



Article

# Bacterial Small RNAs in the Genus *Herbaspirillum* spp.

Amanda Carvalho Garcia <sup>1,\*</sup> , Vera Lúcia Pereira dos Santos <sup>1</sup>,  
Teresa Cristina Santos Cavalcanti <sup>2</sup>, Luiz Martins Collaço <sup>2</sup> and Hans Graf <sup>1</sup>

<sup>1</sup> Department of Internal Medicine, Federal University of Paraná, Curitiba 80.060-240, Brazil; santosvlp@hotmail.com (V.L.P.d.S.); hansgraf@bighost.com.br (H.G.)

<sup>2</sup> Department of Pathology, Federal University of Paraná, PR, Curitiba 80.060-240, Brazil; tecava@yahoo.com.br (T.C.S.C.); lmcollaco@uol.com.br (L.M.C.)

\* Correspondence: amanda.ufpr@gmail.com; Tel.: +55-41-99833-0124

Received: 5 November 2018; Accepted: 12 December 2018; Published: 22 December 2018



**Abstract:** The genus *Herbaspirillum* includes several strains isolated from different grasses. The identification of non-coding RNAs (ncRNAs) in the genus *Herbaspirillum* is an important stage studying the interaction of these molecules and the way they modulate physiological responses of different mechanisms, through RNA–RNA interaction or RNA–protein interaction. This interaction with their target occurs through the perfect pairing of short sequences (*cis*-encoded ncRNAs) or by the partial pairing of short sequences (*trans*-encoded ncRNAs). However, the companion Hfq can stabilize interactions in the *trans*-acting class. In addition, there are Riboswitches, located at the 5' end of mRNA and less often at the 3' end, which respond to environmental signals, high temperatures, or small binder molecules. Recently, CRISPR (clustered regularly interspaced palindromic repeats), in prokaryotes, have been described that consist of serial repeats of base sequences (spacer DNA) resulting from a previous exposure to exogenous plasmids or bacteriophages. We identified 285 ncRNAs in *Herbaspirillum seropedicae* (*H. seropedicae*) SmR1, expressed in different experimental conditions of RNA-seq material, classified as *cis*-encoded ncRNAs or *trans*-encoded ncRNAs and detected RNA *riboswitch* domains and CRISPR sequences. The results provide a better understanding of the participation of this type of RNA in the regulation of the metabolism of bacteria of the genus *Herbaspirillum* spp.

**Keywords:** RNA non-coding (ncRNA); ncRNA *cis*-encoded; *trans*-encoded; riboswitches; CRISPR; mRNA

## 1. Introduction

The genus *Herbaspirillum* was described by Baldani et al. [1] and includes the *Herbaspirillum seropedicae* (*H. seropedicae*) species with several strains isolated from different grasses [1]. *Herbaspirillum seropedicae* is characterized by fixing nitrogen (diazotrophic) and colonizing plants.

A second diazotrophic species, *Pseudomonas rubrisubalbicans*, was reclassified under the name of *Herbaspirillum rubrisubalbicans* and also has the ability to colonize plants [2]. From the 1990s, new species were included in the genus *Herbaspirillum*: *H. autotrophicum* (*syn. Aquaspirillum autotrophicum*) and *H. huttienne* (*syn. Pseudomonas huttienne*) [3] *H. frisingense* [4], *H. lusitanum* [5], *H. cholorophenicum* [6], *H. hiltneri* [7], *H. rhizosphaerae* [8], *H. aquaticum* [9], *H. massiliense* [10], *H. canariense*, *H. aurantiacum*, and *H. soli* [11], and *H. seropedicae* AU14040 [12]. The *H. putei* species was reclassified as *H. huttienne* subspecies *putei* [9].

*H. seropedicae* belongs to the beta class of Proteobacteria, is gram-negative, and displays a characteristic growth pattern in semi-solid culture medium and without nitrogen [1,13]. It is a

bacterium that colonizes intracellular spaces and the xylem of grass roots, such as sugarcane and sorghum [1,14]; and intracellular spaces in seeds and rice roots [14,15]. It can also be isolated from the stem leaves of various grasses [15–17]. The genome of the *H. seropedicae* strain (SmR1) has been completely sequenced and annotated by the GENOPAR (Paraná Genome) consortium. “Wordiness 4,735” Open Reading frames (ORFs) occupy about 88.3% of the genome, on a single circular chromosome.

In this annotation, the genes of non-coding RNAs (apart from rRNAs and tRNAs), are not included. Gene encoding regulatory RNAs, known as small RNAs (sRNAs) or non-coding sRNAs (ncRNAs), are transcribed in *trans* and in *cis* relative to target RNA [18,19]. The work of the researchers Anderson et al. were the first to isolate and characterize the RNA encoded by the *micF* gene in *Escherichia coli* [20].

The genes are located between the coding regions for the proteins, that is, in the intergenic regions of the genome and display signals from the promoters and terminator sequences that are usually Rho-independent [21–23]. The size of the ncRNAs genes ranges from ~50 to ~500 nucleotides and several transcripts were processed by RNase into smaller products [24–27]. They modulate physiological responses through different mechanisms, by RNA-RNA interaction or RNA-protein interactions, and some interactions can be regulated by the Hfq chaperone [28–30].

The most studied ncRNAs are the *cis*-encoded RNAs and the *trans*-encoded RNAs. The *cis*-encoded RNAs are transcribed in *cis* and *encoded* at the same locus as their target mRNA in the antisense sense duplex structure, in addition to having a post-transcriptional gene regulation of responses and a high degree of sequence complementarity, which was considered to be an indication that interaction of the Hfq protein [31,32] would not be required. However, some investigators reported interference of Hfq in the ncRNA, mRNA targeting, and *cis*-encoded target [33–35]. Overall, these ncRNAs act by complementing the mRNA ribosome-binding site, as well as inhibiting the conversion [33,36].

In contrast, *trans*-encoded ncRNAs that are encoded in *trans*, have target mRNAs at different sites in the genome where the formation of the ncRNA occurs: the Hfq—mRNA complex may act positively or negatively on post-transcriptional regulation [37]. This is due to the imperfect base pairing which prevents their eventual degradation by RNase [38,39]. The regulation of gene expression of *trans*-encoded ncRNAs was first discovered by Dalihis, who described the *micF* and *dicF* found in *Escherichia coli*, and the *lin-4* found in the nematode *Caenorhabditis elegans*, which are encoded by genes located at different sites of their target genes and only display a partial complementarity with their target RNAs [40].

The *Riboswitches*, which are located in the 5' UTR region of an mRNA, constitute another class of ncRNAs and lead to transcriptional regulation through their interaction with a linker molecule [41–43]. However, the presence of small, non-coding RNAs in different *Herbaspirillum* species remains undetermined. A total number of 285 ncRNAs are needed to predict the novel candidate sRNAs within the intergenic regions (*Herbaspirillum seropedicae* SmR1). We further analyzed the predicted ncRNAs in great detail in *Herbaspirillum* spp. using Infernal 1.1.1 bioinformatic tools. Moreover, these analyses constitute the very first documented evidence of the presence of sRNAs in *Herbaspirillum* genomes.

## 2. Results

### 2.1. Prediction of New ncRNAs in the Genus *Herbaspirillum* spp.

The computational tool Infernal 1.1.1 was employed for this study (see Section 4.3) and the multiple covariance detection and sequence analysis had as its input the RNA family database (Rfam, available link <http://rfam.xfam.org/>; accessed on 18 February 2016), while the prediction of the 16 species of ncRNAs made up the genus *Herbaspirillum* spp., as shown in Supplementary Material S1.

The complete list of genomes is as follows: *H. seropedicae* Z67, *H. lusitanum* P6-12, *H. hiltneri* N3, *H. frisingense* GSF30, *Herbaspirillum* spp. strain B65, *H. rubrisubalbicans* M1, *H. autotrophicum*

IAM 14942, *H. huttiense* subsp. *putei*, *Herbaspirillum* spp. strain B501, *Herbaspirillum* spp. strain GW103, *Herbaspirillum* spp. strain RV423, *H. rhizosphaerae* UMS-37, *Herbaspirillum* spp. strain Os34, *Herbaspirillum* spp. strain Os45, and *H. seropedicae* AU14040. The results are listed in Table 1. By adopting this approach, no ncRNAs were detected in *H. lusitanum* P6-12 and *Herbaspirillum* spp. strain B65, due to the significant number of scaffolds and contigs that are identified in incomplete genomes. The organisms with the highest number of common ncRNAs were *H. seropedicae* SmR1, *H. seropedicae* Z67, *H. hiltneri* N3, *H. frisingense*, *H. rubrisubalbicans* M1, and *H. seropedicae* AU14040, shown in Supplementary Material S4.

**Table 1.** ncRNAs predicted for bacteria of the genus *Herbaspirillum* and different from those already known in *H. seropedicae* SmR1. The Infernal (version 1.1.1) computational tool was used.

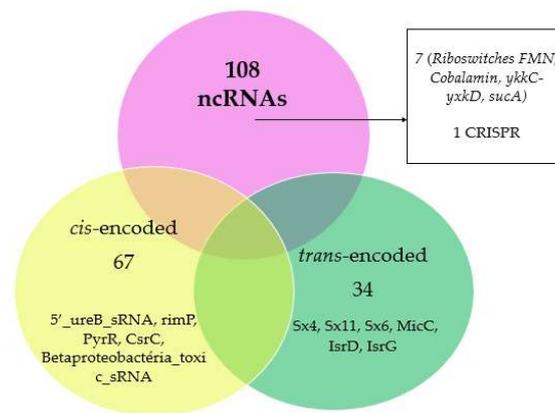
Genus <i>Herbaspirillum</i>	ncRNAs
<i>Herbaspirillum seropedicae</i> Z67	ohsC_RNA, Rhizobiales-2 and UPD-PKc
<i>Herbaspirillum lusitanum</i> P6-12	Absence
<i>Herbaspirillum hiltneri</i> N3	ar14, Acido-1, Alpha_RBS, ar14, asX2, asX3, Bp2_287, csRNA, isrN, istR, Ms_AS-4, Ms_IGR-8, ncr1175, NsiR1, P26, P5, PrrB_RsmZ, RsmY, RydC, sau-5971, SpF41_sRNA, Spot_42, Xoo5, sraA, sX9, TB10Cs4H2, TB11Cs5H2 and TB9Cs1H1
<i>Herbaspirillum frisingense</i> GSF30	Entero_5_CRE, P10 and STnc430
<i>Herbaspirillum</i> spp. <i>estirpe</i> 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 and TB9Cs1H2.
<i>Herbaspirillum huttiense</i> subsp. <i>putei</i> <i>estirpe</i> 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 and veev_FSE.
<i>Herbaspirillum autotrophicum</i> IAM 14942	MIR1444 and SpF10_sRNA.
<i>Herbaspirillum</i> spp. <i>estirpe</i> B501	Absence
<i>Herbaspirillum</i> spp. <i>estirpe</i> GW103	sX2, Flavi_CRE, RyeB, SAM-IV, sau-6072 e sroB.
<i>Herbaspirillum</i> spp. <i>estirpe</i> RV1423	psRNA2, sbcD and STnc40.
<i>Herbaspirillum rhizosphaerae</i> UMS-37	NRF2_IRES, NsiR1, Pseudomon-groES and sX9.
<i>Herbaspirillum rubrisubalbicans</i> spp. <i>estirpe</i> Os34	ryfA, SpR20_sRNA and asX2.
<i>Herbaspirillum rubrisubalbicans</i> spp. <i>estirpe</i> Os45	asX2, ROSE_2 and ryfA.

