464	68	0 Alvos
"CUFF.934"	2519484	2519726	243	0 Alvos
"CUFF.938"	2526174	2526304	131	0 Alvos
"CUFF.967"	2578050	2578397	348	26 Alvos
"CUFF.1011"	2800595	2800705	111	16 Alvos
"CUFF.1094"	3160202	3160357	156	0 Alvos
"CUFF.1117"	3204507	3204585	79	44 Alvos
"CUFF.1142"	3303508	3303560	53	11 Alvos
"CUFF.1223"	3512060	3512227	168	0 Alvos
"CUFF.1716"	4925909	4926013	105	4 Alvos
"CUFF.598"	3781409	3781584	176	0 Alvos

"CUFF.225"	539844	539917	74	0 Alvos
"CUFF.228"	541704	541796	93	20 Alvos
"CUFF.1027"	3181155	3181240	86	23 Alvos
"CUFF.1106"	3382844	3382966	123	5 Alvos
"CUFF.1109"	3382646	3382739	94	25 Alvos
"CUFF.1158"	3553031	3553120	90	2 Alvos
"CUFF.1190"	3614107	3614194	88	0 Alvos
"CUFF.1194"	3630870	3630996	127	44 Alvos
"CUFF.1510"	4402051	4402283	233	0 Alvos
"CUFF.1548"	4499669	4499780	112	11 Alvos
"CUFF.1570"	4587035	4587149	115	23 Alvos
CUFF.1692	4834369	4834486	118	0 Alvos
"CUFF.1713"	4909594	4909823	230	95 Alvos
"CUFF.1720"	4929103	4929157	55	13 Alvos
"CUFF.1751"	5000211	5000493	283	27 Alvos
CUFF.17.51	5000211	5000493	283	27 Alvos
CUFF.1692	4834369	4834486	118	0 Alvos
CUFF.839	2225999	2226140	142	15 Alvos
Hs_noco_479	328428	328515	90	9 Alvos
Hs_noco_516	825901	826300	400	31 Alvos
Hs_noco_517	831492	831611	120	14 Alvos
Hs_noco_519	853978	854161	184	0 Alvos
Hs_noco_567	1624918	1624991	74	19 Alvos
Hs_noco_571	1664205	1664460	256	0 Alvos
Hs_noco_584	1880142	1880548	407	6 Alvos
Hs_noco_631	2579377	2579662	286	23 Alvos
Hs_noco_645	3042974	3043384	441	38 Alvos
Hs_noco_654	3188002	3188408	407	0 Alvos
Hs_noco_662	3290469	3290741	273	20 Alvos
Hs_noco_663	3317231	3317432	202	28 Alvos
Hs_noco_664	3325609	3325783	175	0 Alvos
Hs_noco_671	3392204	3392349	146	8 Alvos
Hs_noco_692	3750872	3751432	561	24 Alvos
Hs_noco_697	3776244	3776620	377	1 Alvo
Hs_noco_712	3989988	3990380	393	10 Alvos
Hs_noco_713	3995857	3996132	276	31 Alvos
Hs_noco_715	4004964	4005107	144	4 Alvos
Hs_noco_716	4010423	4010502	80	0 Alvos
Hs_noco_731	4232898	4233211	314	9 Alvos
Hs_noco_735	4319058	4319138	81	0 Alvos
Hs_noco_736	4321024	4321152	129	0 Alvos
Hs_noco_741	4388096	4388200	105	11 Alvos
Hs_noco_743	4408988	4409170	183	9 Alvos
Hs_noco_753	4489489	4489615	127	0 Alvos
Hs_noco_758	4565726	4565818	93	2 Alvos
Hs_noco_790	5118375	5118838	464	37 Alvos

