In silico prediction and expression profile analysis of small non-coding RNAs in *Herbaspirillum seropedicae* SmR1

Tatiane Dobrzanski[1], Vânia Pobre[2*], Leandro Ferreira Moreno[1], Helba Cirino de Souza Barbosa[3], Rose Adele Monteiro[1,3], Fábio de Oliveira Pedrosa[1], Emanuel Maltempi de Souza[1], Cecília Maria Arraiano[2] and Maria Berenice Reynaud Steffens[1*]

Abstract

Background: *Herbaspirillum seropedicae* is a diazotrophic bacterium from the β-proteobacteria class that colonizes endophytically important gramineous species, promotes their growth through phytohormone-dependent stimulation and can express *nif* genes and fix nitrogen inside plant tissues. Due to these properties this bacterium has great potential as a commercial inoculant for agriculture. The *H. seropedicae* SmR1 genome is completely sequenced and annotated but despite the availability of diverse structural and functional analysis of this genome, studies involving small non-coding RNAs (sRNAs) has not yet been done. We have conducted computational prediction and RNA-seq analysis to select and confirm the expression of sRNA genes in the *H. seropedicae* SmR1 genome, in the presence of two nitrogen independent sources and in presence of naringenin, a flavonoid secreted by some plants.

Results: This approach resulted in a set of 117 sRNAs distributed in riboswitch, *cis*-encoded and *trans*-encoded categories and among them 20 have Rfam homologs. The housekeeping sRNAs tmRNA, ssrS and 4.5S were found and we observed that a large number of sRNAs are more expressed in the nitrate condition rather than the control condition and in the presence of naringenin. Some sRNAs expression were confirmed in vitro and this work contributes to better understand the post transcriptional regulation in this bacterium.

Conclusions: *H. seropedicae* SmR1 express sRNAs in the presence of two nitrogen sources and/or in the presence of naringenin. The functions of most of these sRNAs remains unknown but their existence in this bacterium confirms the evidence that sRNAs are involved in many different cellular activities to adapt to nutritional and environmental changes.

Keywords: *Herbaspirillum seropedicae* SmR1, sRNA, Regulatory RNA, Nitrogen fixation, Diazotrophic bacterium, Bacterial plant interaction

Background

Herbaspirillum seropedicae SmR1 is a diazotrophic and endophytic bacterium that belongs to the β-proteobacteria. This microorganism fixes nitrogen under microaerobic conditions inside the plant tissues of the economically important cereal crops wheat, rice, maize and sorghum [1]. *H. seropedicae* strains also appears associated with sugar cane and forage grasses [2, 3], fruit crops [4] and common bean [5]. Several studies have demonstrated the benefits of *Herbaspirillum*-plant interaction through the increase of the biomass of the inoculated plant [6–10]. Nitrogen cellular sources are the atmospheric dinitrogen and the nitrate, present in several environments. *H. seropedicae* SmR1 fixes nitrogen to ammonia in a reaction catalysed by the enzyme nitrogenase [11, 12]. The bacterial-plant interaction can promote plant growth and increase yield of crops since

* Correspondence: vaniapobre@itqb.unl.pt; steffens@ufpr.br
[2]Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av. da República, 2780-157 Oeiras, Portugal
[1]Department of Biochemistry and Molecular Biology, Universidade Federal do Paraná (UFPR), Av. Coronel. Francisco H. dos Santos, 210, PoBox 19046, Curitiba 81.531-980, Paraná, Brazil
Full list of author information is available at the end of the article

some compounds produced by the bacterium can stimulate the synthesis of phytohormones by plants [13]. Plants also play an important role in establishing this interaction since they produce compounds that affect their associations with microorganisms. One of such compounds is naringenin, a flavonoid produced as secondary metabolite, that can stimulate or inhibit specific genetic responses in different bacteria associated with plants [14, 15]. It was demonstrated that naringenin stimulates the endophytic colonization of *Arabidopsis thaliana* by *H. seropedicae* Z67 [16]. In *H. seropedicae* SmR1, naringenin regulates the expression of several genes, positively or negatively [17]. This microorganism can catabolize naringenin probably to obtain carbon and energy [18].

The single circular chromosome of the *H. seropedicae* SmR1 strain was sequenced and 4804 open reading frames were annotated [19]. Since then, there are many studies focusing the genomic structure, gene expression and physiology of *H. seropedicae* SmR1 [12, 17, 18, 20–22], but the investigation about the presence and function of small non-coding RNAs (sRNAs) was never performed. sRNAs have key regulatory roles in post-transcriptional control of gene expression. They can modulate turnover of target mRNAs and affect their translation [23–25]. They can be found in all three domains of life and are particularly important in bacteria allowing them to rapidly respond to environmental challenges [26, 27]. These molecules are 50–500 nucleotides long and are located predominantly in intergenic or in untranslated regions in the bacterial genomes [28]. sRNAs can be divided in two major categories *trans*-encoded RNAs (traRNA) and *cis*-encoded RNAs (caRNA) [29]. The caRNA can act at transcriptional or translational level and are sensory RNAs elements such as riboswitches that adopt two conformational structures in response to chemical signals such as small ligands [30–33]. The traRNA comprises the *trans*-encoded sRNAs that are partially complementary to their target [23] and the antisense small RNAs (asRNAs) that are totally complementary to their target [34, 35]. There are still traRNAs that bind to proteins to modulate their activity such as 6S RNA [36]. The paring of many traRNAS to mRNA target sites is facilitated by the RNA chaperone Hfq, a Sm family protein, which binds to adenine- and uridine-rich sequences (AU-motif) in sRNA [37, 38]. *H. seropedicae* SmR1 contains a conserved Hfq protein with a classic hexameric ring shape, observed in all available Hfq structures, with sRNA and mRNA contact surfaces [39]. The presence of a variety of types of small non-coding RNAs provides a versatile regulation of metabolic functions [25, 40–42]. The bioinformatic prediction of sRNAs followed by RNA-seq approach made possible genome

screens for sRNAs and has shown that there are much more bacterial regulatory sRNAs than previously thought [43]. In this study we applied in silico approach to predict sRNAs in *H. seropedicae* SmR1 genome and RNA-seq analysis to confirm their expression in bacteria grown in the presence of two nitrogen sources (ammonia or nitrate) and in the presence of naringenin. A set of 117 sRNAs transcripts were confirmed and some of them showed sequence identity with well-characterized sRNAs in other bacteria. Some sRNAs were experimentally detected confirming their existence.

Results

sRNAs in the *H. seropedicae* SmR1 genome

To search sRNAs in the genome of *H. seropedicae* SmR1 we used the nocoRNAc software [44], a bioinformatic tool that predicts sRNAs based on the co-localization of transcriptional terminators and promoter and is not limited to intergenic regions [44]. We identified 769 putative sRNAs. At the same time, we verified the presence of sRNAs transcripts in the RNA-seq data of *H. seropedicae* SmR1 using Cufflinks [45]. We were able to identify 1395 regions being transcribed which could encode sRNAs. Were analysed three RNA-seq data conditions obtained during the exponential growth phase of the bacterium: (i) control (CRT) - bacteria grown in NFbHPN medium containing NH_4Cl as nitrogen source, (ii) presence of naringenin (NAR) - bacteria cultured in NFbHPN medium containing NH_4Cl and the flavonoid naringenin, and (iii) nitrate (NIT) - bacteria grown in NFbHP medium containing KNO_3 as nitrogen source. Using the coverage criterion which established a minimum coverage ≥5 as a confidence level to select sRNAs in at least one of the three culture conditions we were able to verify the expression of 117 sRNAs transcripts in *H. seropedicae* SmR1 which have been termed Hsnc001 to Hsnc117 (Additional file 1). Forty sRNAs (34.5%) resulted only from Cufflinks, 63 sRNAs (54.3%) resulted only from nocoRNAc and 14 (12.1%) appeared in both approaches. These sRNAs range in length from 41 to 560 nucleotides and are equally distributed in the genome of *H. seropedicae* SmR1 being 67 annotated on the sense and 50 on the antisense strand in intergenic regions (Fig. 1a). According to the genomic location the sRNAs were distributed in riboswitch (10), *cis*-encoded (26) and *trans*-encoded (81) categories (Fig. 1b). Regarding base composition of the sRNAs we observed a range of 35 to 75% GC content, with an average of 54.97% (Additional file 1), whereas the genome has about 63.4% GC [19]. We observed that the riboswitches present high GC content (61.8%) and the known housekeeping sRNAs tmRNA, ssrS and 4.5S had 52.3, 53.7 and 65.7% GC content, respectively.

sRNAs of *H. seropedicae* SmR1

Comparison of *H. seropedicae* SmR1 117 sRNAs transcripts with RNA family's data base (Rfam) returned only 20 sequences with hits and information about putative function (Table 1). This result suggests that most of the *H. seropedicae* SmR1 sRNAs may be new or present low level of identity with those deposited in the Rfam database. Among the sRNAs identified the Toxic small RNA (tsRNA) and sucA RNA motif were found essentially in β-proteobacteria [46, 47] whereas YkkC/YxkD leader is present in some Cyanobacteria and Proteobacteria [48]. The sRNAs belonging to the family of small toxic RNAs in *H. seropedicae* SmR1 were Hsnc59, Hsnc63, Hsnc86 and Hsnc107 (Table 1). These small Toxic RNAs were found to be expressed in several strains of *Burkholderia cenocepacia* and, although they do not present a known function, they are capable of inhibiting *Escherichia coli* growth when introduced in a cloning vector [47, 49]. The Hsnc006 was annotated as the sucA 5'UTR which is considered a riboswitch candidate since the ligand that changes its conformation is still unknow [50]. This sequence is the 5′ UTR of *sucAsucBlpd* operon and, according to RNA-seq data, the RNA motif and the operon exhibited proportional expression level in the three conditions analysed (Additional file 1). We found two copies of YkkC/YxkD leader (Hsnc110 and Hsnc116) upstream HSERO_RS22365 and HSERO_RS22370, respectively, which encode two lipid kinases involved in the inorganic ion transport and metabolism. Recently this riboswitch was renamed as guanidin-I riboswitch

since it senses and responds to guanidine and controls genes that modify or pump guanidine as a toxic compound of bacteria [51].