Genus *Herbaspirillum* = 15 species; ncRNAs = non-coding RNAs.

The ctRNA\_p42d, Betaproteobacteria\_toxic\_sRNA, and P10 ncRNAs are examples of ncRNAs found in *Herbaspirillum* spp. RV423, *H. rhizosphaerae* UMS-37, and *Herbaspirillum* spp. OS34 and O45 and common to *H. seropedicae* SmR1. The ncRNAs *Betaproteobacteria\_toxic\_sRNA*, *pfl*, *SX4*, and *alpha\_tmRNA*, AR35, *tmRNA*, among others, are repeated in all the species, except for *Herbaspirillum* spp. B65, *Herbaspirillum* spp. B501, and *Herbaspirillum lusitanum* P6-12. The characterization of the ncRNAs predicted in the genus *Herbaspirillum* spp. suggests that despite the differences within these species, there is sequence preservation among the common ncRNAs.

In *Herbaspirillum seropedicae* SmR1, the prediction of ncRNAs resulted in 108 ncRNAs, as shown in Figure 1. The sequence of ncRNAs predicted in *H. seropedicae* SmR1 by Infernal 1.1.1 is shown in Supplementary Material S2.

All ncRNAs predicted by the Infernal 1.1.1 tool in the genus *Herbaspirillum* spp. are shown in Supplementary Material S4.

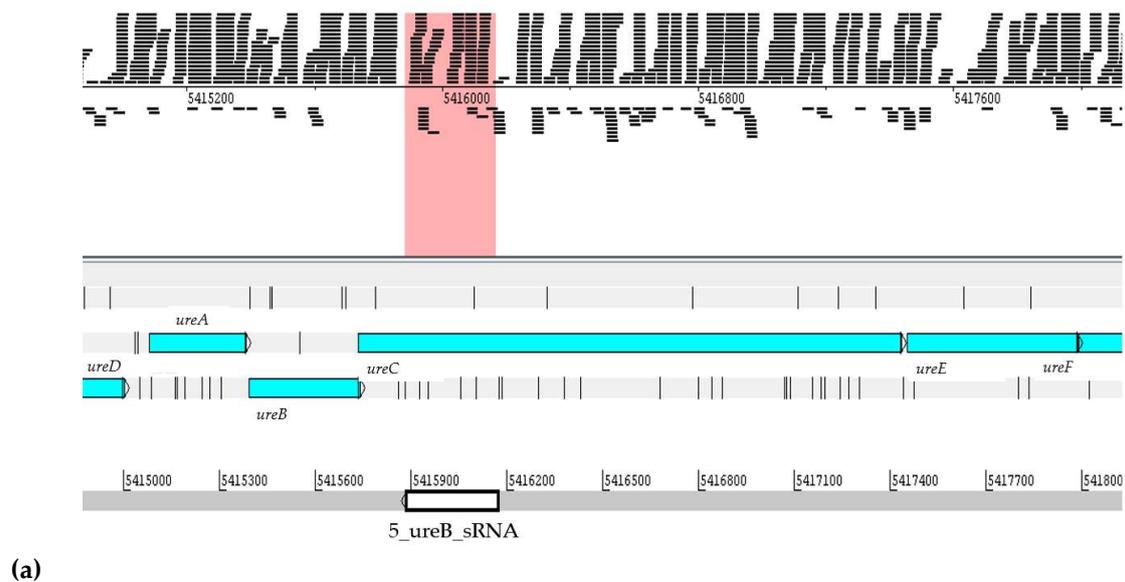


**Figure 1.** Examples of the *cis*-encoded 5'\_ureB\_sRNA, rimP, PyrR, CsrC, Betaproteobacteria\_toxic\_sRNA, riboswitch, FMN, cobalamin, ykkC-yxkD, sucA; and the *trans*-encoded sX4, sX11, sX6, MicC, IsrK, IsrD, IsrG. The CRISPR RNA region predicted the *H. seropedicae* in the SmR1 genome.

2.2. ncRNAs Differential Expression in RNA-Seq

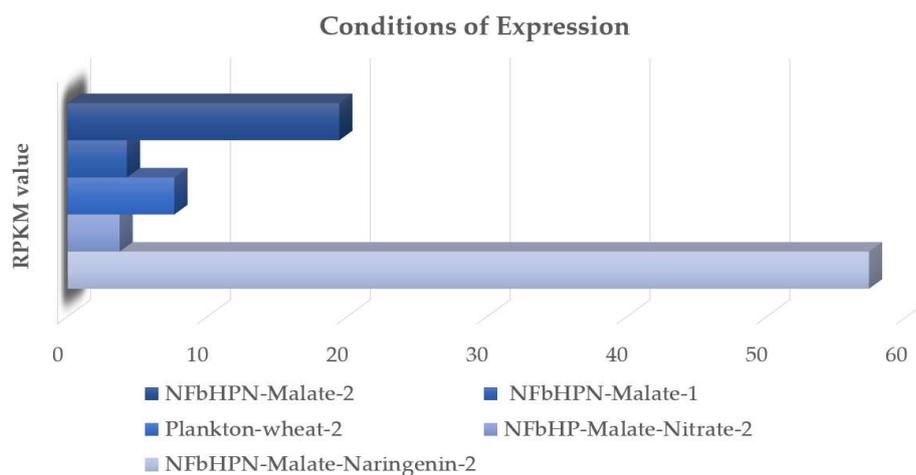
Data from RNA-seq data were used in 11 culture conditions that were tested in *Herbaspirillum seropedicae* SmR1, as described in Supplementary Material S3. Less than five covering values were included and these were regarded as possible predicted ncRNAs. It was observed that there were ncRNAs expressed in all 11 conditions, but others were only expressed in some of the conditions.

The 5'\_ureB\_sRNA in *H. seropedicae* SmR1 has 286 nucleotides and is classified as *cis*-encoded antisense ncRNA depending on its function and localization of the genome, is incorporated with the *ureA*, *ureB*, *ureC*, *ureD*, *ureE*, and *ureF* (sense) genes, is expression the NFbHPN-Malate-Naringenin and NFbHP-Malate-2 conditions, as shown in Figure 2. *H. seropedicae* SmR1 uses naringenin as a carbon source and is found in the plant, where the biosynthesis of flavonoids occurs; its compounds are involved in the plant defense mechanism.



(a)

Figure 2. Cont.

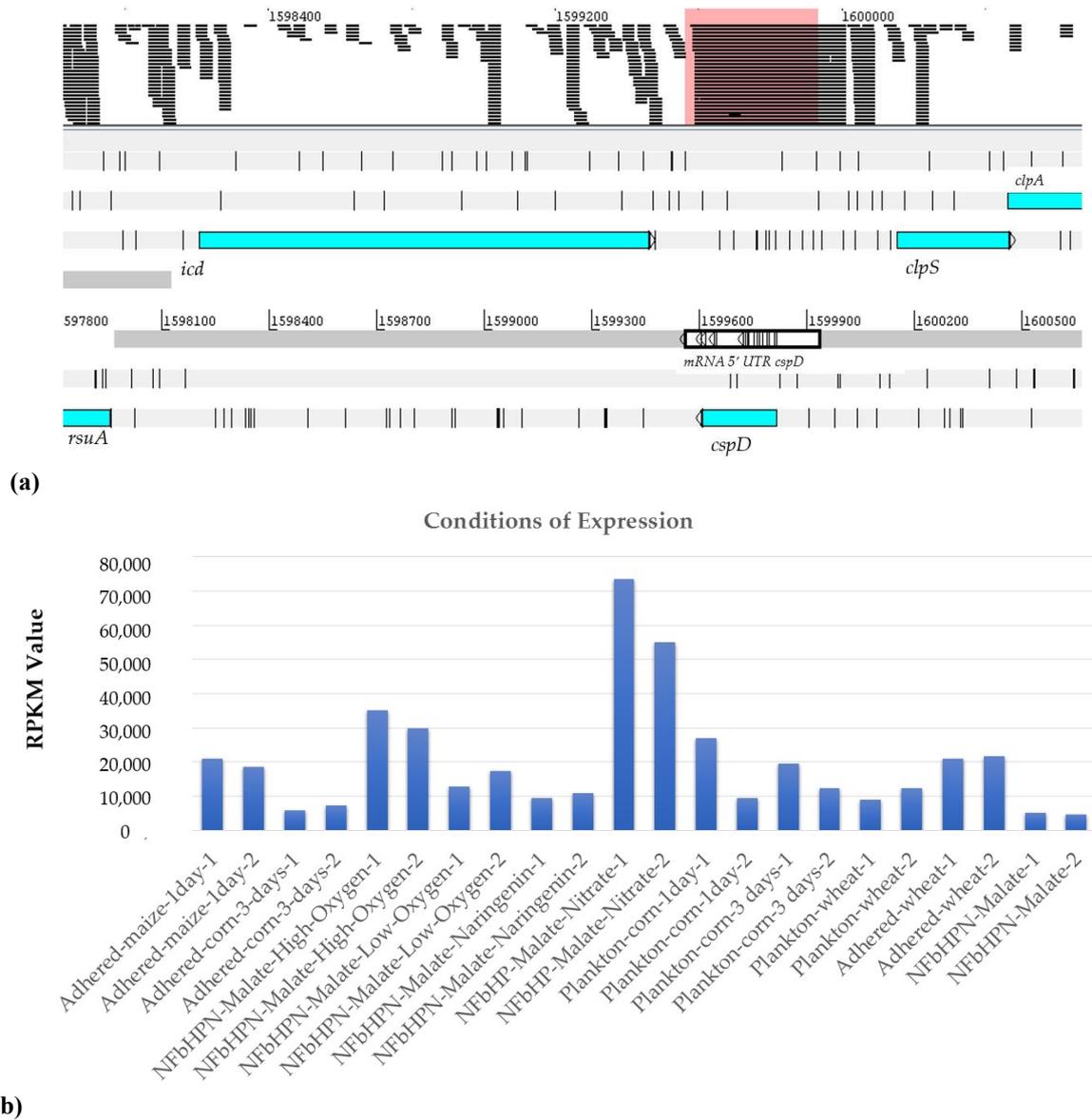


(b)

**Figure 2.** Visualization of position and expression profile of 5\_ureB\_sRNA ncRNA in the genome of *H. seropedicae* SmR1. The sequence was located using the Artemis program. The readings delineated within the pink rectangle refer to: (a) Levels of expression of 5\_ureB\_sRNA ncRNA in RNA-seq assays in *H. seropedicae* SmR1. In color, the analyzed experimental conditions represented are shown, where 1 and 2 represent d.a.i (days after inoculation); (b) the numbers represent the value of the gene expression in reads per kilobase of transcript per million mapped reads (RPKM).