Hs_noco_819	5373859	5374363	505	19 Alvos
Hs_noco_824	5433895	5434068	174	13 Alvos
Hs_noco_834	5511199	5511776	578	22 Alvos
Hs_small_1.2	4408980	4409076	97	5 Alvos
Hs_small_16.2	4873548	4873707	160	18 Alvos
Hs_small_827	5475789	5476188	400	35 Alvos
Hs_sRNA_10.2	4896557	4896712	156	19 Alvos
Hs_sRNA_27.2	1943525	1943726	155	0 Alvos
Hs_sRNA_3.2	3815799	3815921	123	17 Alvos
Hs_sRNA_32.2	2111171	2111326	156	0 Alvos
Hs_sRNA_45.2	2140728	2141053	326	11 Alvos
Hs_sRNA_47.2	4136551	4136707	157	0 Alvos
Hs_sRNA_48.2	2972997	2973111	115	2 Alvos
Hs_sRNA_50.2	551056	551240	185	5 Alvos
Hs_sRNA_52.2	5058503	5058597	95	24 Alvos
Hs_sRNA_53.2	1383582	1383798	217	72 Alvos
Hs_sRNA_57.2	417221	417409	189	12 Alvos
Hs_sRNA_6.2	2939838	2939938	101	3 Alvos
Hs_sRNA_Gs1	3998440	3998528	89	5 Alvos
Hs_sRNA_Gs11	808938	809200	263	24 Alvos
Hs_sRNA_Gs13	4872868	4872949	82	21 Alvos
Hs_sRNA_Gs14	4873545	4873711	167	18 Alvos
Hs_sRNA_Gs2	3990250	3990386	137	5 Alvos
Hs_sRNA_Gs21	2939764	2940103	340	21 Alvos
Hs_sRNA_Gs23	935267	935546	280	0 Alvos
Hs_sRNA_Gs25	1433087	1433246	160	20 Alvos
Hs_sRNA_Gs27	5019552	5019706	202	0 Alvos
Hs_sRNA_Gs31	1596981	1597171	191	25 Alvos
Hs_sRNA_Gs33	2032443	2032547	105	35 Alvos
Hs_sRNA_Gs34	2032256	2032442	187	28 Alvos
Hs_sRNA_Gs4	4127685	4127782	98	36 Alvos
Hs_sRNA_Gs5	2177255	2177347	93	7 Alvos
Hs_sRNA_Gs7	1915912	1916133	222	0 Alvos
Hs_sRNA_Gs8	5433936	5434088	153	19 Alvos
smallRNA_30.2	2973008	2973111	104	38 Alvos
smallRNA_43.2	4564951	4565134	184	26 Alvos
smallRNA_44.2	4564991	4565134	144	6 Alvos
sRNA46	1636668	1636831	164	25 Alvos
tmRNA_tmRNA	3942038	4239574	537	11 Alvos
CsrC_CsrC	3382646	3383016	371	20 Alvos

Os alvos foram preditos pela ferramenta TartgetRNA2.

APÊNDICE III – Os 57 NOVOS ncRNAs PREDITOS EM *H. seropedicae* SmR1

ncRNAs	Rfam 12.0	Tamanho	Posição Inicial	Posição Final	Sentido
5 ureB sRNA	RF02514	286	5416169	5415884	Antisense
6S	RF00013	177	1383796	1383620	Antisense
Afu_254	RF01504	72	4233427	4233498	sense
alpha tmRNA	RF01849	104	2687581	2687684	sense
AS1726	RF01779	69	1536442	1536374	Antisense
ar35	RF02346	162	3551853	3552014	sense
ASdes	RF01781	65	788645	788709	sense
Bacteria_large_SRP	RF01854	96	4109005	4108910	Antisense
Bacteria_small_SRP	RF00169	99	1070733	4108907	Antisense
Betaproteobacteria_toxic_sRNA	RF02278	65	4409043	4408979	Antisense
beta tmRNA	RF01850	322	2687527	2687848	sense
BjrC68	RF02353	100	2560015	2559916	Antisense
Bp1_162	RF02420	65	4404959	4404895	Antisense
cspA	RF01766	373	1599935	1599563	Antisense
crcB	RF01734	74	2800602	2800675	sense
CC3513	RF01528	150	5090400	5090549	sense
CC1840	RF01521	92	3382967	3383058	sense

Chlorobi-RRM	RF01697	68	3830742	3830675	Antisense
Cobalamin	RF00174	247	3038919	3038673	Antisense
ctRNA_pGA1	RF00236	69	3317375	3317443	sense
cyano_tmRNA	RF01851	122	2687572	2687693	sense
drz-agam-1	RF01787	87	3936052	3935966	Antisense
isrD	RF01388	62	1093595	1093656	sense
isrG	RF01390	88	3775718	3775631	Antisense
isrK	RF01394	66	3533695	3533630	Antisense
MicC	RF00121	97	1775910	1775814	Antisense
ncr1241	RF02451	71	3807488	3807558	sense
P13	RF01665	63	3578907	3578845	Antisense
P18	RF01671	110	4331365	4331255	Antisense
P31	RF01676	73	5496836	5496908	sense