We also found the TPP (Hsnc109) and FMN (Hsnc035) riboswitches. The TPP riboswitch is immediately upstream of HSERO_RS02120 (*thiC*) encoding the phosphomethylpyrimidine synthase and is known to bind directly to thiamine pyrophosphate (TPP) turning off TPP biosynthesis [52]. The FMN riboswitch is upstream of HSERO_RS09820 (*ribE*) encoding the 6,7-dimethyl-8-ribityllumazine synthase which catalyses one of last steps in the biosynthesis of riboflavin. FMN binds to the FMN aptamer and regulates the *ribE* expression [53]. *H. seropedicae* SmR1 still presents the SAH riboswitch (Hsnc113), Cobalamin riboswitch (Hsnc115), ZMP / ZTP riboswitch (Hsnc029), Yybp-ykoY (Hsnc112) and Fluoride riboswitch (Hsnc002). The SAH riboswitch is upstream HSERO_RS21435, encoding S-adenosyl- (L) -homocysteine (SAH), and is involved in S-adenosyl- (L) -methionine (SAM) regeneration cycle [54, 55] Cobalamin riboswitch is upstream HSERO_RS13325-HSERO_RS13320 operon encoding the cobalt transporter CbtB-CbtA acting in concert with vitamin B12 biosynthesis systems [56, 57]. ZMP/ZTP riboswitch regulates the expression of carbon metabolism genes [58, 59]. *H. seropedicae* SmR1 showed this riboswitch (Hsnc029) upstream to the HSERO_RS06140 (*glyA*) encoding serine hydroxymethyltransferase, a pyridoxal phosphate-dependent enzyme that plays an important role in the cellular pathways of a carbon [59].
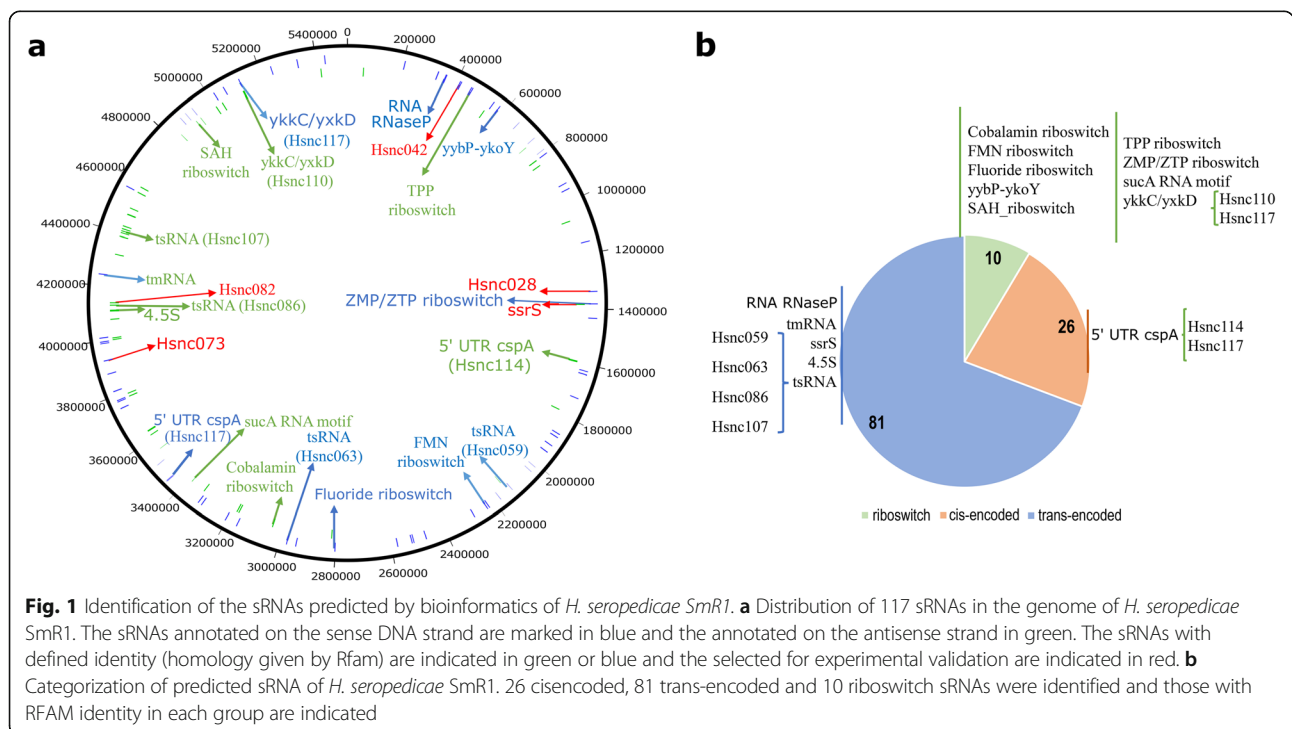

Fig. 1 Identification of the sRNAs predicted by bioinformatics of *H. seropedicae* SmR1. **a** Distribution of 117 sRNAs in the genome of *H. seropedicae* SmR1. The sRNAs annotated on the sense DNA strand are marked in blue and the annotated on the antisense strand in green. The sRNAs with defined identity (homology given by Rfam) are indicated in green or blue and the selected for experimental validation are indicated in red. **b** Categorization of predicted sRNA of *H. seropedicae* SmR1. 26 cisencoded, 81 trans-encoded and 10 riboswitch sRNAs were identified and those with RFAM identity in each group are indicated

Table 1 *H. seropedicae* SmR1 sRNAs identified in Rfam

Predicted			Rfam					
sRNA	size	GC%	ID	Acession	Start	End	Bit score	e-value
Hsnc001	99	65.66	4.5S	RF00169	1	99	76.4	8E-19
Hsnc002	111	60.36	Fluoride riboswitch	RF01734	8	81	50.3	4.4e-10
Hsnc006	94	56.38	sucA RNA motif	RF01070	12	93	79.8	2.3e-16
Hsnc029	100	61.00	ZMP/ZTP riboswitch	RF01750	1	100	59.3	2.6e-09
Hsnc035	169	65.09	FMN riboswitch	RF00050	1	169	112.6	5.3e-28
Hsnc050	177	53.67	ssrS (6S)	RF00013	1	177	67.3	2.5e-14
Hsnc059	93	49.46	Betaproteobacteria toxic RNA	RF02278	25	93	59.6	1.4e-12
Hsnc063	115	43.48	Betaproteobacteria toxic RNA	RF02278	48	115	62.4	3.4e-13
Hsnc083	384	52.34	tmRNA	RF00023	1	381	193.0	4.6e-57
Hsnc086	97	47.42	Betaproteobacteria toxic sRNA	RF02278	34	97	64.7	7.6e-14
Hsnc107	98	43.88	Betaproteobacteria toxic sRNA	RF02278	31	97	61.1	6.1e-13
Hsnc109	293	58.36	TPP riboswitch	RF00059	86	196	55.6	2.3e-10
Hsnc110	101	64.36	ykkC-yxkD	RF00442	1	101	99.0	1.8e-23
Hsnc111	335	61.19	RNA RNaseP	RF00010	1	335	212.1	1.1e-68
Hsnc112	184	63.04	yybP-ykoY	RF00080	16	184	47.7	6E-12
Hsnc113	90	65.56	SAH riboswitch	RF01057	1	90	49.1	9.6e-09
Hsnc114	373	51.47	5′ UTR cspA	RF01766	1	373	94.3	1.6e-24
Hsnc115	247	63.97	Cobalamin riboswitch	RF00174	1	247	111.2	1.6e-30
Hsnc116	100	60.0	ykkC-yxkD	RF00442	1	100	93.4	4.8e-22
Hsnc117	387	48.97	5′ UTR cspA	RF01766	1	388	86.7	2.7e-22

Yybp-ykoY is a manganese riboswitch that binds directly to Mn^{2+} and is associated with YebN/MntP genes [60–62]. In *H. seropedicae* SmR1, this riboswitch is located upstream HSERO_RS02630 encoding the manganese efflux pump MntP. In *Xanthomonas oryzae* this riboswitch acts as an essential Mn^{2+} sensor in infections during interaction with rice [61]. Fluoride riboswitch located upstream of HSERO_RS12335 coding the voltage-gated chloride channel protein. This riboswitch has been experimentally verified by [63] detecting fluoride and triggering the expression of genes that can help *Enterobacter cloacae* FRM to mitigate fluoride toxicity, using a fluorine carrier to expel fluoride from the cells.