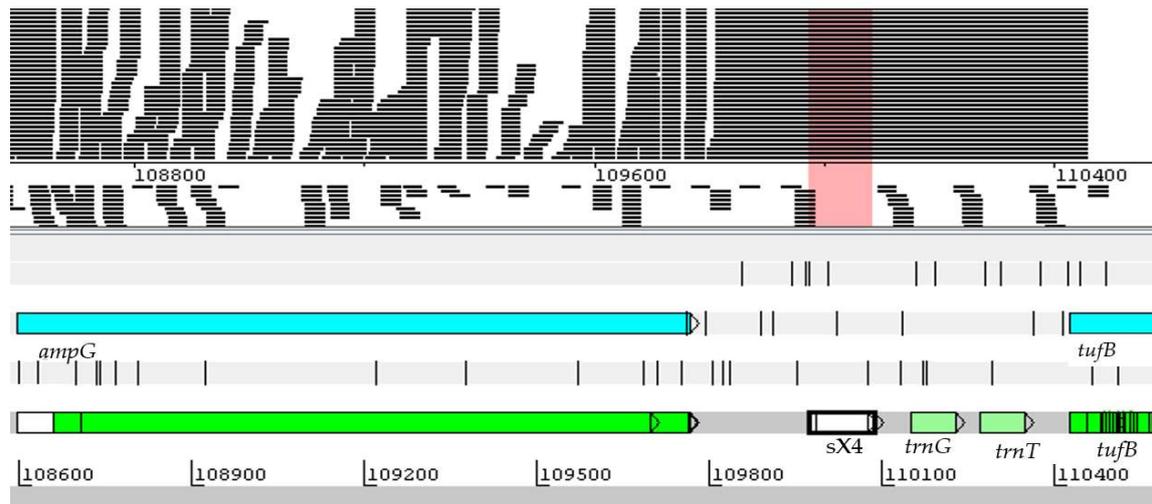
The ncRNA 5'\_ureB\_sRNA is repeated in some genera of *Herbaspirillum* by having 100% sequence identity such as, for example, in *H. seropedicae* Z67, *H. lusitanum*, *H. hiltneri* N3, *H. rubrisubalbicans* M1, and *H. seropedicae* AU14040. The location of the ureABC operon locus in the genome of these species was identified. The presence of ureaseC (alpha), ureaseA (gamma), and ureaseB (beta) subunits was detected and it was found that the 5'-ureaB sRNA ncRNAs were found in these species of *Herbaspirillum* interacts in *cis* of the ureaseC subunit. The ncRNA mRNA 5' UTR cspA is a thermoregulator which has 373 nucleotides in *H. seropedicae* SmR1. It is classified as a *cis*-encoded ncRNA, which appears as a mRNA 5' mRNA sequence of the gene annotated as *cspD* (*Hsero\_1397*). In addition, there is a second *cspD* gene (*Hsero\_3028*) that does not contain the mRNA sequence 5' UTR.

Yamanaka and Inouye [44] used the *cspD-lacZ* fusion and found that the *cspD* expression induced by stationary phase growth does not depend on sigma factor 6S. Thus, the *cspD* gene expression is inversely dependent on the growth, rates induced by glucose starvation. They also observed that the 5' UTR cspA mRNA is expressed in the all assay conditions and showed a higher expression in seq RNA assays at NFbHPN-Malate-Nitrate and NFbHP-Malate-Nitrate-2 and NFbHPN-Malate-High-Oxygen-1, as shown in Figure 3. NFbHPN-Malate-Nitrate-1 is where *H. seropedicae* SmR1 rises to 30 °C in NFbHP medium, using malate as a carbon source and 10 mM nitrate as a nitrogen source [45].

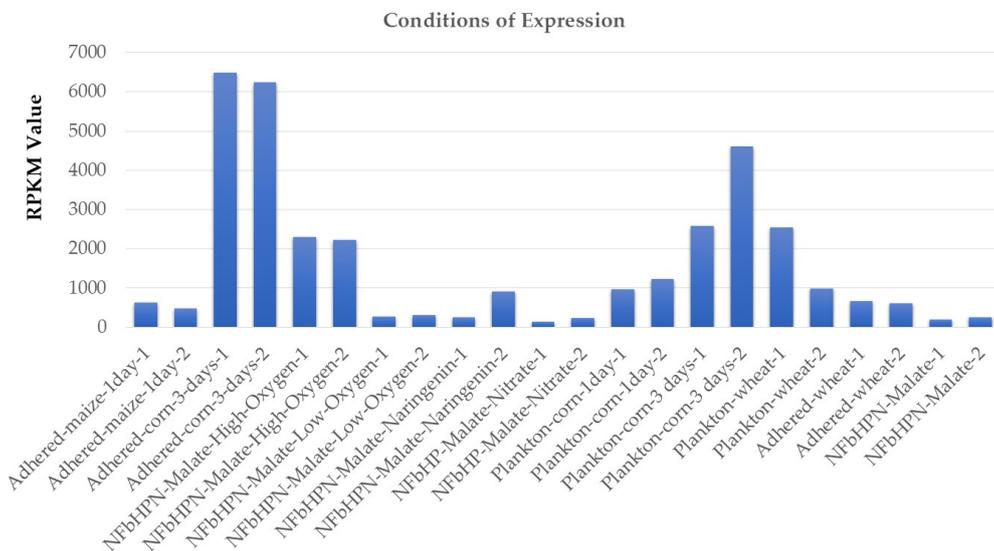


**Figure 3.** Visualization of the position and expression profile of mRNA 5' UTR *cspA* in the genome of *H. seropedicae* SmR1. The sequence was located with the aid of the Artemis program. The readings delineated within the pink rectangle refer to: (a) Levels of expression of mRNA 5' UTR *cspA* in RNA-seq assays in *H. seropedicae* SmR1. In the colored part, the 11 analyzed experimental conditions are represented, where 1 and 2 represent d.a.i (days after inoculation); (b) the numbers represent the value of the gene expression in RPKM.

The ncRNA *sX4* has 113 nucleotides and is classified as *trans*-encoded antisense ncRNA. It is co-located with the *ampG*, *tufB*, and *trnG* genes. The designation *sX4* refers to the ncRNAs described in *Xanthomonas campestris* *pv. vesicatoria* [46]. In *H. seropedicae* SmR1, the *sX4* ncRNA was expressed in all the tested conditions, but the highest expression in Adhered-corn-3-days conditions, followed by the Plankton-corn-3-days-2. It was found that the greatest expression occurred during the colonization of maize, as shown in Figure 4. Thus, it should be noted that the greatest expression occurred during the colonization of corn.



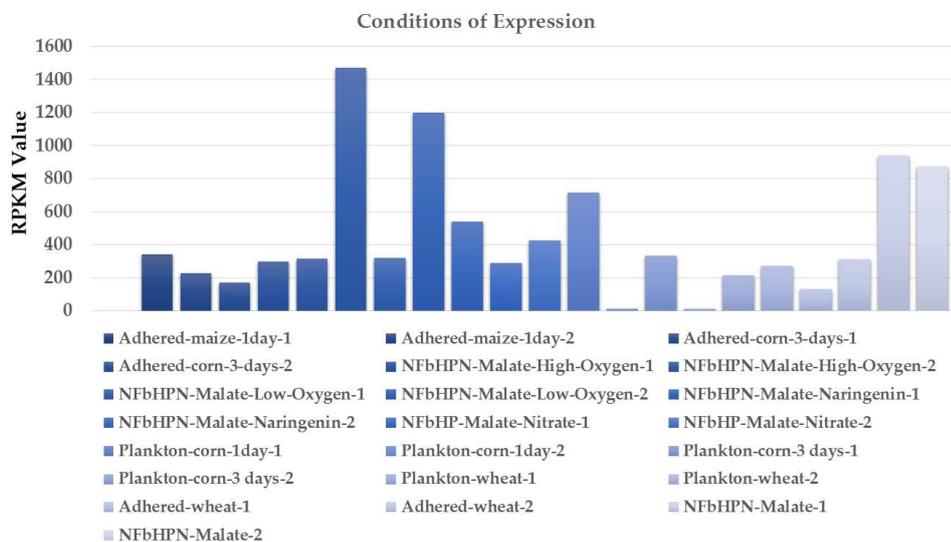
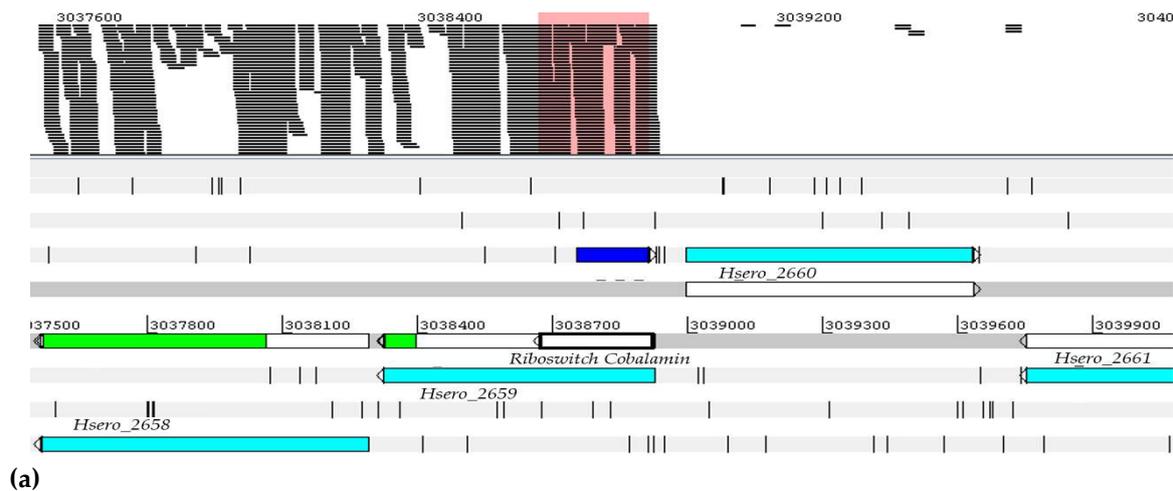
(a)