pfl	RF01750	100	1381288	1381387	sense
PyrR	RF00515	117	752894	752778	Antisense
rimP	RF01770	73	48992	48920	Antisense
RNAI	RF00106	77	1330725	1330649	Antisense
rsmX	RF02144	136	4409948	4409813	Antisense
SAH riboswitch	RF01057	90	4908206	4908117	Antisense
SpF25_sRNA	RF02432	68	1792848	1792781	Antisense
SpF36_sRNA	RF02433	51	4450248	4450298	sense
STnc240	RF02074	79	3961785	3961863	sense
STnc310	RF02067	61	5371582	5371642	sense
STnc320	RF02066	56	5252517	5252462	Antisense
STnc350	RF02063	68	4257901	4257834	Antisense
STnc370	RF02064	53	3751391	3751443	sense
STnc430	RF02053	88	245950	245863	Antisense
STnc50	RF02059	94	3365471	3365564	sense
suhB	RF00519	77	2556590	2556514	Antisense
sX11	RF02230	111	2548906	2549016	sense
sX4	RF02223	113	109972	110084	sense
sX6	RF02225	257	1412066	1411810	Antisense
symR	RF01809	82	3045340	3045259	Antisense
TarA	RF02537	102	655476	655577	sense
Xoo2	RF02241	74	3686760	3686833	sense
Yfr10	RF02362	56	4959639	4959584	Antisense
ykkC-III	RF01763	65	4503620	4503684	sense
yybP-ykoY	RF00080	184	576481	576664	sense

ncRNAs	Código do Rfam 12.0	Descrição	Número de alvos
5_ureB_sRNA	RF02514	ncRNA 5' ureB_s RNA, localizado na 5' do gene ureB compõem clusters de genes da urease e é antisenso ao ureB.	Ausência
6S	RF00013	6S / SsrS RNA	84 Alvos
Afu_254	RF01504	ncRNA Afu_254 (<i>A. fumigatus</i>)	32 Alvos (recA)
alpha_tmRNA	RF01849	mRNA de transferência (Alphaproteobacteria)	Ausência
AS1726	RF01779	AS1726 ncRNA (<i>M. tuberculosis</i>)	4 Alvos
ar35	RF02346	Alphaproteobacterial ncRNA ar35	Ausência
ASdes	RF01781	ASdes TB ncRNA	Ausência
Bacteria_large_SRP	RF01854	Partícula de reconhecimento de sinal ncRNA bacteriano	76 Alvos
Bacteria_small_SRP	RF00169	Partícula de reconhecimento de sinal ncRNA bacteriano	76 Alvos
Betaproteobacteria_toxic_sRNA	RF02278	Betaproteobacteria toxic small RNA (tncRNA)	12 Alvos
beta_tmRNA	RF01850	mRNA de transferência (Betaproteobacteria)	7 Alvos
BjrC68	RF02353	ncRNA BjrC68 (<i>Bradyrhizobiaceae</i>)	1 Alvo
Bpl_162	RF02420	ncRNA Bpl_Cand612_SIPHT (<i>Burkholderia</i>)	54 Alvos
cspA	RF01766	cspA termoregulador	Ausência
crcB	RF01734	crcB RNA	32 Alvos (livGM, narU)