Table 2 Oligonucleotides used in radiolabelling reactions

Probe	Sequence (5′ – 3′)
ssrS-F	CCGTGTTCGCGATTGCC
ssrS-T7	<u>TAATACGACTCACTATAGG</u>CCGGCATCCTGAACCTG
Hsnc042-F	GATGCCCGACTGCTGAAACG
Hsnc042-T7	<u>TAATACGACTCACTATAGG</u>TAGCGTCGGAATCGCGTTCCTG
Hsnc073-F	GCAATAACCAATGCGCAGG
Hsnc073-T7	<u>TAATACGACTCACTATAGG</u>GCATCATCAAGGGATGCCAG
Hsnc028	AAATCAGGCGTTTGTCATGGTTCGGTAAG
Hsnc082	AACGATGGAAGTACGGTGGTTCGCGTGATG

The T7 promoter sequence in the oligos is underlined

We found only two sRNAs with hits in the *cis*-encoded category, the *cspA* 5′UTR mRNAs Hsnc114 and Hsnc117. These sRNAs were identified in *H. seropedicae* SmR1 in untranslated regions of genes HSERO_RS07020 and HSERO_RS15195 encoding cold-shock proteins. These elements are known to be involved in the expression of *cspA* in response to temperature shift [64, 65] however it was already demonstrated that they might have a role in stress tolerance [66]. Both *H. seropedicae* *cspA* 5 'UTR mRNAs contain single-strand AU-motif in mRNA which could be binding sites of the RNA chaperone Hfq, as demonstrated in *E. coli* [67].

Some well-preserved housekeeping sRNAs were also found in *H. seropedicae* SmR1, ssrS or 6S RNA (Hsnc050), 4.5S RNA (Hsnc001) and tmRNA (Hsnc083). These sRNAs associate with proteins and are highly expressed in the cell. The 6S interacts with the primary form of holoenzyme of RNA polymerase, negatively regulates transcription and is involved in modulating stress and optimizing survival during nutrient limitation [68]. The 4.5S is part of the signal recognition particle (SRP) ribonucleoprotein complex [69]. In most bacteria the SRP consists of an RNA molecule (4.5S) and the Ffh protein that bind to ribosome stopping protein synthesis. The tmRNA (transfer messenger RNA) forms a ribonucleoprotein complex (tmRNP) that binds to bacterial

ribosomes that are blocked in the middle of protein synthesis; it is able to recycle the blocked ribosome by bringing a stop codon and g a proteolysis-inducing tag to unfinished polypeptides [70].

Expression of sRNA in *H. seropedicae* SmR1

Bacteria have a versatile system to respond quickly to environmental changes. In this process, many sRNAs are often expressed to regulate gene expression in a specific or different conditions and stages of growth. Thus, we wanted to determine how sRNAs are expressed in different culture conditions. The reading count of exponential phase in RNA-seq data was performed for each sRNA followed by normalization by Reads Per Kilobase Million (RPKM). A heat map that includes the 117 sRNA reveals different expression profiles according to the growth condition of *H. seropedicae* SmR1 (Fig. 2). We observed that a large cluster with 62 sRNAs are more expressed in the nitrate condition than in the control and naringenin conditions. Another cluster with 24 sRNAs is more expressed in the naringenin condition than in the control and nitrate conditions. This demonstrates the role of environment in the expression of sRNAs and the influence of sRNAs when the bacterium is exposed to certain nutritional conditions.

Experimental validation of *H. seropedicae* SmR1 sRNAs

We also used the coverage criterion of RNA-seq reads greater than or equal to five, in at least one of the three culture conditions (Additional file 1), to select Hsnc050, Hsnc028, Hsnc042, Hsnc073 and Hsnc082 sRNAs and validate their expression by Northern blot. Although coverage values do not represent expression quantification (as opposed to RPKM quantification), they can indicate that a given sRNA may be expressed. ssrS (Hsnc050) was chose as a control since it is a conserved housekeeping sRNA among bacterial species [68, 71, 72].
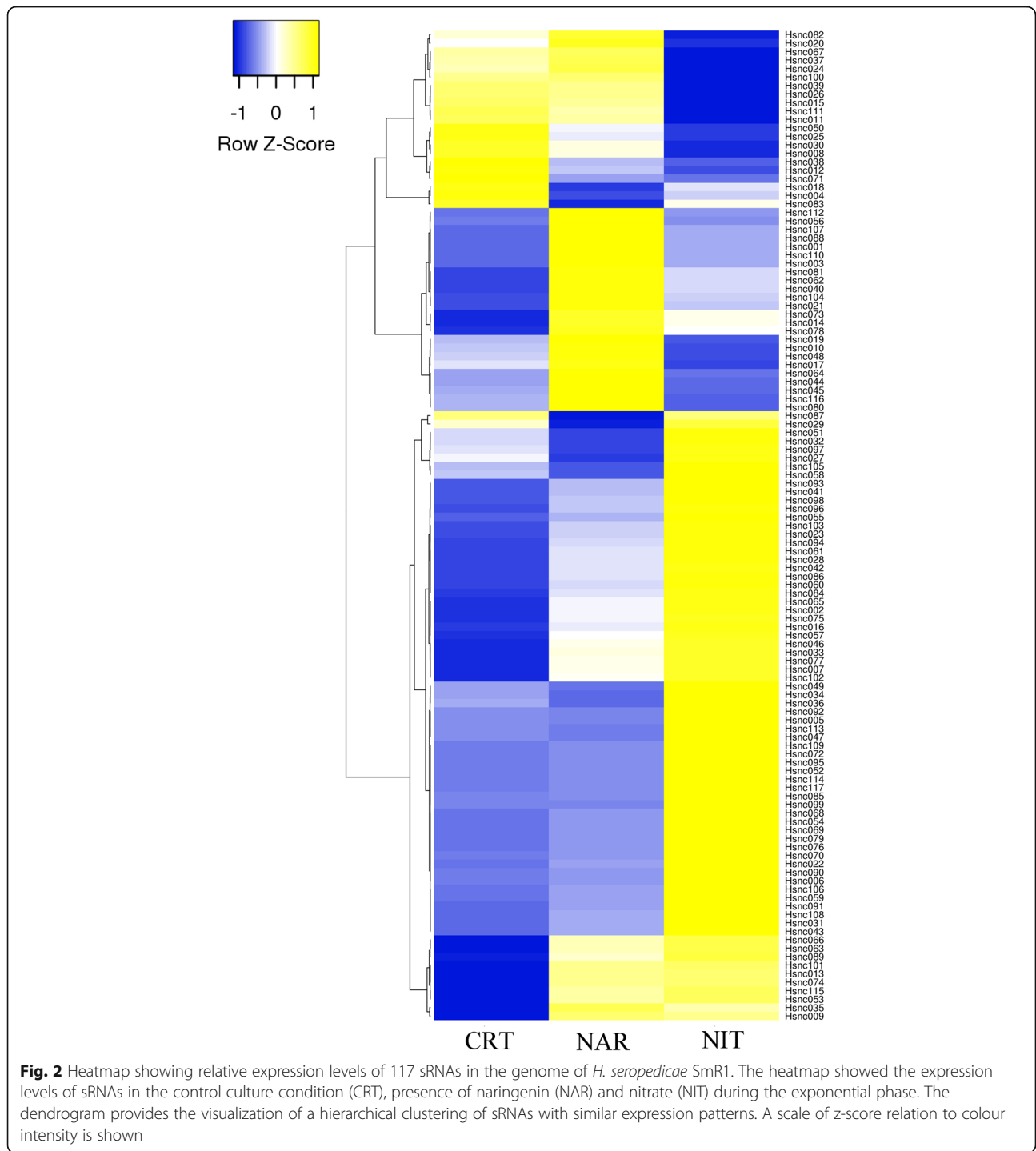
Considering that many sRNAs are induced under stress conditions, such as the lack of nutrients in the stationary phase, we evaluate the expression of the sRNAs in two phases of growth, exponential (OD_{600} = 0.7) and stationary phase (after 10 h of culture). *H. seropedicae* SmR1 was cultured under the CRT, NAR and NIT conditions and the total RNA was extracted and hybridized with specific radiolabelled probes. We were able to confirm that the five selected sRNAs are expressed in *H. seropedicae* SmR1 (Fig. 3). It is notable that all RNAs were expressed in all analysed growth conditions (CRT, NAR and NIT) as well as in the exponential and stationary phases (Fig. 3). We quantified at least three replicates of northern blots for each sRNA that was experimentally validated. The

CRT conditions of each growth phase were used as the standard of comparison between the different treatments. The expression of sRNAs did not show significant differences between culture conditions and growth phases (Fig. 3). However, a small decrease in expression for the Hsnc042 sRNA in the NAR condition in stationary phase and for Hsnc073 sRNA under NAR and NIT conditions in stationary phase was observed.