(b)

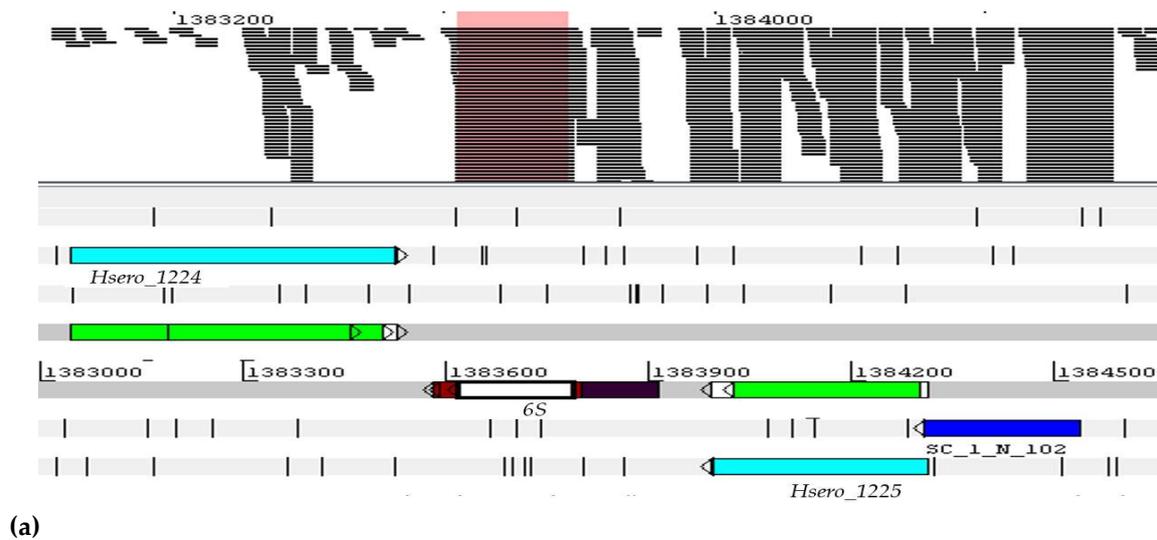
**Figure 4.** Visualization of the position and expression profile of the sX4 ncRNA in the genome of *H. seropedicae* SmR1. The sequence was located by means the Artemis program. The readings delineated by the pink rectangle are as follows: (a) Levels of sX4 ncRNA expression in RNA-seq assays in *H. seropedicae* SmR1. In the colored section, the 11 analyzed experimental conditions are represented, where 1 and 2 represent d.a.i (days after inoculation); (b) the numbers represent the value of the gene expression in RPKM.

In *H. seropedicae* the *cobalamin riboswitch* has 113 nucleotides and is classified as *trans*-encoded antisense ncRNA or cobalamin element. It is co-located with the *Hsero\_2659* gene (cobalt transporter) as a sequence 5' UTR, and with the genes *Hsero\_2660* and *Hsero\_2661* (antisense sense) genes. Coenzyme B12 or cobalamin riboswitches are control elements that are widespread in prokaryotes. For example, the *metE* gene contains a cobalamin *riboswitch* and binding of coenzyme B12 to this element that leads to the formation of a transcriptional terminator and the repression of *metE* expression. In the absence of the coenzyme, the *metE* mRNA is synthesized [47]. NFbHPN-Malate-Low-Oxygen-1 and NFbHPN-Malate-Naringenin-1 are shown in Figure 5. These results indicate that there is likely to be no binding of the CbtB (cobalt carrier subunit) linker to the *hsero\_2659* mRNA in the aptamer region, where an alternate clamp loader formed.

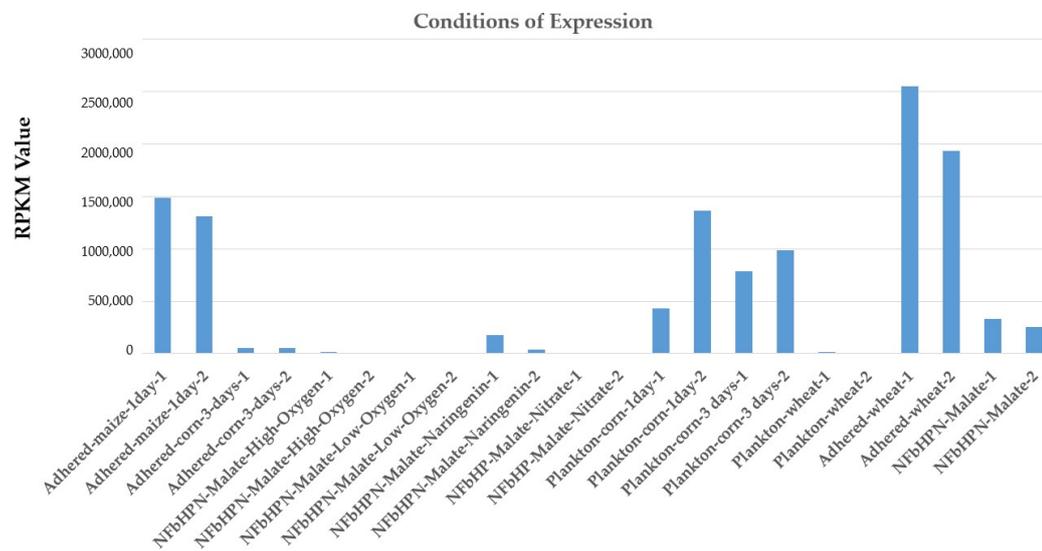


**Figure 5.** Visualization of the position and expression profile of the cobalamin ncRNA in the *H. seropedicae* genome SmR1. The sequence was located by means of the Artemis program. The readings are delineated by the pink rectangle (a). Levels of ncRNA cobalamin expression in RNA-seq assays in *H. seropedicae* SmR1. In the colored section, the 11 analyzed experimental conditions represented, where 1 and 2 represent d.a.i (days after inoculation). (b) The numbers represent the value of the gene expression in RPKM.

Non-coding RNAs were regarded as essential (*housekeeping* ncRNAs), such as 6S ncRNA, and were also identified by the Infernal 1.1.1 tool. In *Herbaspirillum seropedicae* SmR1, mcRNA 6S has a sequence of 177 nucleotides and expression in RNA-seq material. The ncRNA 6S in *E. coli* is responsible for regulating the activity of the RNA polymerase (RNAP) that is dependent on the sigma factor 70, and necessary for the recognition of the promoter and initiation of the transcription. In addition, it globally regulates the gene expression response to growth from the exponential phase to the stationary phase [48]. It is significant that the level of expression was found to be high relative to its neighbors (*Hsero\_1224* and *Hsero\_1225*) under the conditions tested, and this confirms that this ncRNA is essential. In addition, the 6S ncRNA was expressed in all the conditions tested, with an emphasis on the Adhered-wheat-1 condition, as shown in Figure 6.



(a)

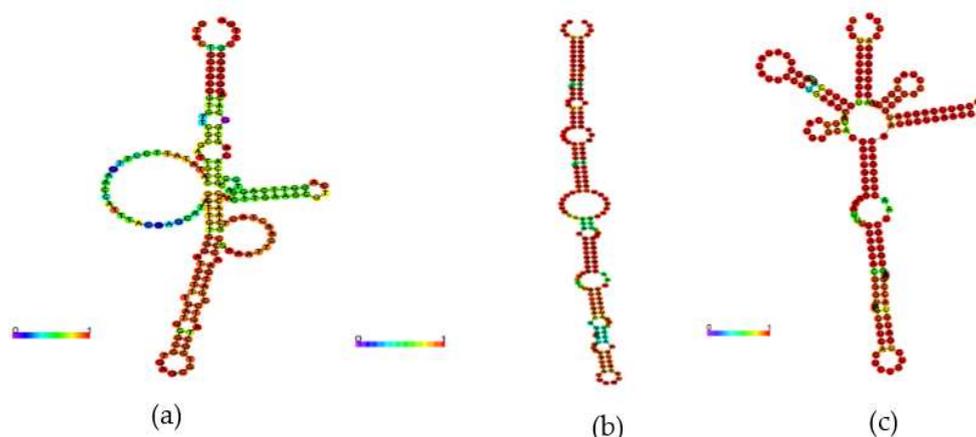


(b)

**Figure 6.** The sequence was located by means of the Artemis program. The readings delineated by the pink rectangle were as follows: (a) Visualization of expression levels of ncRNA 6S in RNA-seq assays in *H. seropedicae* SmR1. In the colored section, the 11 analyzed experimental conditions represented, where 1 and 2 show d.a.i (days after inoculation); (b) the numbers represent the value of the gene expression in RPKM.

### 2.3. Comparative Analysis of the Secondary Structure of 6S of *H. seropedicae* SmR1

Given the importance of this ncRNA, there was a comparative analysis of the secondary structure of the 6S of *H. seropedicae* SmR1, together with the 6S of *Janthinobacterium* spp. and *Burkholderia* CCGE1003. The analysis was conducted with the RNAFold [49] tool, and based on thermodynamic RNA-RNA interactions, where the secondary structures of ncRNA 6S were generated. The three structures shown in Figure 7 are basically similar (a long rod shown in Figure 7B for the 6S of *Janthinobacterium* spp.). However, there are variations in the presence and number of loop out structures that may or may not have a functional meaning. A multiple sequence alignment was carried out by the CLUSTALW tool [50], as shown in Figure 8. It is observed that there is a difference in the secondary structure of ncRNA 6S between these species, even if they present a high degree of conservation in the aligned nucleotide sequence.



**Figure 7.** Visualization of the secondary structure of ncRNA 6S. (a) *H. seropedicae* SmR1; (b) *Janthinobacterium* sp.; (c) *Burkholderia* CCGE1003. A predictive analysis of the secondary structures of identified ncRNAs was conducted by means of RNAFold.

```

Hsero          GUACUCCUGCCGUGUUCGCGAUUGCCAUAUAUUCUUGAACCAUU--UACGAGCAUAGGU
Janthinobacteri GCGCUCCUGCCGUGUUCGCCAUUGCCAUAUUAUUCUUGAACCAUU--UAUGCGCAGCGGC
Burkholderia   GCUUCCUGCCGUGUUCGCCAAGGUCUAUAUUCUUGAACCAUUGCCAUGUGCA-CGGU
                *      * * * * *      * * * * *      * * * * *      * * * * *

Hsero          UGUGGGA--UCGUUUGAUGGGCGUGAGCGUCGCU-AGUCCGAUGAACCCGAAUUGAACU
Janthinobacteri UGCGGAAUUUUGUUCGACGGGAGUGCGCGUCGCU-AGUCCGAUGAACCCGAAGUUGAACU
Burkholderia   UGCGGAAUUUUGUAGCACGGGUGUGCGGUCACUCUGUCUGAUGUACCCGAAGUGCGUCU
                ** * * . * * * *      * * * * *      * * * * *      * * * * *

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

Hsero          UGA
Janthinobacteri UG-
Burkholderia   CAU
    
```

**Figure 8.** Multiple alignment of 6S ncRNA gene sequences by the CLUSTALW tool. Hsero = *Herbaspirillum seropedicae* SmR1; Janthinobacteri = *Janthinobacterium* sp.; and Burkholderia = *Burkholderia* CCGE1003.