CC3513	RF01528	ncRNA CC3513 (<i>caulobacter</i>)	Ausência
CC1840	RF01521	ncRNA CC1840 (<i>caulobacter</i>)	8 Alvos
Chlorobi-RRM	RF01697	Chlorobi-RRM RNA	37 Alvos
Cobalamin	RF00174	Cobalamina Riboswitch é um elemento cis-regulador amplamente distribuído em regiões 5' UTR relacionados à vitamina B12 em bactérias	12 Alvos
ctRNA_pGAl	RF00236	CtRNA	Ausência
cyano_tmRNA	RF01851	mRNA de transferência = 10S RNA ou SsrA (<i>Cyanobacteria</i>)	36 Alvos
drz-agam-1	RF01787	Riboswitch drz-agam-1	30 Alvos
isrD	RF01388	Interage com a chaperona Hfq	Ausência
isrG	RF01390	Interage com a chaperona Hfq	12 Alvos
isrK	RF01394	Interage com a chaperona Hfq	56 Alvos
MicC	RF00121	ncRNA MicC	30 Alvos
ncr1241	RF02451	Bacillus ncRNA ncr1241	Ausência
P13	RF01665	ncRNA P13 (<i>Pseudomonas</i>)	Ausência
P18	RF01671	ncRNA P18 (<i>Pseudomonas</i>)	Ausência
P31	RF01676	ncRNA P31 (<i>Pseudomonas</i>)	9 Alvos
pfl	RF01750	ncRNA pfl	2 Alvos
PyrR	RF00515	PyrR possui sítio de ligação	1 Alvos
rimP	RF01770	ncRNA rimP (<i>Gammaproteobacteria</i>)	Ausência
RNAI	RF00106	RNAI	Ausência
rsmX	RF02144	rsmX	Ausência
SAH riboswitch	RF01057	Riboswitch S-adenosil-L-homocisteína	68 Alvos
SpF25_sRNA	RF02432	ncRNA SpF25 (<i>Streptococcus</i>)	1 Alvo
SpF36_sRNA	RF02433	ncRNA SpF36 (<i>Streptococcus</i>)	Ausência
STnc240	RF02074	ncRNA STnc240 (<i>Enterobacteria</i>)	9 Alvos
STnc310	RF02067	ncRNA STnc310 (<i>Salmonella enterica</i>)	Ausência
STnc320	RF02066	ncRNA STnc320 (<i>Salmonella enterica</i>)	20 Alvos
STnc350	RF02063	ncRNA STnc350 (<i>Salmonella enterica</i>)	Ausência
STnc370	RF02064	ncRNA STnc370 (<i>Enterobacteria</i>)	1 Alvo (acpD)

STnc430	RF02053	ncRNA STnc430 (Enterobacteria)	10 Alvos
STnc50	RF02059	ncRNA STnc40 (<i>Salmonella enterica</i>)	20 Alvos
suhB	RF00519	ncRNA	25 Alvos
sX11	RF02230	ncRNA sX11 (Proteobacteria)	28 Alvos
sX4	RF02223	ncRNA sX4 (Proteobacteria)	16 Alvos
sX6	RF02225	ncRNA sX6 (Proteobacteria)	Ausência
symR	RF01809	SymR antitoxina	Ausência
TarA	RF02537	RNA TarA ativado do (<i>Vibrio</i>) ToxT	Ausência
Xoo2	RF02241	ncRNA Xoo2 (<i>Xanthomonadaceae</i>)	Ausência
Yfr10	RF02362	RNA 10 funcional (Cyanobacterial)	10 Alvos
ykkC-III	RF01763	ykkC-III RNA	Ausência
yybP-ykoY	RF00080	yybP-ykoY (SraF)	Ausência

Os alvos foram preditos pela ferramenta TartgetRNA2

ncRNAs Rfam 12.0	Vizinhos Sense	Vizinhos antisense
5 ureB sRNA RF02514	familia ure (A,B,C,D,E,F,G,J), glpD1, Hseo_4727)	glpD1, Hsero_4227
6S RF00013	Hsero_1224, glyA, tsaA, tgt, CUFF_552	exbD, tolQ, Hsero_1221,Hsero_1222, Hsero 1225, Hsero_1226, recG
Afu 254 RF01504	Hsero_3696, Hsero_3699, ubiG, gph	noco_731, gyrA, Hsero_3693, serC
alpha tmRNA RF01849	pvdE, Hsero_2352, Hsero_2353, Hsero_2354, Hsero_2355	fhuA, Hsero_2346, Hsero_2347, Hsero_2348, Hsero_2356
AS1726 RF01779	Hsero_1343, ncnH, Hsero_1347, fabG, Hsero_1349, Hsero_1350, Hsero_1351, Hsero_1352	OmpC, Hsero_1354
ar35 RF02346	Hsero_3091, Hsero_3092, panC, cheA, cheZ, Hsero_3098, cysI, Hsero_3100, cysH, cysD, cysN	
ASdes RF01781	Hsero_0725, Hsero_0726, Hsero_0727	Hsero_0724, Hsero_0721, miaB, phoH, tsr, corC, lnt
Bacteria large SRP RF01854	Hsero_3579, glk, Hsero_3582	rhlE, Hsero_3578, noco_782, 4.5S, ubiD, Hsero_3583, nusB, ribH, ribAB.
Bacteria_small_SRP RF00169	Hsero_3579, glk, Hsero_3582	rhlE, Hsero_3578, noco_782, 4.5S, ubiD, Hsero_3583, nusB, ribH, ribAB.
Betaproteobacteria toxi c sRNA RF02278	hemL, Hsero_3851	glpD1, gltK, gltL, pyrC, ubiF, rplM, rpsI, argC, Hsero_3845, Hsero_3847, ccmA, Hs_small_1.2, Hs_noco_743, gltJ, Hsero_3857
beta tmRNA RF01850	pvdE, Hsero_2352, Hsero_2353, Hsero_2354, Hsero_2355	fhuA, Hsero_2346, Hsero_2347, Hsero_2348, Hsero_2356
BjrC68 RF02353	Hsero_2265, Hsero_2270, Hsero_2271	poxB, Hsero_2267, Hsero_2266, Hsero_2264, Hsero_2263, Hsero_2262
Bp1 162 RF02420	thiD hemL, Hsero_3851	pilI, Hsero_3842, Hsero_3845, ccmA, Hsero_3847, argC, rpsI, rplM, ubiF, Hs_small_1.2, Hs_noco_743
cspA RF01766	rpsT, icd, clpS, clpA	Hsero_1400, cspD, rsuA, Hsero_ 1394, mviN, Hsero_1391, Hs_ncRNA_Gs31, Hs_ noco_565.
crcB RF01734	Hsero_2465, eriC, Hsero_2469, CUFF.1011	Hsero_2471, cheY, Hsero_2468, Hsero_2466, cyoA, cyoB