The function of these sRNAs is yet to be elucidated. In addition, we notice that Hsnc073 and Hsnc082 sRNAs showed lengths in the northern blot smaller ($\sim$ 78 nt and $\sim$ 120 nt respectively) than the length initially predicted by bioinformatics (182 nt and 157 nt respectively) (Fig 4). We suggest the smaller lengths of Hsnc073 and Hsnc082 sRNAs may be due to sRNA processing since most of the sRNAs are transcribed with larger length and then later processed by RNases for smaller functional lengths [28]. This difference can be also due to an imprecision in the prediction of sRNAs by the bioinformatics tools used in this work. Previously, we already observed that the nocoRNAc tool can predict sRNAs with larger than expected lengths, as the case of ssrS (6S). This sRNA was initially predicted with a length of 327 nucleotides by nocoRNAc, however based on RNA-seq we corrected its length to 177 nucleotides and, in fact, an RNA band with around this length was obtained in the northern blot (Fig. 3). Bacterial ssrS (6S) RNAs are generally transcribed as pre-6S RNA and then processed in 5′ end by ribonucleases that cuts a short sequence to mature form [73]. The *H. seropedicae* 6S RNA length is very close to the 6S of *Neisseria meningitidis* MC58 and *Pseudomonas aeruginosa* which are about 180 nucleotides [71]. The Hsnc028 (126 nucleotides) and Hsnc073 (182 nucleotides) also presented well distributed coverage in the RNA-seq profile (data not shown).
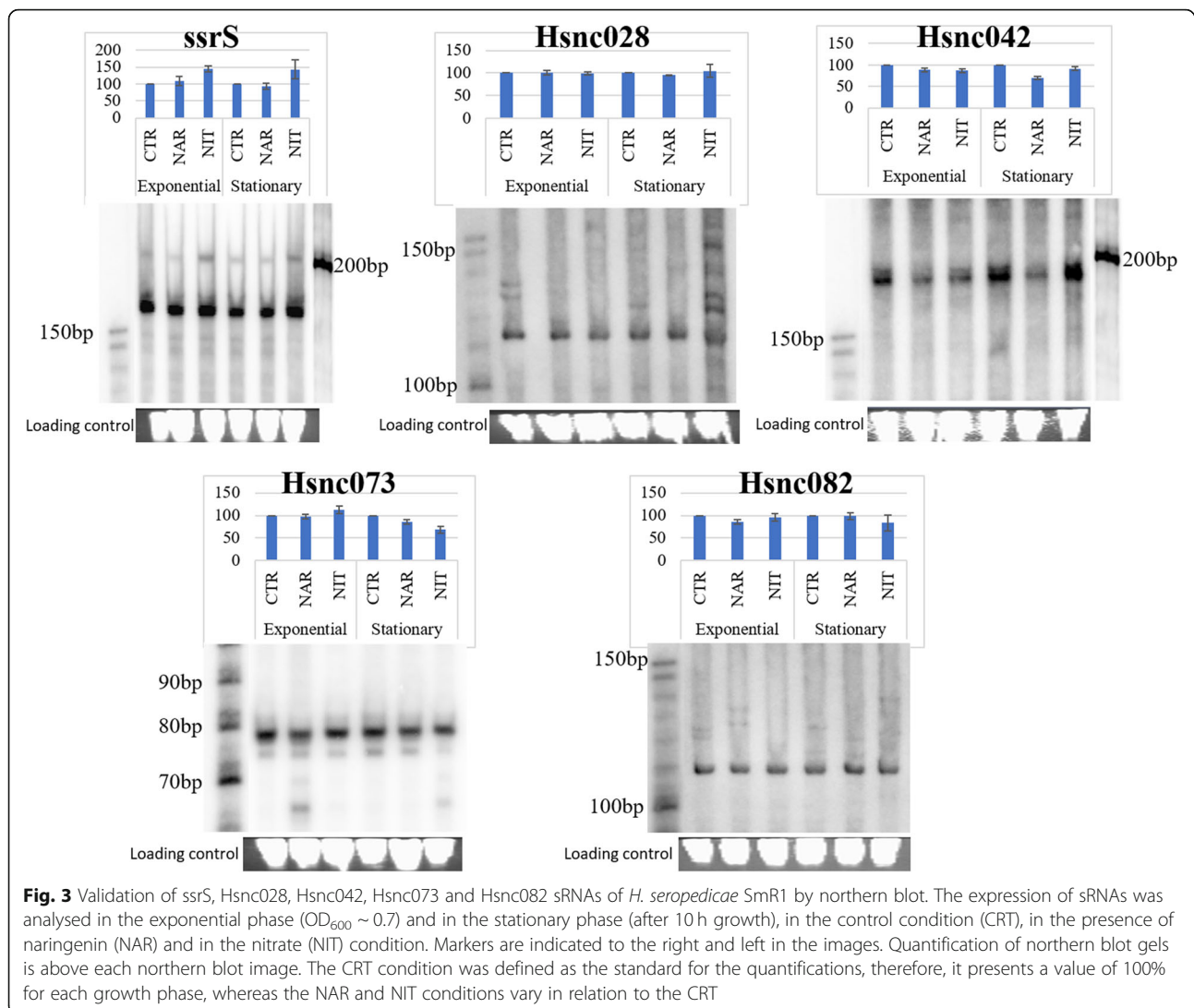
Discussion

In this work we identified and validated sRNAs in β-proteobacteria *H. seropedicae* SmR1, a highly versatile diazotrophic bacterium capable of metabolizing a wide range of carbon and nitrogen sources. Initially, we showed that 2164 sRNAs were predicted by bioinformatics tools Cufflinks and nocoRNAc that uses a specific prediction method for prokaryotic sRNA. To confirm which of these sRNAs are expressed in *H. seropedicae* SmR1, we analysed RNA-seq data from three culture conditions (CRT, NAR and NIT) and verified the expression of 117 sRNAs, each expressed in at least one of the three conditions analysed. The number of sRNAs predicted by bioinformatics tools varies among species as already described for *Streptomyces*

Dobrzanski *et al. BMC Genomics*   (2020) 21:134

Page 6 of 13

Fig. 2 Heatmap showing relative expression levels of 117 sRNAs in the genome of *H. seropedicae* SmR1. The heatmap showed the expression levels of sRNAs in the control culture condition (CRT), presence of naringenin (NAR) and nitrate (NIT) during the exponential phase. The dendrogram provides the visualization of a hierarchical clustering of sRNAs with similar expression patterns. A scale of z-score relation to colour intensity is shown

coelicolor (843 sRNAs) [44], *Burkholderia pseudomallei* (1306 sRNAs) [74] and *B. cenocepacia* J2315 (213 sRNAs) [75]. We can attribute this variation to the different methods of total RNA purification and to the sRNA prediction and validation methods employed. Concerning to *H. seropedicae* SMR1, we should also consider the stringent criterion adopted as a confidence level for sRNA expression (minimum coverage ≥5) to avoid data noise and

false-positives from the biocomputational prediction. Besides that, considering the size of the *H. seropedicae* SmR1 genome and the number of ORFs, we can expect a larger number of different sRNAs being expressed in other culture metabolic conditions. When we submitted the expressed sequences of the *H. seropedicae* SmR1 sRNAs (Additional file 1) to Rfam database the sequences we found only 20 sRNAs with hit to some RNA families

Fig. 3 Validation of ssrS, Hsnc028, Hsnc042, Hsnc073 and Hsnc082 sRNAs of *H. seropedicae* SmR1 by northern blot. The expression of sRNAs was analysed in the exponential phase ($OD_{600} \sim 0.7$) and in the stationary phase (after 10 h growth), in the control condition (CRT), in the presence of naringenin (NAR) and in the nitrate (NIT) condition. Markers are indicated to the right and left in the images. Quantification of northern blot gels is above each northern blot image. The CRT condition was defined as the standard for the quantifications, therefore, it presents a value of 100% for each growth phase, whereas the NAR and NIT conditions vary in relation to the CRT

(Table 1). We were able to confirm the expression of about 26 *cis*-encoded sRNAs in *H. seropedicae* SmR1 e two of them had hit to the family of *cspA* sRNAs according to Rfam. All 10 riboswitches showed good Rfam alignment scores (Table 1). Of the 81 trans-encoded sRNAs, 8 showed similarity given by Rfam. The occurrence of a large number of *H. seropedicae* SmR1 sRNAs with no hit in Rfam is not surprising since the database is populated with sRNAs from the most characterized model organisms, such as *E. coli*, *Salmonella enterica* and *Staphylococcus aureus* [76], and others that are phylogenetically distant from *H. seropedicae*. Considering the genome size of this bacterium we can expect a larger number of different sRNAs being expressed in other culture metabolic conditions. Thus, we suggest that *H. seropedicae* SmR1 may exhibit specific sRNAs with closer similarity to other organisms belonging to the same class. In β-proteobacteria the sRNAs have been described in *Burkholderia* species

although they were not well functionally characterized [47, 77, 78].