### 2.4. Candidate sRNA Target Identification

In *H. seropedicae* SmR1, the consensus ribosomal binding site (RBS) sequence and its variations were examined, since this is the RBS AGGA nucleus and the CAAGGACA sequence. Take as an example the nucleotide sequence of ncHSmR1\_04, which has 40 possible targets. They include the *serS* gene that encodes a seryl-tRNA synthetase.

It is worth noting that the base pairing between the ncHSmR1\_04 and the mRNA *serS*, occurs in the RBS region of the latter, as shown in Figure 9. This is a strong indication that the ncHSmR1\_04 negatively regulate the conversion of seryl-tRNA synthetase. The 4.5S RNA is a common ncRNA in bacteria and its purpose is to the function of directing proteins containing signal peptide to the secretory apparatus, and thus form a signal-recognizing particle with the Fhh protein. In addition, the 4.5S RNA takes part in the conversation when interacting with the G stretch factor of G [51].

```

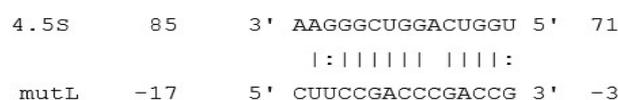
ncHSmR1_04      36      3' AGCAAGAAAGAAGGGCUGCG 5'      17
                |||| :| | | | | | | | | | |
serS            -15      5' UCGUAUUUUUCUCCCC-AUGA 3'      4
    
```

**Figure 9.** It is worth noting that the base pairing between the ncHSmR1\_04 and the mRNA *serS* occurs in the RBS region of the latter.

In *E. coli*, the 4.5S RNA is encoded by the gene (*ffs*) and contains 114 nucleotides. It is the bifunctional molecule that is involved in the conversion and secretion of protein by binding to elongation factor (EFG) and Fth the protein respectively [52]. Apparently, the 4.5S RNA facilitates the release of the EF-G-GTP from the ribosome by competing with the 23S rRNA for binding with the EF-G [51,53].

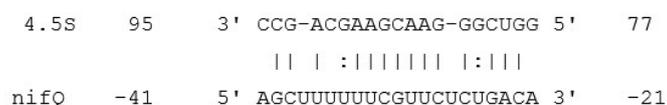
Buskiewicz et al. [54] analyzed the functions of the SRP (signal-recognizing particle) in *E. coli* that were based on the binding of the Fth protein to the 4.5 S RNA and the Fth-4.5 S complex. The researchers suggested that the free Fth, the 4.5S RNA, and FtsY binding sites are occluded by strong domain–domain interactions that must be disrupted by the SRP or Fth-FtsY complex formation.

The 4.5S RNA of *H. seropedicae* SmR1 contains 146 nucleotides. Among the targets of 4.5S, as predicted by Target RNA2, is the *mutL* mRNA that encodes a protein involved in DNA repair. Again, the matching of the ncRNA occurs close to that of the RBS of the target RNA and in this specific case, starts at the third base upstream of the adenosine of the AUG codon, as shown in Figure 10.



**Figure 10.** It is worth noting that the base pairing between the 4.5S and the mRNA *mutL* start at the third base upstream of the adenosine of the AUG codon.

The analysis of the ncRNA 4.5S conducted by the TargetRNA2 tool resulted in 77 target mRNAs in this set and includes the *nifQ* gene whose product is involved in nitrogen metabolism, as shown in Figure 11.



**Figure 11.** It is worth noting that the base pairing between the 4.5S and the mRNA *nifQ* predicted pairing starts 21 bases upstream of the AUG codon, and hence is outside of the RBS region.

It can be seen that the predicted pairing starts 21 bases upstream of the AUG codon, and hence is outside of the RBS region, as shown in Figure 12. This means that according to TargetRNA2, when compared with the previous example, an interaction energy value is obtained for the pairing:  $-16.21$  for *mutL* and  $-8.31$  for *nifQ*. These two factors (the pairing position and interaction energy value) make *nifQ* an unlikely target for 4.5S.

```

5'GTCCTTCTTCAGGCAGGACGACATGAAGGCCTTGC GTTCGTCCCTTCTTGCC
   CGCGGCGTCAGCATTGCACTGCTTCACTTGCTGCTGGGTGGCTGGCTGGGCGC
   GTCCTTCTTCAGGCAGGACGACATGAAGGCCTTGC GTTCGTCCCTTCTTGCC3'

```

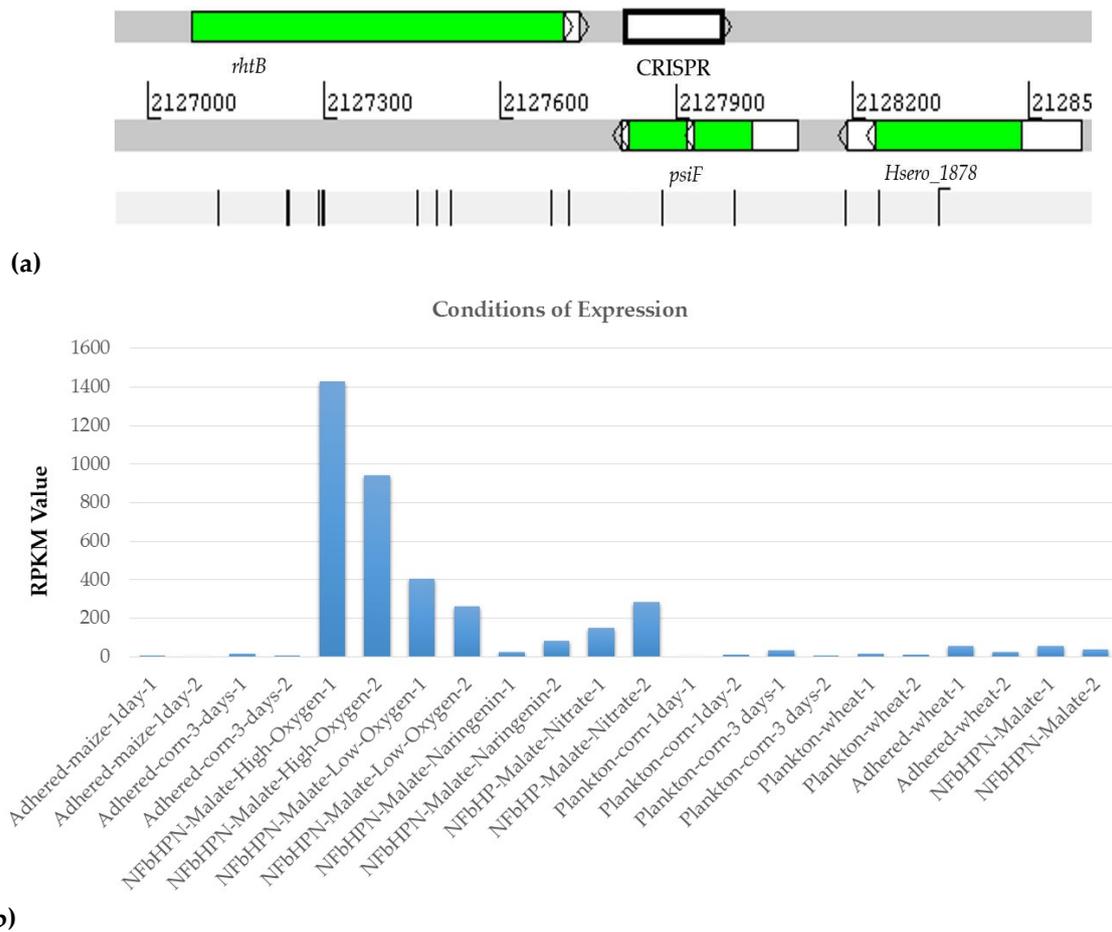
**Figure 12.** Visualization of the CRISPR locus in the genome of *H. seropedicae* SmR1. Sense sense in the genome, with 164 nucleotides divided into DR (54 nucleotides) palindromic repetitions and spacing region (57 nucleotides) underlined.

## 2.5. Prediction of CRISPR

A new category of non-coding RNA, CRISPR RNA, has been described for both the Bacteria domain and the Archaea domain [55,56]. The CRISPR transcription factor results in small fragments of RNA that recognize a specific exogenous DNA and that can guide the Cas nuclease, that is responsible for the cleavage of this DNA if it re-contacts the microorganism. Thus, prokaryotes have a defense mechanism against invasive DNAs [57]. In this study, the CRISPRFinder and CRISPRmap tools were used to predict CRISPR in the genome of *H. seropedicae* SmR1. A CRISPR locus was identified with

164 nucleotides divided into DR (54 nucleotides) and SPA (57 nucleotides) palindromic repetitions, as shown in Figure 12.

The CRISPR locus is co-located with the *rhtB* gene (sense-sense) and with the *psiF* and *Hsero\_1878* genes (in the antisense-sense), as shown in Figure 13. With regard to this, CRISPR is expressed in some conditions, with the highest level being expression in the condition NFbHPN-Malate-Alto-Oxygen-1 a condition where *H. seropedicae* SmR1 is using malate as a carbon source up to OD of 0.4. This result suggests that the expression of the CRISPR locus in the NFbHPN-Malate-High-Oxygen-1 condition acts as a signal for the nitrogen source that the bacterium is using, by assisting in controlling the pH of the medium, since the *psiF* gene is expressed during the phosphate deprivation.



**Figure 13.** Visualization of the position of CRISPR in the genome of *H. seropedicae* SmR1 (a). Visualization of expression levels of CRISPR in RNA-seq assays in *H. seropedicae* SmR1. In the colored section, the 11 analyzed experimental conditions are represented, where 1 and 2 represent d.a.i (days after inoculation); (b) the numbers represent the value of the gene expression in RPKM.

In addition, the CRISPRfinder tool identified the CRISPR-associated protein, while the Cmr4 Family RAMP in *H. seropedicae* SmR1 interacted with casRAMP\_Cmr4, which participates in the process of developing immunological e memory. The Cas RAMP complex modulates between the cleavage of the RNAs [58–61]. The results suggest that the CRISPR-Cas system is a prokaryotic defense mechanism against foreign genetic elements. The main features of this defense system are Cas proteins and CRISPR RNA. However, since the different nitrogen sources have variations in the amount of nitrogen and CaCO<sub>3</sub> equivalents, acidification and the reduction of soil pH can become significant.