CC3513 RF01528	Hsero_4449, rhaI, Hsero_4458	Hsero_4451, Hsero_4452, Hsero_4455, serA, Hsero_4456, Hsero_4457
CC1840 RF01521	CsrC	sucA, sucB, gltA, Hsero_2971, sdhA, sdhB, sdhC, Hsero_297, CUFF.1109, CUFF.1106
Chlorobi-RRM RF01697	serA, Hsero_3338, Hsero_3340, Hsero_3337, Hsero_3339, Hsero_3340	Hsero_3344, Hsero_3343, betA, Hsero_3341
Cobalamin RF00174	Hsero_2660, Hs_noco_645	pgm, cheM, Hsero_2661, Hsero_2669, Hsero_2658, cobN
ctRNA_pGAl RF00236	ppa, hemy	prfC, Hs_noco_663
cyano_tmRNA RF01851	pvdE, Hsero_2350, Hsero_2351	Hsero_2348
drz-agam-1 RF01787	Hsero_1381, Hsero_1376	gecG, carA, carB, greA, ftsJ, hflB, folP
isrD RF01388	Hsero_0991, Hsero_0995, mmr	groEL1, Hsero_0993, Hsero_0992, Hsero_0989, Hsero_0988, mltB
isrG RF01390	Hs_noco_697	fimV, asd, leuB, leuD, trupA, trupF, trupB
isrK RF01394	Hsero_3071, era, rpoE, Hsero_3074, metG	metG, mrp, gltX, caiD, Hsero_3075
MicC RF00121	rscB (CheY, OmpR, NtrC, PhoB), gst	fumA, murI, Hsero_1539, trnH
ncr1241 RF02451	Hsero_3317, Hsero_3318, Hsero_3319, nhaP	Hsero_3324, Hsero_3323, Hsero_3322, Hsero_3321, Hsero_3316, phaB
P13 RF01665	prpR, Hsero_3123, cheD	Hsero_3118, prpC, prpB, Hsero_3122, ntrC
P18 RF01671	Hsero_3774	ggt, Hsero_3780, kdpA, kdpB, kdp
P31 RF01676	mDoG	betA, Hsero_4796, Hsero_4795, Hsero_4794, Hsero_4793, Hsero_4792, Hsero_4791, gidA, gidB, blc, rhtB, para.
pfl RF01750	glyA, Hsero_1224	recG, Hsero_1226, Hsero_1225, Hsero_1222, Hsero_1221, tolQ, exbD, tolA, Hsero_1217.
PyrR RF00515	Hsero_0627, pntAa	Hsero_0695, trmU, Hsero_0693, bfd, exbB, exbD, bfr, purB, cybB, Hsero_0684, Hsero_0683, pdxH
rimP RF01770	Hsero_0038	Hsero_0037, Hsero_0036, fecI, fecR
RNAI RF00106	Hsero_1158, Hsero_1162, dadA, lpxK, kdsB, exbB2, exbD2, Hsero_1163, Hsero_1161	xseA, sadB, Hsero_1160, Hsero_1159
rsmX RF02144	Hsero_3851	gltL, pyrC, ubiF, Hs_small_1.2, Hs_noco_743
SAH riboswitch RF01057	Hsero_4286, htrB, dapF	Hsero_4294, CUFF.1713, metK, achY, Hsero_4288, set, metF, Hsero_4284
SpF25_sRNA RF02432	Hsero_1552	Hsero_1560, Hsero_1559, Hsero_1558
SpF36_sRNA RF02433	uspA	apaG, mltA, paaF