Many processes could be controlled by bacterial sRNAs induced by specific metabolic or environmental signals [79] or expressed constitutively under different growth conditions, such as housekeeping sRNAs (tmRNA and 4.5S) that participate in the regulation of genes that are constitutively expressed [80]. Interestingly, all *H. seropedicae* SmR1 sRNAs had a mean GC content (54.97% GC) below the mean of the genome content (63.4% GC). However, there are some sRNAs with slightly higher GC content such as riboswitches (61.81% GC) and known housekeeping sRNAs (ssrS, tmRNA and 4.5S with an average 57.22% GC). RNAs that exhibit high GC content generally exhibit a more rigid and conserved structure wich is probably necessary for the regulation of specific target molecules, such as maintenance genes or molecules that hardly evolve. In contrast, sRNAs that have a more flexible

Fig. 4 Coverage profiles of *H. seropedicae* SmR1 sRNAs in the exponential phase detected by RNA-seq. The y-axes indicate the reading coverage for each sRNA in the control conditions (purple lines), in the presence of naringenin (orange lines) and nitrate (blue lines). The x-axes denote the genomic positions according to the *H. seropedicae* SmR1 genome coordinates
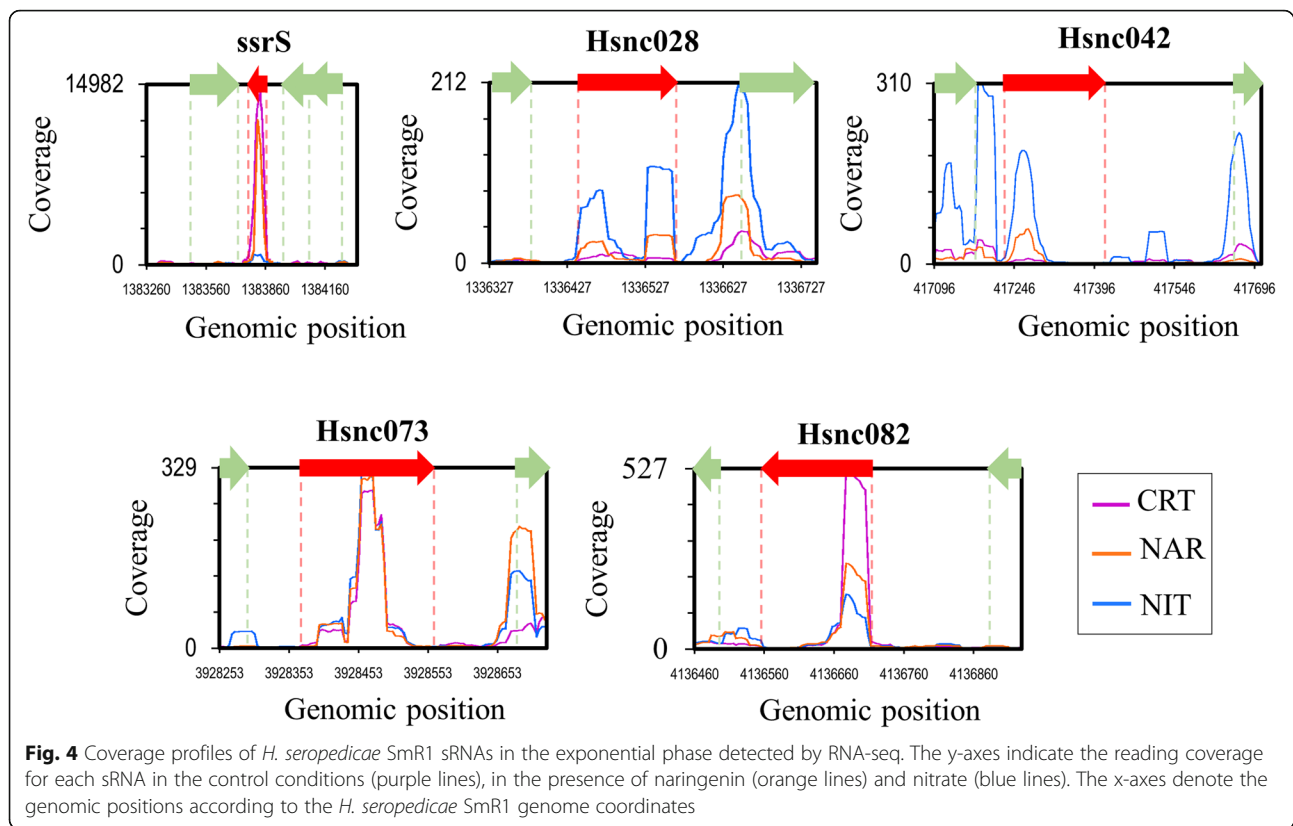
structure with a low GC content are probably involved in regulating the expression of several different genes or molecules that frequently evolve [77, 80, 81]. We have observed that the four Toxic sRNAs have low GC content (Hsnc063 43.48%, Hsnc086 47.42%, Hsnc107 43.88% and Hsnc059 49.46%). The Toxic sRNAs currently found exclusively in β-proteobacteria are *trans*-encoded and [49] and possibly interact with several mRNAs through alignment with ribosomal binding site (RBS). Then we suggest that these sRNAs have a more flexible structure with low GC content to control the regulation of different targets.

We created a heatmap to observe the expression of sRNAs between control and different culture conditions and observed a large cluster of sRNAs more expressed in the NIT condition. The main source of nitrogen for most organisms is ammonium and, in the absence or low concentration of ammonium, bacteria need to mobilize alternative nitrogen sources to maintain growth and increase chances of survival. In the absence or restriction of ammonium, *H. seropedicae* SmR1 can assimilate nitrate [12]. The RNA-seq profile revealed that the change in nitrogen source from ammonium to nitrate caused modifications in the pattern of gene expression in *H. seropedicae* SmR1, more than 37% of the genes were differentially expressed in the nitrate condition and the carbon consumption was increased [12]. Since sRNAs may play an important role in nutritional

deprivation [79], our data suggest that when the nitrate is the only nitrogen source many sRNAs may be influencing the post-transcriptional regulation of genes involved in carbon and nitrogen metabolism. When *H. seropedicae* SmR1 was cultivated in the presence of nitrate Bonato and collaborators [12] reported the increased expression of citric acid cycle genes with emphasis on *sucA* and *sucB* (7.3 and 7.7 respectively). In our work we observed the sucA RNA motif (Hsnc006) is highly expressed in the presence of nitrate (10 mmol/L) when compared to the control (20 mml/L NH_4Cl) (Additional file 1). Since this motif typically appears at 5′ end of β-proteobacteria *sucA* genes [46, 47] this could indicate a unique adaptation of the energetic metabolism of this bacterial Class to the changes in nitrogen sources. This hypothesis should be investigated.

A small cluster of sRNAs more expressed in the presence of naringenin was also observed in the heatmap. Naringenin is a plant-derived flavonoid and may act as a signal molecule during endophytic colonization by *H. seropedicae* SmR1 [82]. In this bacterium, naringenin triggers a change in gene expression to reduce motility and flagella synthesis [17]. An extensive bacterial sensory system for adaptation and survival to the plant environment was also observed during the early stages of colonization of maize [20]. Our data suggest that many sRNAs may be involved in the post-transcriptional regulation of genes related to the adaptation

and endophytic colonization of *H. seropedicae* SmR1 in plants. Further investigation is required to determine if some sRNAs may be related to flagella synthesis and bacterial motility as shown in other organisms [83, 84].

We were able to validate by Northern blot the expression of the Hsnc050 (ssrS RNA), Hsnc028, Hsnc042, Hsnc073 and Hsnc082 sRNAs. According to the Northern blot Hsnc073 and Hsnc082 sRNAs had smaller lengths than those predicted by bioinformatics. As we said earlier this difference could be due to sRNA processing or even imprecision in the prediction. Klein et al. (2202) also observed a length difference of sRNAs predicted for *Pyrococcus furiosus*, of the 11 sRNAs identified in Northern blot only one showed length corresponding to the initially predicted size [85]. *Sinorhizobium meliloti* also showed sRNAs with smaller lengths than predicted for sRNAs ssrS and sra25 (15 nt of difference) as well as for sm84 (30 nt of difference) and sm270 (10 nt of difference) [86]. The sRNA processing requires enzymatic cleavage to remove extra residues by ribonucleases to generate functional stable forms [87]. *H. seropedicae* SmR1 presents RNase E (Hsero_RS09410) an endoribonuclease which preferentially cleaves AU-rich regions [88] and affect sRNA biogenesis [89]. PNPase (Hsero_RS08755) could also trim sRNAs contributing to their maturation and/or degradation [90]. In fact, this difference in size can also account for the differences in expression observed between the RNA-Seq data and the Northern blots.

Conclusions
We reported the expression of several sRNA in *H. seropedicae* SmR1 genome in the presence of two nitrogen sources and/or in the presence of naringenin. The functions of the novel sRNAs remain unknown but their existence in this bacterium confirms the evidence that sRNAs are involved in many different cellular activities to adapt bacterium to nutritional and environmental changes. Some of them may participate in the regulation of nitrogen metabolism or in the bacterial-plant interaction. The discovery and knowledge of these sRNA molecules in this nitrogen fixation bacterium is very important due to its biotechnological significance.

Methods
Bacterial growth
H. seropedicae SmR1 (NCBI sequence: NC_014323.1) was grown at 30 °C and with agitation of 120 rpm in NFbHPN medium containing 80 µg/mL streptomycin [91]. Three growth conditions were used: (i) Control condition (CRT), bacteria grown on NFbHPN medium using malate as carbon source and 20 mmol/L NH_4Cl as

nitrogen source; (ii) Naringenin condition (NAR), bacteria grown in NFbHPN medium in the presence of flavonoid naringenin (100 µM); (iii) Nitrate condition (NIT), bacteria grown in NFbHP medium, using malate as carbon source and 10 mmol/L KNO3 as nitrogen source [12, 17].