### 3. Discussion

In the RNA-seq assays, the highest expression of 5\_ureB\_sRNA ncRNA in the NFbHPN-Malate-Naringenin condition was observed, which may suggest that this 5\_ureB\_sRNA ncRNA is involved in the regulation of flavonoid biosynthesis. These act as a protection against oxidizing agents in the plant and involve the use of externally generated and internally generated urea as a source of carbon. Urea is metabolized by urease, a nickel-dependent metalloenzyme, which catalyzes the hydrolysis of urea to form ammonia and carbon dioxide, and can be found in bacteria, fungi, and plants. [62].

In *Klebsiella aerogenes*, the ureABC operon encodes the three subunits of the pimple enzyme [62]. In *Helicobacter pylori*, the urease cluster consists of two operons (ureAB and ureIEFGH). This bacterium is the only one described to date that contains the 5\_ureB\_sRNA, an antisense *cis*-encoded ncRNA with 290 nucleotides, which negatively regulates ureAB expression when paired at the 5'-region of the ureB mRNA [63]. *H. seropedicae* SmR1 has a probable ureABC operon and a 5'\_ureB ncRNA, with 286 nucleotides. According to the prediction of Infernal 1.1.1, this ncRNA is a *cis*-encoded antisense that is positioned close to the ureC gene. This difference can be explained by the fact that (a) *H. pylori* does not have the ureC gene, (b) the *H. seropedicae* SmR1 gene sequences are larger than those of *H. pylori*, and (c) the metabolism of urea involves different proteins.

According to Yamanaka and Inouye [44], in *E. coli* there are new homologues of cspA, denominated cspA to cspI, and the cspD gene does not have the 5' UTR sequence. This information suggests that in *H. seropedicae* SmR1, the genes cspD (*Hsero\_1397*) is actually the cspA gene. This gene is co-located with the genes *rsuA*, *icd*, *clpS*, *clpA* (sense sense). Jiang et al. [43] analyzed the function of the unconverted 5' UTR sequence CspA mRNA from the major cold shock adaptation protein to CspA in *E. coli*. The researchers observed that adaptation to cold shock is blocked when the 5' UTR CspA mRNA region is overexpressed, through the synthesis not only of the CspA protein, but also of proteins from the CspA family such as CspB, CspC, CspD, and CspE. The cspD gene in *E. coli* codes for a protein of homologous sequence with the cold shock protein to the CspA, but the expression of cspD is not induced by the cold shock [44].

The sX4 ncRNA whose expression is dependent on the HrpG and HrpX proteins, suggests that the target (s) of this ncRNA can be found in the interaction of the bacterium with the plant [46,64]. Studies with *HrpG* and *HrpX* mutants in *X. campestris* pv. *vesicatoria* reveal that these genes are essential for pathogenicity and assist in the survival of epiphytic bacteria [65]. These microorganisms are gram-negative and pathogenic  $\gamma$ -proteobacteria of many plant species of economic interest.

The ncRNA MicC has 97 nucleotides, is classified as *trans*-encoded antisense, and is co-located with the *murI* and *smoked* genes and with the *gst* gene (sense). In *E. coli*, the MicC ncRNA is located between the *ompN* and *ydbk* genes, and regulates OmpC protein expression by pairing with the mRNA leader sequence and inhibiting ompC mRNA binding to the ribosome [47,65].

In the genome of *H. seropedicae* SmR1, the *ompC* gene is located at the *Hsero\_1282* locus, the *ompA* gene is located at the *Hsero\_1287* locus and the *ompR* gene is located at the *Hsero\_1518* locus. With the aid of the TargetRNA2 tool, it can be observed that the MicC ncRNA in *H. seropedicae* SmR1 targets the *rscB* gene located at the *Hsero\_1538* locus, which acts as a transcriptional regulator that has a receptor domain for the OmpR protein. The ncRNA MicC was expressed in all the tested conditions and achieved higher expression observed in Adhered-corn-3-days-2 and Aderido-maize-3-days-1 conditions, with a higher level of expression with regard to the other conditions. In addition, it is involved in the transcriptional expression of the *rscB* gene when it is adhered to maize.

In *H. seropedicae* SmR1, the *IsrD* ncRNA has 62 nucleotides, and is co-located with the *mmr* and *Hsero\_0991* genes, together with the *Hsero\_0992* gene (antisense sense). It has the highest expression profile in RNA-seq material under Planktonic-corn-3 days-2 and NFbHPN Malate-High-Oxygen-2 conditions. The *IsrG* ncRNA has 88 nucleotides, is co-located with the *fimV* and *asd* genes, and with *Hs\_noco\_697* (sense sense), and is only expressed in the NFbHPN Malate-High Oxygen condition. In *Salmonella typhimurium* ncRNA, *IsrG* is involved in the response to cold shock and acid pH, in

the stationary phase of growth [66]. According to the literature, genes from the ncRNA-*isr* family have a promoter sequence superimposed on the 5' end or the 3' end of the neighboring gene [67]. In *H. seropedicae* SmR1, it was observed that the *IsrD* gene is superimposed on the 3' end of the *Hsero\_0991* gene, already for *IsrG*, expressed in the NFbHPN Malate-High Oxygen condition; it was also observed that it overlaps the *fimV* gene.

The TargetRNA2 computational tool was used to predict ncRNA targets in *H. seropedicae* SmR1. This program was specifically designed for the prediction of mRNA targets of ncRNAs with activity in *trans*. As a result, a wide range of mRNA targets were observed, regardless of the ncRNA that was analyzed, because after transcription the *cis*-encoded ncRNA make short and perfect pairing with their targets and the *trans*-encoded ncRNAs make long and imperfect pairing [68].

The number of targets for each ncRNA was variable. For example, ncHSmR1\_02 had 11 targets, ncHSmR1\_33 had 45 targets, and ncHSmR1\_23 with 99 targets. In the case of some ncRNAs, such as ncHSmR1\_01, there was no prediction of targets. The mRNA targets predicted in *H. seropedicae* SmR1 triggered by different biochemical mechanisms, for example, some of the targets of ncRNA 4.5S are as follows: *argB* (acetylglutamate kinase), *Hsero\_3355* (lipoprotein), *glk* (glucokinase), *htpX* (heat shock protein HtpX), *Hsero\_2759* (transcriptional regulatory protein).

The ncHSmR1\_45 with 11 targets, ncHSmR1\_149 with 45 targets, and ncHSmR1\_104 with 99 targets, ncHSmR1\_01 with 16 targets, ncHSmR1\_140 with 5 targets. Moreover, in the case of some non-coding RNAs, such as ncHSmR1\_42, there was no prediction of targets. This was probably because the mRNA sequences deposited in the TargetRNA2 tool, did not match ncHSmR1\_42. The mRNA targets predicted in *H. seropedicae* SmR1 are related to different biochemical mechanisms, for example, among the 4.5S RNA targets are: *argB* (acetylglutamate kinase), *Hsero\_3355* (lipoprotein), *glk* (glucokinase), *htpX* and HtpX (Heat shock protein), and *Hsero\_2759* (transcriptional regulatory protein). These results should lead to a better understanding of the participation of this type of RNA in the regulation of the metabolism of *Herbaspirillum seropedicae* SmR1.

Among the mRNA ratios for the predicted targets, an analysis was conducted of the likely target ncRNA-RNA interaction site, with an emphasis on those cases in which this interaction occurs near or over the RBS region. The first experimental evidence for the binding of an ncRNA to the ribosomal binding site of a messenger RNA and blocking the binding site, as a result of the base pairing between *micF* antisense RNA and the target *mFNA* mRNA, has been described in *Escherichia coli* in the final 5' regions of the latter [69].

Three ncRNAs were identified in *H. seropedicae* SmR1, with expression observed in RNA-seq material. The results indicate the existence of *cis*-encoded ncRNA, *trans*-encoded ncRNA, *riboswitch*, and CRISPR in *H. seropedicae* SmR1. *Micc*, *IsrK*, *IsrD*, and *IsrG* Hfq-dependent ncRNAs were detected among these. The number of target mRNAs predicted by the TargetRNA2 tool for the *H. seropedicae* SmR1 ncRNAs varied considerably and there was also a variation in the position of the mating region in relation to the AUG codon. The Infernal 1.1.1 tool did not detect ncRNAs in *H. lusitanum* P6-12 and *Herbaspirillum* sp. B65.

## 4. Materials and Methods

### 4.1. Genomes of Bacteria of the Genus *Herbaspirillum*

The genomes of *H. seropedicae* SmR1, *H. seropedicae* Z67, *H. lusitanum* P6-12, *H. hiltneri*, *H. frisingense* GSF30 N3, *Herbaspirillum* spp. strain B65, *H. rubrisubalbicans* M1, *H. autotrophicum* IAM 14942, *H. huttiense* subsp. *putei*, *Herbaspirillum* spp. strain B501, *Herbaspirillum* spp. strain GW103, *Herbaspirillum* spp. strain RV1423, *H. rhizosphaerae* UMS-37, *H. rubrisubalbicans* spp. strain Os34, *H. rubrisubalbicans* spp. strain Os45, and *Herbaspirillum* AU14040 were obtained from the database of the National Center for Biotechnology Information (NCBI) search. Species of *Herbaspirillum*, with access codes and date described, are in Supplementary Material S1.

#### 4.2. RNA-Seq Data

The researchers of the Biological Nitrogen Fixation (BNF) Centre kindly provided gene expression data of *H. seropedicae* SmR1. The in-plant culture and/or inoculation conditions are described in Supplementary Material S3.

#### 4.3. Prediction of ncRNAs in Silico

The *H. seropedicae* SmR1 genome was screened with the INFERENCE tools of RNA Alignment by employing a covariance model to show probabilistic profiles of sequences and secondary structures of RNA families [70]. The covariance model is a special case of a probabilistic mechanism that seeks to combine the consensus sequence with the secondary consensus structure of a given RNA. In many cases it is able to identify homologous RNAs that have a conserved secondary structure but low primary sequence conservation [71]. Infernal 1.1.1 available link <http://eddylab.org/infernal/>; accessed on 18 February 2016), consists of five programs: cmbuild, cmcalibrate, cmsearch, cmsearch, and cmalign. The following command line was used: cmsearch-o sequenceSMR1.out-tblout sequenceSMR1.tbl-T 24-notrunc Rfam.cm.sequenceSMR1 where sequenceSMR1, which represents the genome of the *H. seropedicae* SmR1 bacterium, has been replaced by the genomes listed in Section 4.1.