STnc240 RF02074	ediA, Hsero_3454	chrA
STnc310 RF02067	def, Hsero_4682	Hsero_4681, smf
STnc320 RF02066	Hsero_4582	Hsero_4582, Hsero_4581, Hsero_4580, Hsero_4578
STnc350 RF02063	Hsero_3718, Hsero_3719	Hsero_3720
STnc370 RF02064	Hsero_3274, Hs_noco_692	Hsero_3274, Hsero_3275
STnc430 RF02053	phbA2, Hsero_0238	Hsero_0240, Hsero_0241, Hsero_0242
STnc50 RF02059	Hsero_2957	RpoS, rumT, ndk
suhB RF00519	Hsero_2265, Hsero_2270, Hsero_2271	Hsero_2266, Hsero_2264, Hsero_2263
sX11 RF02230	Hsero_2257	dnaE, Hsero_2259
sX4 RF02223	ampG, tufB secE, trnG	
sX6 RF02225	Hsero_1247	Hsero_1246, Hsero_1245
symR RF01809	GlyA, Hsero_1224	Hsero_2668, Hsero_2665, pgm, cheM, Hsero_2661
TarA RF02537	Hsero_0596, Hsero_0598, Hsero_0599, potA, potC, Hsero_0602	leuS, rlpB, hola
Xoo2 RF02241	pyrD, ChaC, Hsero_3227, Hsero_3228	Hsero_3225
Yfr10 RF02362	rdgC, fusA	rmpH, rnpA, Hsero_4337, yidC, thdF, nfnB, ycfQ, Hsero_4330
ykkC-III RF01763	Hsero_3930, ansB, mdoG	amn
yybP-ykoY RF00080	Hsero_0527, Hsero_0528	Hsero_0529, Hsero_0530

ncRNAs Código do Rfam 12.0	Condições de Expressão
5_ureB_sRNA RF02514	AltoO2_1,AltoO2_2,Nitrato_1,Nitrato_2,PS3_1
6S RF00013	AD1_1, AD1_2, AD3_1, AD3_2, AltoO2_1, AltoO2_2, BaixoO2_1, BaixoO2_2, Nar_1, Nar_2, Nitrato_1, Nitrato_2, PS1_1, PS1_2, PS3_1, PS3_2 PT1, PT2 WR1, WR2, Controle1, Controle2
Afu_254 RF01504	AD1_1, AD1_2, AD3_1, AD3_2, AltoO2_1, AltoO2_2, BaixoO2_1, BaixoO2_2, Nar_1, Nar_2, Nitrato_1, Nitrato_2, PS1_1, PS1_2, PS3_1, PS3_2 PT1, PT2 WR1, WR2, Controle1, Controle2
alpha_tmRNA RF01849	AD3_2, AltoO2_1, BaixoO2_2
AS1726 RF01779	Ausência
ar35 RF02346	AD3_1,AD3_2, AltoO2_1,AltoO2_2, BaixoO2_1, BaixoO2_2,Nar1,Nar2,Nitrato_1,Nitrato_2,PS1_1,PS1_2,PS3_1,PS3_2,PT1,PT2, WR2,Controle1,Controle2
ASdes RF01781	Nar_1 e PS1_2
Bacteria_large_SRP RF01854	AD1_1, AD1_2, AD3_1, AD3_2, AltoO2_1, AltoO2_2, BaixoO2_1, BaixoO2_2, Nar_1, Nar_2, Nitrato_1, Nitrato_2, PS1_1, PS1_2, PS3_1, PS3_2 PT1, PT2 WR1, WR2, Controle1, Controle2
Bacteria_small_SRP RF00169	AltoO2_1
Betaproteobacteria_toxic_sRNA RF02278	AD1_1,AD3_1, AD3_2, AltoO2_1,AltoO2_2, BaixoO2_2,BaixoO2_2, Nar1,Nar2,Nitrato_1,Nitrato_2,PS1_2, PS3_2, PT1, PT2,WR2.
beta_tmRNA RF01850	AD3_1, AltoO2_1, Nitrato_1, PS1_2.
Bjrc68 RF02353	Baixo O2
Bp1_162 RF02420	AltoO2_1, BaixoO2_1, Nitrato_2, PS3_1.
cspA RF01766	AD1_1, AD1_2, AD3_1, AD3_2, AltoO2_1, AltoO2_2, BaixoO2_1, BaixoO2_2, Nar_1, Nar_2, Nitrato_1, Nitrato_2, PS1_1, PS1_2, PS3_1, PS3_2 PT1, PT2 WR1, WR2, Controle1, Controle2
crcB RF01734	AD1_2, AD3_2, AltoO2_1, AltoO2_2, BaixoO2_2, Nar_1, Nar_2, Nitrato_1, Nitrato_2,PT1, PT2, Controle1, Controle2