Screening of small RNAs by nocoRNAc and Cufflinks
The *H. seropedicae* SmR1 genome was screened with the computational tool nocoRNAc [44] to search sRNAs by features that include promoter sequence, Rho-independent terminator and regions of sequence conservation. Furthermore, regions with high level of transcription, free of encoded proteins, were assessed in three individual RNA-seq data sets using *Cufflinks* [45] to verify the presence/expression of sRNAs. The nocoRNAc (non-coding RNA characterization) tool predicts putative sRNAs based on analyses not limited to intergenic regions. In order to find the location of candidates, firstly the SIDD sites, which are destabilized regions in the genomic DNA, were identified as putative promoter regions. The SIDD calculation was conducted using default values. After that, the putative Rho-independent terminators were predicted by the TranstermHP program, which is integrated in nocoRNAc. The tool was run using the standard protocol, with the option overwrite set up as described in the user guide [44]. The coordinates of SIDD sites with Rho-independent terminator are used for generating a list of putative sRNAs. Thus, the nocoRNAc tools may detect putative sRNAs in whole genome, even those which are encoded antisense from protein genes.

We assessed RNA-seq data to uncover sRNAs based only on read alignments. The reads from RNA-seq were mapped to *H. seropedicae* SmR1 genome with the tRNA, mRNA, rRNA and their 50 flaking nucleotides masked. We used the default parameters on the Cufflinks program to localize transcribed/expressed regions on the genome likely to encode sRNAs.

Mapping and visualization of sequence reads and analyses of predicted sRNAs
Before mapping the short reads to the *H. seropedicae* SmR1 reference genome, the rRNA sequences were masked using the cross-match program. Recursive trimming of the reads at 5′ and 3′ to 35 nucleotides were performed using a Perl script and the Mate-Paired reads were aligned to the reference genome using the alignment tool Short Read Mapping Program SHRiMP [92]. The program was set up to tolerate 3 mismatches. The maximal number of hits to each read was 1. Samtools [93] was used to convert data into SAM/BAM format. Mapped RNA-seq reads in BAM format were visualized

in the genome browse Artemis [94]. A sequence predicted as a sRNA was considered expressed when the minimum read coverage was 5-fold.

RNA -Seq data analysis

The RNA-seq data sets used in this work are available in the ArrayExpress database (www.ebi.ac.uk/arrayexpress) under accession number E-MTAB-3435 e E-MTAB-3646. Small non-coding RNA expression profiles was obtained with Artemis [95]. We employed an RPKM normalized expression values of three RNA-seq conditions describe in [12, 17] for the heatmap and hierarchical clustering. We used Pearson correlation as distance measure and average linkage as clustering method in Heatmapper [96]. Genome coordinate plot of non-coding RNA was performed with DNAPlotter [97].

RNA extraction and northern blot analysis

Overnight cultures grown in the CRT, NAR and NIT conditions were diluted in fresh medium to an initial $OD_{600} = 0.1$ and grown to exponential (OD_{600} 0.7) and stationary phase (10 h of growth). Culture samples were withdrawn and mixed with an equal volume of RNA stop buffer (10 mM Tris at pH 7.2, 5 mM $MgCl_2$, 25 mM NaN_3, and 500 μg/mL chloramphenicol). The total RNA extraction followed the protocol of cell lysis and phenol: chloroform extraction (adapted from [90]. After a precipitation step in ethanol and 300 mM sodium acetate, RNA was resuspended in MilliQ™-water. The integrity of RNA samples was evaluated by agarose gel electrophoresis. When necessary, Turbo DNase (Ambion) treatment following a new phenol: chloroform step was used to remove contaminant DNA. Next, 10–20 μg of total RNA was used to analyse small RNA expression on 10% polyacrylamide gels in TBE 1x buffer. RNA was transferred onto Hybond-N+ membrane (Amersham Biosciences) using TAE 1x as transfer buffer. RNAs were UV cross-linked to the membrane with a UVC 500 apparatus for 3 min (Amersham Biosciences). DNA templates carrying a T7 promoter sequence for in vitro transcription were generated by PCR using genomic DNA of *H. seropedicae* SmR1 and the primers listed in Table 2. Hsnc028 and Hsnc082 were detected by 5′ -end labelling of an antisense primer (Table 2). Radiolabelled probes using rUTP α-^{32}P (T7 probes) or ɣ-^{32}P ATP (primer probes) were purified on G25 Microspin columns (GE Healthcare). Hybridizations were carried out overnight at 42 °C or 68 °C with the PerfectHyb Plus Hybridization Buffer (Sigma). RNA Decade marker (Ambion) was used when detecting non-coding RNAs up to 150 nt; for longer transcripts, the 100–1000 bp Ladder (Biotools) was used. All radiochemicals were purchased from Perkin-Elmer.

Supplementary information

Additional file 1. *Herbaspirillum seropedicae* SmR1 predicted sRNAs. Expression of the 117 sRNAs of *H. seropedicae* SmR1 are shown in the table by Coverage and RPKM using RNA-seq data under CRT, NAR and NIT conditions. Defined identity sRNAs have the Rfam code listed. All sRNAs have features as the tools by which the sRNA was predicted (nocoRNAc and / or Cufflinks), category (trans-encoded, cis-encoded or riboswitch), percentage of G + C in the sRNA sequence, initial and final position of the sRNA found in the genome of *H. seropedicae* SmR1 as well as the sense of the DNA strand (sense or antisense) and the size of the sRNA.

Abbreviations

asRNA: Antisense small RNA; caRNA: Cis-encoded; RNA; CRT: Control condition; mRNA: Messenger RNA; NAR: Presence of naringenin condition; NIT: Nitrate condition; RBS: Ribosomal binding site; Rfam: RNA family database; RNA-seq: RNA sequencing; RPKM: Reads per kilobase million; rRNA: Ribosomal RNA; SAH: S-adenosyl- (L) –homocysteine; SAM: S-adenosyl-(L) –methionine; SIDD: Stress-induced (DNA) duplex destabilization; sRNA: Small non-coding RNA; SRP: Signal recognition particle; tmRNA: Transfer messenger RNA; tmRNP: Ribonucleoprotein complex; TPP: Thiamine pyrophosphate; traRNA: Trans-encoded RNA; tRNA: Transfer RNA; tsRNA: Toxic small RNA

Availability of data and materials

The RNA-Seq data analysed during the current study are available in the ArrayExpress database (www.ebi.ac.uk/arrayexpress) under accession number E-MTAB-2842 e E-MTAB-3435.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Department of Biochemistry and Molecular Biology, Universidade Federal do Paraná (UFPR), Av. Coronel. Francisco H. dos Santos, 210, PoBox 19046, Curitiba 81.531-980, Paraná, Brazil. [2]Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av. da República,

2780-157 Oeiras, Portugal. [3]Graduate Program in Bioinformatics, Universidade Federal do Paraná (UFPR), Rua Alcides Vieira Arcoverde, 1225, Curitiba 81520-260, Brazil.