#### 4.4. ncRNA Sequence Annotation

The ncRNAs sequences were annotated using the Rfam online database. This database contains a collection of ncRNA families represented by sequences of manually edited alignments, consensus secondary structures, and annotations taken from taxonomic sources and ontology [70]. The base is a broad and diverse source of ncRNAs and includes 2791 families of ncRNAs, with information on various types of ncRNAs throughout the three life domains and viruses. Infernal 1.1.1 was used to make multiple sequence alignments by means of the covariance model. In addition to the annotation of ncRNAs, Rfam 14.0 classifies ncRNAs and provides bibliographical references for each family, links to the PDB (Protein Data Bank), ENA (European Nucleotide Archive), and Gene Ontology (GO). The ncRNA candidates were also compared with the online Ribe dataset [72] so as to identify putative *riboswitches*. This database is capable of recognizing conserved sequences and known *riboswitches* located upstream of orthologous genes in multiple genera [73]. Also, the ncRNA candidates were compared with the NCBI database using the BlastX/N webserver.

#### 4.5. Mapping and Visualization of Sequence Reads

We masked the rRNA sequences through a cross-match program and carried out a recursive read trimming at 5' and 3' to 35 nucleotides using a Perl script to map the short reads required for targeting the genome. The resulting Mate-Paired reads were aligned to the *H. seropedicae* SmR1 genome using the SHRiMP alignment tool [74]. The program was set up to tolerate three mismatches. We used SAMtools available link <http://samtools.sourceforge.net/>; accessed on 17 February 2016) [75] to convert data into the SAM/BAM format. Mapped RNA-seq reads in BAM format were visualized in the Artemis genome browser [76].

#### 4.6. Transcriptome Analysis

The standardization of the samples sequenced on the SoliD platform was carried out by the RPKM (reads per kilobase of transcript per million mapped reads) method, which estimates the gene expression value of a gene by measuring the density of the readings in a gene region of interest, and standardizing the readings counts in their exonic regions, compared with the original size of the gene or exon [77].

#### 4.7. Target Prediction

The TargetRNA2 prediction tool was employed to target mRNAs that base-pair with ncRNAs [68,78]. It calculates the hybridization score and the statistical significance of interactions between mRNA-ncRNA [79]. The tool uses a wide range of features to identify the target mRNAs, among them the conservation of ncRNA. This feature makes a comparison of the available sequence in GenBank in terms of the deposited genome (replicon) and indication of regions that have a greater conservation of sequences that are prone to undergoing mRNA/cRNA interactions. Another feature is the accessibility structure of the mRNA/mRNA/ncRNA, which is weighted by its stability and regions of interactions. In addition, account is taken of the energy hybridization regions of the mRNAs with a low energy index for hybridization in one or more regions of interactions with potential targets [68].

#### 4.8. Identifying CRISPR

A computational tool to identify CRISPR classification repeat conservation was used. This allows the independent grouping and determination of conserved sequence families, structural motifs suitable for endonucleases, and evolutionary relations [80]. CRISPRFinder is an interface that provides a detailed analysis of CRISPR genomic sequences [81].

### 5. Conclusions

The method used by Infernal1.1.1. for prediction of ncRNAs in the genus *Herbaspirillum* spp. prioritizes the identity of the various classes of ncRNAs. This RNA alignment inference has perspective combining the biological premise on associations of DNA sequences to RNA structure and sequence similarities. At the same time incorporation of the covariance models (CMs) scores a combination of sequence consensus and secondary RNA structure consensus, therefore, ncRNAs in the genus *Herbaspirillum* spp. were identified, given the importance of ncRNA expression in seq RNA in the strain *Herbaspirillum seropedicae* SmR1. The results provide a better understanding of the mechanism of post-transcriptional regulation of ncRNAs due to the interaction of mRNAs and the diversity of ncRNAs classification in the genus *Herbaspirillum* spp.

**Supplementary Materials:** Supplementary materials can be found at <http://www.mdpi.com/1422-0067/20/1/46/s1>.

**Author Contributions:** A.C.G. predicted of ncRNAs in the genus *Herbaspirillum* spp. and wrote the manuscript. V.L.P.d.S., T.C.S.C., L.M.C. and H.G. contributed to the funding acquisition and revision of the manuscript.

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

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

### References

1. Baldani, J.I.; Baldani, V.L.D.; Seldin, L.; Döbereiner, J. Characterization of *Herbaspirillum seropedicae* nov. sp. a root associated nitrogen fixing bacterium. *Int. J. Syst. Evol. Microbiol.* **1986**, *36*, 86–93. [CrossRef]
2. Baldani, J.I.; Pot, B.; Kirchof, G.; Falsen, E.; Baldani, V.L.D.; Olivares, F.L.; Hoste, B.; Kersters, K.; Hartmann, A.G.; Döbereiner, 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. *Int. J. Syst. Evol. Microbiol.* **1996**, *46*, 802–810.
3. 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. *Int. J. Syst. Evol. Microbiol.* **2004**, *54*, 2223–2230. [PubMed]

4. Kirchof, 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.* **2001**, *51*, 157–168. [[CrossRef](#)] [[PubMed](#)]
5. 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.* **2003**, *53*, 1979–1983. [[PubMed](#)]
6. Im, W.T.; Bae, H.S.; Yokota, A.; Lee, S.T. *Herbaspirillum chlorophenolicum* sp. nov., a 4-chlorophenol-degrading bacterium. *Int. J. Syst. Evol. Microbiol.* **2004**, *54*, 851–855. [[CrossRef](#)] [[PubMed](#)]
7. Rothballer, M.; Schmid, M.; Klein, I.; Gattinger, A.; Grundmann, S.; Hartmann, A. *Herbaspirillum hiltneri* sp. nov., isolated from surface-sterilized wheat roots. *Int. J. Syst. Evol. Microbiol.* **2006**, *56*, 1341–1348.
8. 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*. *Int. J. Syst. Evol. Microbiol.* **2007**, *57*, 2284–2288. [[CrossRef](#)]
9. Dobritsa, 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.* **2009**, *60*, 1418–1426.
10. Lagier, J.C.; Gimenez, G.; Robert, C.; Raoult, D.; Fournier, P.E. Non-contiguous finished genome sequence and description of *Herbaspirillum massiliense* sp. *Stand. Genom. Sci.* **2012**, *7*, 200–209. [[CrossRef](#)]
11. Carro, L.; 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. *Int. J. Syst. Evol. Microbiol.* **2012**, *62*, 1300–1306. [[CrossRef](#)] [[PubMed](#)]
12. 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.; et al. Biochemical characteristics, adhesion, and cytotoxicity of environmental and clinical isolates of *Herbaspirillum*. *J. Clin. Microbiol.* **2015**, *53*, 302–308. [[CrossRef](#)] [[PubMed](#)]
13. Baldani, J.I.; Baldani, V.L.D.; Sampaio, M.J.A.M.; Döbereiner, J. A fourth azospirillum species from cereal roots. In *Anais da Academia Brasileira de Ciência*; Brazilian Academy of Sciences: Rio de Janeiro, Brasil, 1984; Volume 56, pp. 351–365.
14. James, E.K.; Olivares, F.L.; Baldani, J.I.; Döbereiner, J. *Herbaspirillum*, an endophytic diazotroph colonizing vascular tissue in leaves of *Sorghum Bicolor* L. Moench. *J. Exper. Bot.* **1997**, *48*, 785–797. [[CrossRef](#)]
15. 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* **1992**, *13*, 65–73.
16. Döbereiner, J.; Baldani, V.L.D.; Baldani, J.I. *Como isolar e identificar bactérias diazotróficas de plantas não-leguminosas*; EMBRAPA-SPI: Brasília, Brazil, 1995; p. 60.
17. Olivares, F.L.; Baldani, V.L.D.; Reis, V.M.; Baldani, J.I.; Döbereiner, J. Occurrence of the endophytic diazotrophs *Herbaspirillum* spp. in roots, stems, and leaves, predominantly of Gramineae. *Biol. Fertil. Soils.* **1996**, *21*, 197–200. [[CrossRef](#)]
18. Storz, G.; Altuvia, S.; Wassarman, K.M. An abundance of RNA regulators. *Annu. Rev. Biochem.* **2005**, *74*, 199–217. [[CrossRef](#)]
19. Liu, J.M.; Camilli, A. A broadening world of bacterial small RNAs. *Curr. Opin. Microbiol.* **2010**, *13*, 18–23. [[CrossRef](#)]
20. Andersen, J.; Delihias, N.; Ikenaka, K.; Green, P.J.; Pines, O.; Ilercil, O.; Inouye, M. The isolation and characterization of RNA coded by the micF gene in *Escherichia coli*. *Nucleic Acids Res.* **1987**, *15*, 2089–2101. [[CrossRef](#)]
21. Ermolaeva, M.D.; Khalak, H.G.; White, O.; Smith, H.O.; Salzberg, S.L. Prediction of Transcription Terminators in Bacterial Genomes. *Mol. Biol.* **2000**, *301*, 27–33. [[CrossRef](#)]
22. 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* **2001**, *11*, 941–950. [[CrossRef](#)]
23. Ciampi, M.S. Rho-dependent terminators and transcription termination. *Microbiology* **2006**, *152*, 2515–2528. [[CrossRef](#)] [[PubMed](#)]
24. Wassarman, K.M.; Kiley, P.J. Global approaches for finding small RNA and small open reading frame functions. *J. Bacteriol.* **2010**, *192*, 26–28. [[CrossRef](#)] [[PubMed](#)]