CC3513 RF01528	Ausência
CC1840 RF01521	AD1_1, AD1_2, AD3_1, AD3_2, AltoO2_1, AltoO2_2, BaixoO2_1, BaixoO2_2, Nar_1, Nar_2, Nitrato_1, Nitrato_2, PS1_1, PS1_2, PS3_1, PS3_2 PT1, PT2 WR1, WR2, Controle1, Controle2
Chlorobi-RRM RF01697	AD1_2, Nar_2, Nitrato_1, Nitrato_2, PS3_1, PT1, PT2 WR1, WR2, Controle1, Controle2.
Cobalamin RF00174	AD1_2, AD3_1, AD3_2, AltoO2_1, AltoO2_2, BaixoO2_1, BaixoO2_2, Nar_1, Nar_2, Nitrato_1, Nitrato_2, PS1_1, PS3_1, PT1, PT2 WR1, WR2, Controle1, Controle2
ctRNA_pGA1 RF00236	AD3_1,AD3_2, AltoO2_1,AltoO2_2, BaixoO2_2, Nar2,Nitrato1,PS3_2,PT1,PT2.
cyano_tmRNA RF01851	AD3_1,AD3-2, AltoO2_1, BaixoO2_2,PS3_1
drz-agam-1 RF01787	AD1_1, AD1_2, AD3_1, AD3_2, AltoO2_1, AltoO2_2, BaixoO2_1, BaixoO2_2, Nar_1, Nar_2, Nitrato_1, Nitrato_2, PS1_1, PS1_2, PS3_1, PS3_2 PT1, PT2 WR1, WR2, Controle1, Controle2
isrD RF01388	AD3_1, AD3_2, AltoO2_2, BaixoO2_2, Nitrato_1, PS3_1, PS3_2, PT2, WR2
isrG RF01390	AltoO2_1
isrK RF01394	AD1_1, AD1_2, AD3_1, AD3_2, AltoO2_1, AltoO2_2, BaixoO2_1, BaixoO2_2, Nar_1, Nar_2, Nitrato_1, Nitrato_2, PS1_2, PS3_1, PS3_2 PT1, PT2 WR1, WR2, Controle1, Controle2
MicC RF00121	AD1_1, AD1_2, AD3_1, AD3_2, AltoO2_1, AltoO2_2, BaixoO2_1, BaixoO2_2, Nar_1, Nar_2, Nitrato_1, Nitrato_2, PS1_1, PS1_2, PS3_1, PS3_2 PT1, PT2 WR1, WR2, Controle1, Controle2
ncr1241 RF02451	BaixoO2_1,Nar2,Nitrato2.
P13 RF01665	Ausência
P18 RF01671	BaixoO2_2,WR1,WT2.
P31 RF01676	AD3_1,AltoO2_2,BaixoO2_2,Nar_1, Nar_2, Nitrato_1, Nitrato_2,PS3_2,PT1, PT2,WR2.
pfl RF01750	AD1_1,AD3_1, AD3_2, AltoO2_1, AltoO2_2, BaixoO2_1, BaixoO2_2, Nar_1, Nitrato_1, Nitrato_2, PS1_1, PS1_2, PS3_1, PS3_2 PT1, PT2 WR1,Controle1, Controle2
PyrR RF00515	AD1_1,AD3_1, AD3_2, AltoO2_1 AltoO2_2, BaixoO2_1, BaixoO2_2, Nar1,Nar2, Nitrato_1, Nitrato_2,PS1_2,PS3_2, PT1, PT2,WR2.
rimP RF01770	Ausência
RNAI RF00106	AD3_1, AD3_2, AltoO2_1 AltoO2_2, BaixoO2_2, Nar1,Nar2, Nitrato_1, Nitrato_2,PS3_1,PS3_2,PT1, PT2,Controle1,Controle2.
rsmX RF02144	AD1_1,AltoO2_2,AltoO2_1,BaixoO2_2, Nar_1, Nitrato_1, Nitrato_2,WR1,WR2,Controle1.
SAH_riboswitch RF01057	AD1_1,AD1_2,AD3_1, AD3_2, AltoO2_1 AltoO2_2, BaixoO2_1, BaixoO2_2, Nar1,Nar2, Nitrato_1, Nitrato_2, PS1_1,PS1_2, PS3_1, PS3_2, PT1, PT2,WR1,WR2,Controle1,Controle2.
SpF25_sRNA RF02432	AD3_2,AltoO2_2, BaixoO2_1,Nitrato_1