References

1. Baldani JI, Pot B, Kirchhof G, Falsen E, Baldani VLD, Olivares FL, et al. Emended description of *Herbaspirillum*; inccusion 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 Bacteriol. 1996;46:802–10.
2. Pimentel JP, Olivares FL, Pitard RM, Urquiaga S, Akiba F, Döbereiner J. Dinitrogen fixation and infection of grass leaves by *Pseudomonas rubrisubalbicans* and *Herbaspirillum seropedicae*. In: Nitrogen fixation. Netherlands: Kluwer Academic Publishers; 1991. p. 225–9.
3. Olivares FL, Baldani VLD, Reis VM, Baldani JI, Döbereiner J. Occurrence of the endophytic diazotrophs *Herbaspirillum* spp. in roots, stems, and leaves, predominantly of Gramineae. Biol Fertil Soils. 1996;21:197–200.
4. Cruz LM, de Souza EM, Weber OB, Baldani JI, Döbereiner J, de Oliveira Pedrosa F. 16S ribosomal DNA characterization of nitrogen-fixing bacteria isolated from banana (Musa spp.) and pineapple (Ananas comosus (L.) Merril). Appl Environ Microbiol. 2001;67:2375–9.
5. Schmidt MA, Souza EM, Baura V, Wassem R, Yates MG, Pedrosa FO, et al. Evidence for the endophytic colonization of *Phaseolus vulgaris* (common bean) roots by the diazotroph *Herbaspirillum seropedicae*. Braz J Med Biol Res. 2011;44:182–5.
6. Gyaneshwar P, James EK, Reddy PM, Ladha JK. *Herbaspirillum* colonization increases growth and nitrogen accumulation in aluminium-tolerant rice varieties. New Phytol. 2002;154:131–45.
7. Baldani JI, Baldani VLD. History on the biological nitrogen fixation research in graminaceous plants: special emphasis on the Brazilian experience. An Acad Bras Cienc. 2005;77:549–79.
8. Estrada GA, Baldani VLD, de Oliveira DM, Urquiaga S, Baldani JI. Selection of phosphate-solubilizing diazotrophic *Herbaspirillum* and *Burkholderia* strains and their effect on rice crop yield and nutrient uptake. Plant Soil. 2013;369: 115–29.
9. Alves GC, Videira SS, Urquiaga S, Reis VM. Differential plant growth promotion and nitrogen fixation in two genotypes of maize by several *Herbaspirillum* inoculants. Plant Soil. 2014;387:307–21.
10. Kandel S, Joubert P, Doty S. Bacterial endophyte colonization and distribution within plants. Microorganisms. 2017;5:77.
11. Chubatsu LS, Monteiro RA, de Souza EM, de Oliveira MAS, Yates MG, Wassem R, et al. Nitrogen fixation control in *Herbaspirillum seropedicae*. Plant Soil. 2012;356:197–207.
12. Bonato P, Batista MB, Camilios-Neto D, Pankievicz VCS, Tadra-Sfeir MZ, Monteiro RA, et al. RNA-seq analyses reveal insights into the function of respiratory nitrate reductase of the diazotroph *Herbaspirillum seropedicae*. Environ Microbiol. 2016;18:2677–88.
13. Bastián F, Cohen A, Piccoli P, Luna V, Baraldi R, Bottini R. Production of indole-3-acetic acid and gibberellins A1and A3by *Acetobacter diazotrophicus* and *Herbaspirillum seropedicae* in chemically-defined culture media. Plant Growth Regul. 1998;24:7–11.
14. Balachandar D, Sandhiya GS, Sugitha TCK, Kumar K. Flavonoids and growth hormones influence endophytic colonization and *in planta* nitrogen fixation by a diazotrophic *Serratia* sp. in rice. World J Microbiol Biotechnol. 2006;22:707–12.
15. Hassan S, Mathesius U. The role of flavonoids in root-rhizosphere signalling: opportunities and challenges for improving plant-microbe interactions. J Exp Bot. 2012;63:3429–44.
16. Gough C, Galera C, Vasse J, Webster G, Cocking EC, Dénarié J. Specific flavonoids promote intercellular root colonization of *Arabidopsis thaliana* by *Azorhizobium caulinodans* ORS571. Mol Plant-Microbe Interact. 1997;10:560–70.
17. Tadra-Sfeir MZ, Faoro H, Camilios-Neto D, Brusamarello-Santos L, Balsanelli E, Weiss V, et al. Genome wide transcriptional profiling of *Herbaspirillum seropedicae* SmR1 grown in the presence of naringenin. Front Microbiol. 2015;6:1–8.
18. Marin AM, Souza EM, Pedrosa FO, Souza LM, Sassaki GL, Baura VA, et al. Naringenin degradation by the endophytic diazotroph *Herbaspirillum seropedicae* SmR1. Microbiology. 2013;159:167–75.
19. Pedrosa FO, Monteiro RA, Wassem R, Cruz LM, Ayub RA, Colauto NB, et al. Genome of *Herbaspirillum seropedicae* strain SmR1, a specialized diazotrophic endophyte of tropical grasses. PLoS Genet. 2011;7(5): e1002064.
20. Balsanelli E, Tadra-Sfeir MZ, Faoro H, Pankievicz VC, de Baura VA, Pedrosa FO, et al. Molecular adaptations of *Herbaspirillum seropedicae* during colonization of the maize rhizosphere. Environ Microbiol. 2016; 18:2343–56.
21. Pankievicz VCS, Camilios-Neto D, Bonato P, Balsanelli E, Tadra-Sfeir MZ, Faoro H, et al. RNA-seq transcriptional profiling of *Herbaspirillum seropedicae* colonizing wheat (*Triticum aestivum*) roots. Plant Mol Biol. 2016;90:589–603.
22. Batista MB, Teixeira CS, Sfeir MZT, Alves LPS, Valdameri G, Pedrosa FO, et al. PHB biosynthesis counteracts redox stress in *Herbaspirillum seropedicae*. Front Microbiol. 2018;9:472.
23. Waters LS, Storz G. Regulatory RNAs in Bacteria. Cell. 2009;136:615–28.
24. Lalaouna D, Simoneau-Roy M, Lafontaine D, Massé E. Regulatory RNAs and target mRNA decay in prokaryotes. Biochim Biophys Acta. 2013;1829:742–7.
25. Carrier M-C, Lalaouna D, Massé E. Broadening the definition of bacterial small RNAs: characteristics and mechanisms of action. Annu Rev Microbiol. 2018;72:141–61.
26. Dennis PP, Omer A. Small non-coding RNAs in Archaea. Curr Opin Microbiol. 2005;8:685–94.
27. Zaratiegui M, Irvine DV, Martienssen RA. Noncoding RNAs and gene silencing. Cell. 2007;128:763–76.
28. Storz G, Vogel J, Wassarman KM. Regulation by small RNAs in bacteria: expanding frontiers. Mol Cell. 2011;43:880–91.
29. Zorgani MA, Quentin R, Lartigue MF. Regulatory RNAs in the less studied streptococcal species: from nomenclature to identification. Front Microbiol. 2016;7:1161.
30. Winkler WC, Breaker RR. Regulation of bacterial gene expression by Riboswitches. Annu Rev Microbiol. 2005;59:487–517.
31. Henkin TM. Riboswitch RNAs: using RNA to sense cellular metabolism. Genes Dev. 2008;22:3383–90.
32. Narberhaus F, Vogel J. Regulatory RNAs in prokaryotes: here, there and everywhere. Mol Microbiol. 2009;74:261–9.
33. Demolli S, Geist MM, Weigand JE, Matschiavelli N, Suess B, Rother M. Development of β -lactamase as a tool for monitoring conditional gene expression by a tetracycline-riboswitch in *Methanosarcina acetivorans*. Archaea. 2014;2014:725610.
34. Brantl S. Regulatory mechanisms employed by cis-encoded antisense RNAs. Curr Opin Microbiol. 2007;10:102–9.
35. Saberi F, Kamali M, Najafi A, Yazdanparast A, Moghaddam MM. Natural antisense RNAs as mRNA regulatory elements in bacteria: a review on function and applications. Cell Mol Biol Lett. 2016;21:6.
36. Wassarman KM, Storz G. 6S RNA regulates *E. coli* RNA polymerase activity. Cell. 2000;101:613–23.
37. Sauer E. Structure and RNA-binding properties of the bacterial LSm protein Hfq. RNA Biol. 2013;10:610–8.
38. Updegrove TB, Shabalina SA, Storz G. How do base-pairing small RNAs evolve? FEMS Microbiol Rev. 2015;39:379–91.
39. Kadowaki MAS, Iulek J, Barbosa JARG, Pedrosa FDO, De Souza EM, Chubatsu LS, et al. Structural characterization of the RNA chaperone Hfq from the nitrogen-fixing bacterium *Herbaspirillum seropedicae* SmR1. Biochim Biophys Acta. 2012;1824:359–65.
40. Gottesman S, Storz G. 2010 bacterial small RNA regulators: versatile roles and rapidly evolving variations. Cold Spring Harb Perspect Biol. 2011;3: a003798.
41. Bobrovskyy M, Vanderpool CK. Regulation of bacterial metabolism by small RNAs using diverse mechanisms. Annu Rev Genet. 2013;47:209–32.
42. Dutta T, Srivastava S. Small RNA-mediated regulation in bacteria: a growing palette of diverse mechanisms. Gene. 2018;656:60–72.
43. Kåhrström CT. Bacterial transcription: introducing the excludon. Nat Rev Microbiol. 2012;10:443.
44. Herbig A, Nieselt K. NocoRNAc: characterization of non-coding RNAs in prokaryotes. BMC Bioinformatics. 2011;12:40.
45. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and cufflinks. Nat Protoc. 2012;7:562–78.

46. Greenlee EB, Stav S, Atilho RM, Brewer KI, Harris KA, Malkowski SN, et al. Challenges of ligand identification for the second wave of orphan riboswitch candidates. RNA Biol. 2018;6286:1–14.

47. Sass AM, Van Acker H, Förstner KU, Van Nieuwerburgh F, Deforce D, Vogel J, et al. Genome-wide transcription start site profiling in biofilm-grown *Burkholderia cenocepacia* J2315. BMC Genomics. 2015; 16:1–15.

48. Barrick JE, Corbino KA, Winkler WC, Nahvi A, Mandal M, Collins J, et al. New RNA motifs suggest an expanded scope for riboswitches in bacterial genetic control. Proc Natl Acad Sci. 2004;101:6421–6.

49. Kimelman A, Levy A, Sberro H, Kidron S, Leavitt A, Amitai G, et al. A vast collection of microbial genes that are toxic to bacteria. Genome Res. 2012; 22:802–9.

50. Sun EI, Leyn SA, Kazanov MD, Saier MH, Novichkov PS, Rodionov DA. Comparative genomics of metabolic capacities of regulons controlled by cis-regulatory RNA motifs in bacteria. BMC Genomics. 2013;14:597.

51. Nelson JW, Atilho RM, Sherlock ME, Stockbridge RB, Breaker RR. Metabolism of free guanidine in bacteria is regulated by a widespread riboswitch class. Mol Cell. 2017;65:220–30.

52. Mehdizadeh Aghdam E, Sinn M, Tarhriz V, Barzegar A, Hartig JS, Hejazi MS. TPP riboswitch characterization in *Alishewanella tabrizica* and *Alishewanella aestuarii* and comparison with other TPP riboswitches. Microbiol Res. 2017; 195:71–80.

53. Pedrolli D, Langer S, Hobl B, Schwarz J, Hashimoto M, Mack M. The ribB FMN riboswitch from *Escherichia coli* operates at the transcriptional and translational level and regulates riboflavin biosynthesis. FEBS J. 2015;282: 3230–42.