25. Arnvig, K.B.; Cortes, T.; Young, D.B. Noncoding RNA in Mycobacteria. *Microbiol. Spectrum* **2014**, *2*, 1–24. [[CrossRef](#)] [[PubMed](#)]
26. Updegrave, T.B.; Zhang, A.; Storz, G. Hfq: The flexible RNA matchmaker. *Curr. Opin. Microbiol.* **2016**, *30*, 133–138. [[CrossRef](#)] [[PubMed](#)]
27. 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* **2015**, *16*, 1075.
28. Pichon, C.; Felden, B. Proteins that interact with bacterial small RNA regulators. *Fems. Microbiol. Rev.* **2007**, *31*, 614–625. [[CrossRef](#)]
29. Livny, J.; Waldor, M.K. Identification of small RNAs in diverse bacterial species. *Curr. Opin. Microbiol.* **2007**, *10*, 1096–1101. [[CrossRef](#)]
30. Hüttenhofer, A.; Schattner, P.; Polacek, P. Non-coding RNAs: Hope or hype? *Trends Genet.* **2005**, *21*, 289–297. [[CrossRef](#)]
31. Brantl, S. Regulatory mechanisms employed by *cis*-encoded antisense RNAs. *Curr. Opin. Microbiol.* **2007**, *10*, 102–109. [[CrossRef](#)]
32. Sittka, A.; Lucchini, S.; Papenfort, K.; Sharma, C.M.; Rolle, K.; Binnewies, T.T.; Hinton, J.C.; Vogel, J. Deep sequencing analysis of small non coding RNA and mRNA targets of the global post-transcriptional regulator Hfq. *PLoS Genet.* **2008**, *224*, 1–20.
33. Lorenz, R.; Bernhart, S.H.; Höner, Z.; Siederdisen, C.; Tafer, H.; Flamm, C.; Stadler, P.F.; Hofacker, I.L. Vienna RNA Package 2.0. *Algorithms. Mol. Biol.* **2011**, *6*, 1–14. [[CrossRef](#)] [[PubMed](#)]
34. 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* **2016**, *22*, 1373–1385. [[CrossRef](#)] [[PubMed](#)]
35. Repoila, F.; Darfeuille, F. Small regulatory non-coding RNAs in bacteria: Physiology and mechanistic aspects. *Biol. Cell* **2009**, *101*, 117–131. [[CrossRef](#)] [[PubMed](#)]
36. Faner, M.A.; Feig, A.L. Identifying and characterizing Hfq–RNA interactions. *Methods* **2013**, *63*, 144–159. [[CrossRef](#)] [[PubMed](#)]
37. Storz, G.; Vogel, J.; Wassarman, K.M. Regulation by small RNAs in bacteria: Expanding frontiers. *Mol. Cell* **2011**, *43*, 880–891. [[CrossRef](#)] [[PubMed](#)]
38. Waters, L.S.; Storz, G. Regulatory RNAs in Bacteria. *Cell* **2009**, *136*, 615–628. [[CrossRef](#)]
39. Delihias, N. Regulation of gene expression by *trans*-encoded antisense RNAs. *Mol. Microbiol.* **1995**, *15*, 411–414. [[CrossRef](#)]
40. Winkler, W.C. *Riboswitches* and the role of noncoding RNAs in bacterial metabolic control. *Curr. Opin. Chem. Biol.* **2005**, *9*, 594–602. [[CrossRef](#)]
41. Meyer, M.M.; Hammond, M.C.; Salinas, Y.; Roth, A.; Sudarsan, N.; Breaker, R.R. Challenges of ligand identification for riboswitch candidates. *RNA Biol.* **2011**, *8*, 5–10. [[CrossRef](#)]
42. Izar, B.; Mraheil, M.A.; Hain, T. Identification and Role of Regulatory Non-Coding RNAs in *Listeria monocytogenes*. *Int. J. Mol. Sci.* **2011**, *12*, 5070–5079. [[CrossRef](#)]
43. Jiang, W.; Fang, L.; Inouye, M. The role of the 5'-end untranslated region of the mRNA for CspA, the major cold-shock protein of *Escherichia coli*, in cold-shock adaptation. *J. Bacteriol.* **1996**, *178*, 4919–4925. [[CrossRef](#)] [[PubMed](#)]
44. Yamanaka, K.; Inouye, M. Growth-phase-dependent expression of *cspD*, encoding a member of the CspA family in *Escherichia coli*. *J. Bacteriol.* **1997**, *179*, 5126–5130. [[CrossRef](#)] [[PubMed](#)]
45. 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, Brazil, 2012.
46. 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.* **2012**, *40*, 2020–2031. [[CrossRef](#)] [[PubMed](#)]
47. Chen, S.; Zhang, A.; Blyn, L.B.; Storz, G. MicC, a second small-RNA regulator of Omp protein expression in *Escherichia coli*. *J. Bacteriol.* **2004**, *186*, 6689–6697. [[CrossRef](#)] [[PubMed](#)]
48. Wasserman, K.M. 6S RNA, a Global Regulator of Transcription. *Microbiol. Spectr.* **2018**, *6*. [[CrossRef](#)]
49. Higgins, D.G.; Sharp, P.M. CLUSTAL: A package for performing multiple sequence alignment on a microcomputer. *Gene* **1988**, *73*, 237–244. [[CrossRef](#)]

50. 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.* **2001**, *276*, 22844–22849. [[CrossRef](#)] [[PubMed](#)]
51. 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* **2001**, *2*, 293–301. [[CrossRef](#)]
52. Gu, S.Q.; Jöckel, J.; Beinker, P.; Warnecke, J.; Semenov, Y.P.; Rodnina, M.V.; Wintermeyer, W. Conformation of 4.5S RNA in the signal recognition particle and on the 30S ribosomal subunit. *RNA* **2005**, *11*, 1374–1384. [[CrossRef](#)]
53. 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* **2005**, *11*, 947–957. [[CrossRef](#)]
54. 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.* **2000**, *36*, 244–246. [[CrossRef](#)] [[PubMed](#)]
55. Jansen, G.; Leberer, E.; Thomas, D.Y.; Whiteway, M. Use of dominant negative mutations in analysis of G protein function in *Saccharomyces cerevisiae*. *Methods Enzymol.* **2002**, *344*, 82–91. [[PubMed](#)]
56. Barrangou, R.; Marraffini, L.A. CRISPR-Cas systems: Prokaryotes upgrade to adaptive immunity. *Mol. Cell* **2014**, *54*, 234–244. [[CrossRef](#)] [[PubMed](#)]
57. Hale, C.R.; Majumdar, S.; Elmore, J.; Pfister, N.; Compton, M.; Olson, S.; Resch, A.M.; Glover, C.V.R.D.; Graveley, B.R.; Terns, R.M.; et al. Essential features and rational design of CRISPR RNAs that function with the Cas RAMP module complex to cleave RNAs. *Mol. Cell* **2012**, *45*, 292–302. [[CrossRef](#)] [[PubMed](#)]
58. Estrella, M.A.; Kuo, F.T.; Bailey, S. RNA-activated DNA cleavage by the Type III-B CRISPR-Cas effector complex. *Genes. Dev.* **2016**, *30*, 460–470. [[CrossRef](#)] [[PubMed](#)]
59. Taylor, D.W.; Zhu, Y.; Staals, R.H.; Kornfeld, J.E.; Shinkai, A.; van der Oost, J.; Nogales, E.; Doudna, J.A. Structural biology. Structures of the CRISPR-Cmr complex reveal mode of RNA target positioning. *Science* **2015**, *348*, 581–585. [[CrossRef](#)] [[PubMed](#)]
60. Liu, L.; Li, X.; Ma, J.; Li, Z.; You, L.; Wang, J.; Wang, M.; Zhang, X.; Wang, Y. The Molecular Architecture for RNA-Guided RNA Cleavage by Cas13a. *Cell* **2017**, *170*, 714–726. [[CrossRef](#)] [[PubMed](#)]
61. Carter, E.L.; Boer, J.L.; Farrugia, M.A.; Flugge, N.; Towns, C.L.; Hausinger, R.P. Function of UreB in *Klebsiella aerogenes* urease. *Biochemistry* **2011**, *50*, 9296–9308. [[CrossRef](#)]
62. Wen, Y.; Feng, J.; Sachs, G. *Helicobacter pylori* 5 ureB-sRNA, a cis-encoded antisense small RNA, negatively regulates ureAB. *J. Bacteriol.* **2013**, *195*, 444–452. [[CrossRef](#)]
63. Abendroth, U.; Schmidtke, C.; Bonas, U. Small non-coding RNAs in plant-pathogenic *Xanthomonas*. *RNA Biol.* **2014**, *11*, 457–463. [[CrossRef](#)]
64. Dam, S.; Pagès, J.M.; Masi, M. Dual regulation of the small RNA MicC and the quiescent porin Omp in response to antibiotic stress in *Escherichia coli*. *Antibiotics* **2017**, *6*, 33. [[CrossRef](#)] [[PubMed](#)]
65. 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.* **2012**, *4*, 469–488. [[CrossRef](#)] [[PubMed](#)]
66. 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.* **2008**, *36*, 1913–1927. [[CrossRef](#)] [[PubMed](#)]
67. Kery, M.B.; Feldman, M.; Livny, J.; Tjaden, B. TargetRNA2: Identifying targets of small regulatory RNAs in bacteria. *Nucleic Acids Res.* **2014**, *42*, 124–129. [[CrossRef](#)] [[PubMed](#)]
68. Schmidt, M.; Zheng, P.; Delihias, N. Secondary structures of *Escherichia coli* antisense MicF RNA, the 5-end of the target OmpF mRNA, and the RNA/RNA duplex. *Biochemistry* **1995**, *34*, 3621–3631. [[CrossRef](#)] [[PubMed](#)]
69. Nawrocki, E.P.; Kolbe, D.L.; Eddy, S.R. Infernal 1.0: Inference of RNA alignments. *Bioinformatics* **2009**, *25*, 1335–1337. [[CrossRef](#)]
70. Nawrocki, E.P.; Eddy, S.R. Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* **2013**, *29*, 2933–2935. [[CrossRef](#)]
71. Abreu-Goodger, C.; Merino, E. RibEx: A web server for locating riboswitches and other conserved bacterial regulatory elements. *Nucleic Acids Res.* **2005**, *1*, 690–692. [[CrossRef](#)]
72. Raghavan, R.; Groisman, E.A.; Ochman, H. Genome-wide detection of novel regulatory RNAs in *E. coli*. *Genome Res.* **2011**, *21*, 1487–1497. [[CrossRef](#)]

73. David, M.; Dzamba, M.; Lister, D.; Ilie, L.; Brudno, M. SHRiMP2: Sensitive yet practical SHort read mapping. *Bioinformatics* **2011**, *27*, 1011–1012. [[CrossRef](#)]
74. 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* **2009**, *25*, 2078–2079. [[CrossRef](#)] [[PubMed](#)]
75. 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* **2010**, *26*, 676–677. [[CrossRef](#)] [[PubMed](#)]
76. Mortazavi, A.; Williams, B.A.; Mccue, K.; Schaeffer, L.; Wold, B. Mapping and Quantifying Mammalian Transcriptomes by RNA-Seq. *Nat. Methods* **2008**, *5*, 621–628. [[CrossRef](#)] [[PubMed](#)]
77. 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.* **2006**, *34*, 2791–2802. [[CrossRef](#)] [[PubMed](#)]
78. Tjaden, B. TargetRNA: A tool for predicting targets of small RNA action in bacteria. *Nucleic Acids Res.* **2008**, *36*, 109–113. [[CrossRef](#)] [[PubMed](#)]
79. Backofen, R.; Hess, W.R. Computational prediction of sRNAs and their targets in bacteria. *RNA Biol.* **2010**, *7*, 1–10. [[CrossRef](#)]
80. Lange, S.J.; Alkhnbashi, O.S.; Rose, D.; Will, S.; Backofen, R. CRISPRmap: An automated classification of repeat conservation in prokaryotic adaptive immune systems. *Nucleic Acids Res.* **2013**, *41*, 8034–8044. [[CrossRef](#)]
81. Grissa, I.; Vergnaud, G.; Pourcel, C. CRISPRFinder: A web tool to identify clustered regularly interspaced short palindromic repeats. *Nuclei. Acids Res.* **2007**, *35* (Suppl. 2), W52–W57. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).