SpF36 sRNA RF02433	AD1_1, AltoO2_1, AltoO2_2, BaixoO2_2, Nar_2, Nitrato_1, Nitrato_2, Controle2.
STnc240 RF02074	AD3_1, AD3_2, AltoO2_2, Nar_1, Nar_2, Nitrato_1, Nitrato_2, PT1, PT2.
STnc310 RF02067	AD3_2, AltoO2_2.
STnc320 RF02066	AD3_2, AltoO2_2, PS3_1.
STnc350 RF02063	AD3_2, AltoO2_1, Nar1, Nar2, Nitrato_1, Nitrato_2, PS3_1, PS3_2, PT1, PT2, WR2, Controle1, Controle2.
STnc370 RF02064	AD3_1, AltoO2_1, AltoO2_2, Nitrato_1
STnc430 RF02053	Nar2, Nitrato1
STnc50 RF02059	AD1_2, AD3_1, AD3_2, AltoO2_1, AltoO2_2, BaixoO2_1, BaixoO2_2, Nar1, Nar2, Nitrato_1, Nitrato_2, PS1_1, PS1_2, PS3_1, PT1, PT2, WR1, Controle1, Controle2.
suhB RF00519	AD1_1, AD3_1, AD3_2, AltoO2_1, AltoO2_2, BaixoO2_1, BaixoO2_2, Nar1, Nar2, Nitrato_1, Nitrato_2, PS1_1, PS1_2, PS3_1, PT1, PT2, WR1, WR2
sX11 RF02230	AltoO2_1
sX4 RF02223	AD1_1, AD1_2, AD3_1, AD3_2, AltoO2_1, AltoO2_2, BaixoO2_1, BaixoO2_2, Nar_1, Nar_2, Nitrato_1, Nitrato_2, PS1_1, PS1_2, PS3_1, PS3_2, PT1, PT2, WR1, WR2, Controle1, Controle2
sX6 RF02225	AD3_1, AD3_2, AltoO2_1, AltoO2_2, BaixoO2_1, BaixoO2_2, Nar1, Nar2, Nitrato_1, Nitrato_2, PS1_1, PS3_1, PS3_2, PT1, PT2, WR1, WR2, Controle1, Controle2.
symR RF01809	Ausência
TarA RF02537	Nar2, PT1
Xoo2 RF02241	AD3_1, AltoO2_2, PS3_1, PS3_2, PT1, PT2, WR1.
Yfr10 RF02362	AD1_1, AD1_2, AD3_1, AD3_2, AltoO2_1, BaixoO2_2, Nar1, Nar2, Nitrato_1, Nitrato_2, PS1_1, PS1_2, PS3_1, PS3_2, PT1, PT2, WR1, WR2, Controle2.
ykkC-III RF01763	AD3_2, AltoO2_2, BaixoO2_1, BaixoO2_2, Nar1, Nar2, Nitrato_1, Nitrato_2, Controle2.
yybP-ykoY RF00080	AD1_1, AD3_1, AD3_2, AltoO2_1, AltoO2_2, BaixoO2_2, Nar1, Nar2, Nitrato_2, PS1_1, PS1_2, PS3_1, PT1, PT2, WR1.