54. Wang JX, Lee ER, Morales DR, Lim J, Breaker RR. Riboswitches that sense S-adenosylhomocysteine and activate genes involved in coenzyme recycling. Mol Cell. 2008;29:691–702.

55. Edwards AL, Reyes FE, Héroux A, Batey RT. Structural basis for recognition of S-adenosylhomocysteine by riboswitches. RNA. 2010;16:2144–55.

56. Rodionov DA, Vitreschak AG, Mironov AA, Gelfand MS. Comparative genomics of the vitamin B12 metabolism and regulation in prokaryotes. J Biol Chem. 2003;278:41148–59.

57. Polaski JT, Webster SM, Johnson JE, Batey RT. Cobalamin riboswitches exhibit a broad range of ability to discriminate between methylcobalamin and adenosylcobalamin. J Biol Chem. 2017;292:11650–8.

58. Ducker GS, Rabinowitz JD. ZMP: a master regulator of one-carbon metabolism. Mol Cell. 2015;57:203–4.

59. Kim PB, Nelson JW, Breaker RR. An ancient riboswitch class in bacteria regulates purine biosynthesis and one-carbon metabolism. Mol Cell. 2015;57:317–28.

60. Dambach M, Sandoval M, Updegrove TB, Anantharaman V, Aravind L, Waters LS, et al. The ubiquitous yybP-ykoY riboswitch is a manganese-responsive regulatory element. Mol Cell. 2015;57:1099–109.

61. Li C, Liu S, Liu P, Wang Y, Xu C, Tao J, et al. The YebN leader RNA acts as a Mn^{2+} sensor required for the interaction of *Xanthomonas oryzae* and rice. Mol Plant-Microbe Interact. 2018;31(9):932–39.

62. Price IR, Gaballa A, Ding F, Helmann JD, Ke A. Mn^{2+}-sensing mechanisms of yybP-ykoY orphan riboswitches. Mol Cell. 2015;57:1110–23.

63. Liu X, Tian J, Liu L, Zhu T, Yu X, Chu X, et al. Identification of an operon involved in fluoride resistance in *Enterobacter cloacae* FRM. Sci Rep. 2017;7:1–9.

64. 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;178:4919–25.

65. Zhang Y, Burkhardt DH, Rouskin S, Li GW, Weissman JS, Gross CA. A stress response that monitors and regulates mRNA structure is central to cold shock adaptation. Mol Cell. 2018;70:274–286.e7.

66. Derman Y, Söderholm H, Lindström M, Korkeala H. Role of *csp* genes in NaCl, pH, and ethanol stress response and motility in *Clostridium botulinum* ATCC 3502. Food Microbiol. 2015;46:463–70.

67. Hankins JS, Denroche H, Mackie GA. Interactions of the RNA-binding protein Hfq with cspA mRNA, encoding the major cold shock protein. J Bacteriol. 2010;192:2482–90.

68. Steuten B, Hoch PG, Damm K, Schneider S, Köhler K, Wagner R, et al. Regulation of transcription by 6S RNAs insights from the *Escherichia coli* and *Bacillus subtilis* model systems. RNA Biol. 2014;11:508–21.

69. Mercier E, Holtkamp W, Rodnina MV, Wintermeyer W. Signal recognition particle binds to translating ribosomes before emergence of a signal anchor sequence. Nucleic Acids Res. 2017;45:11858–66.

70. Sundermeier TR, Dulebohn DP, Cho HJ, Karzai AW. A previously uncharacterized role for small protein B (SmpB) in transfer messenger RNA-mediated trans-translation. Proc Natl Acad Sci. 2005;102:2316–21.

71. Barrick JE, Sudarsan N, Weinberg Z, Ruzzo WL, Breaker RR. 6S RNA is a widespread regulator of eubacterial RNA polymerase that resembles an open promoter. RNA. 2005;11:774–84.

72. Trotochaud AE, Wassarman KM. A highly conserved 6S RNA structure is required for regulation of transcription. Nat Struct Mol Biol. 2005;12:313–9.

73. Burenina OY, Elkina DA, Hartmann RK, Oretskaya TS, Kubareva EA. Small noncoding 6S RNAs of bacteria. Biochemistry. 2015;80:1429–46.

74. Khoo J, Chai S, Mohamed R, Nathan S, Firdaus-Raih M. Computational discovery and RT-PCR validation of novel *Burkholderia* conserved and *Burkholderia pseudomallei* unique sRNAs. BMC Genomics. 2012;13(Suppl 7):S13.

75. Coenye T, Drevinek P, Mahenthiralingam E, Shah SA, Gill RT, Vandamme P, et al. Identification of putative noncoding RNA genes in the *Burkholderia cenocepacia* J2315 genome. FEMS Microbiol Lett. 2007;276:83–92.

76. Sharma CM, Heidrich N. Small RNAs and virulence in bacterial pathogens. RNA Biol. 2012;9:361–3.

77. Sass A, Kiekens S, Coenye T. Identification of small RNAs abundant in *Burkholderia cenocepacia* biofilms reveal putative regulators with a potential role in carbon and iron metabolism. Sci Rep. 2017;7:1–13.

78. Ghosh S, Dureja C, Khatri I, Subramanian S, Raychaudhuri S, Ghosh S. Identification of novel small RNAs in *Burkholderia cenocepacia* KC-01 expressed under iron limitation and oxidative stress conditions. Microbiology. 2017;163:1924–36.

79. Michaux C, Verneuil N, Hartke A, Giard JC. Physiological roles of small RNA molecules. Microbiology. 2014;160(Part 6):1007–19.

80. Hershberg R, Altuvia S, Margalit H. A survey of small RNA-encoding genes in *Escherichia coli*. Nucleic Acids Res. 2003;31:1813–20.

81. Zhao T, Zhang R, Wang M. Prediction of candidate small non-coding RNAs in by *Agrobacterium* by computational analysis. J Biomed Res. 2010;24:33–42.

82. Tadra-Sfeir MZ, Souza EM, Faoro H, Müller-Santos M, Baura VA, Tuleski TR, et al. Naringenin regulates expression of genes involved in cell wall synthesis in *Herbaspirillum seropedicae*. Appl Environ Microbiol. 2011;77: 2180–3.

83. Bak G, Lee J, Suk S, Kim D, Lee JY, Kim KS, et al. Identification of novel sRNAs involved in biofilm formation, motility, and fimbriae formation in *Escherichia coli*. Sci Rep. 2015;5:1–19.

84. Fuentes DN, Calderón PF, Acuña LG, Rodas PI, Paredes-Sabja D, Fuentes JA, et al. Motility modulation by the small non-coding RNA SroC in *Salmonella typhimurium*. FEMS Microbiol Lett. 2015;362:1–8.

85. Klein RJ, Misulovin Z, Eddy SR. Noncoding RNA genes identified in AT-rich hyperthermophiles. Proc Natl Acad Sci. 2002;99:7542–7.

86. Valverde C, Livny J, Schlüter J-P, Reinkensmeier J, Becker A, Parisi G. Prediction of *Sinorhizobium meliloti* sRNA genes and experimental detection in strain 2011. BMC Genomics. 2008;9:416.

87. Matos RG, Casinhas J, Bárria C, Dos Santos RF, Silva IJ, Arraiano CM. The role of ribonucleases and sRNAs in the virulence of foodborne pathogens. Front Microbiol. 2017;8:1–7.

88. Arraiano CM, Andrade JM, Domingues S, Guinote IB, Malecki M, Matos RG, et al. The critical role of RNA processing and degradation in the control of gene expression. FEMS Microbiol Rev. 2010;34:883–923.

89. Chao Y, Li L, Girodat D, Förstner KU, Said N, Corcoran C, et al. In vivo cleavage map illuminates the central role of RNase E in coding and non-coding RNA pathways. Mol Cell. 2017;65:39–51.

90. Andrade JM, Pobre V, Matos AM, Arraiano CM. The crucial role of PNPase in the degradation of small RNAs that are not associated with Hfq. RNA. 2012; 18:844–55.

91. Klassen G, Pedrosa FO, Souza EM, Funayama S, Rigo LU. Effect of nitrogen compounds on nitrogenase activity in *Herbaspirillum seropedicae* SMR1. Can J Microbiol. 1997;43:887–91.

92. Rumble SM, Lacroute P, Dalca AV, Fiume M, Sidow A, Brudno M. SHRiMP: accurate mapping of short color-space reads. PLoS Comput Biol. 2009;5: e1000386.

93. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. Bioinformatics. 2009;25: 2078–9.

94. Carver T, Harris SR, Berriman M, Parkhill J, McQuillan JA. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. Bioinformatics. 2011;28:464–9.

95. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, et al.
     Artemis: sequence visualization and annotation. Bioinformatics. 2000;16:944–5.
96. Babicki S, Arndt D, Marcu A, Liang Y, Grant JR, Maciejewski A, et al.
     Heatmapper: web-enabled heat mapping for all. Nucleic Acids Res. 2016;44:
     W147–53.
97. Carver T, Thomson N, Bleasby A, Berriman M, Parkhill J. DNAPlotter: circular
     and linear interactive genome visualization. Bioinformatics. 2009;25:119–20.

Publisher's Note