NsiR1, a small RNA with multiple copies, modulates heterocyst differentiation in the cyanobacterium *Nostoc* sp. PCC 7120

Manuel Brenes-Álvarez [ID], **Marina Minguet,**[†]
Agustín Vioque [ID][*] **and Alicia M. Muro-Pastor** [ID]
Instituto de Bioquímica Vegetal y Fotosíntesis, Consejo Superior de Investigaciones Científicas and Universidad de Sevilla, Sevilla, Spain.

Summary

Upon nitrogen starvation, filamentous cyanobacteria develop heterocysts, specialized cells devoted to the fixation of atmospheric nitrogen. Differentiation of heterocyst at semi-regular intervals along the filaments requires complex structural and functional changes that are under the control of the master transcriptional regulator HetR. NsiR1 (nitrogen stress-induced RNA 1) is a HetR-dependent non-coding RNA that is expressed specifically in heterocysts from a very early stage of differentiation. In the genome of *Nostoc* sp. PCC 7120 there are 12 tandem copies of *nsiR1* (*nsiR1.1* to *nsiR1.12*), seven of them with identical sequence (*nsiR1.3* to *nsiR1.9*) and the others slightly divergent. *nsiR1.1* is transcribed antisense to the 5′ UTR of *hetF*, a gene required for heterocyst development. Here, we show that binding of NsiR1.1 inhibits translation of the *hetF* mRNA by inducing structural changes in its 5′ UTR. Altered levels of NsiR1 result in different phenotypic alterations including enlarged cell size and delayed heterocyst development that could be related to a reduced amount of HetF.

Introduction

Cyanobacteria are the only prokaryotes able to perform oxygenic photosynthesis and important contributors to global primary productivity. They constitute a coherent phylogenetic group with diverse morphology and wide ecological distribution. *Nostoc* sp. PCC 7120 (a.k.a *Anabaena* sp. PCC 7120) is a filamentous cyanobacterium that under

nitrogen deprivation differentiates heterocysts, specialized cells devoted to fixation of atmospheric nitrogen. In the absence of combined nitrogen, filaments are composed of photosynthetic vegetative cells and regularly spaced (every 10–15 cells) nitrogen-fixing heterocysts. Heterocyst differentiation involves a complex developmental program (Flores and Herrero, 2010; Muro-Pastor and Hess, 2012; Herrero *et al*., 2016; Flores *et al*., 2019) that is ultimately under the control of NtcA, the global regulator of nitrogen assimilation (Herrero *et al*., 2001), and also under the control of HetR, a specific regulator of cell differentiation (Buikema and Haselkorn, 1991; Black *et al*., 1993). At the initiation of heterocyst differentiation, induction of the expression of NtcA and HetR depends on each other (Muro-Pastor *et al*., 2002) and leads to increased expression of HetR in those cells becoming heterocysts. The expression of *hetR* is regulated at several levels. *hetR* is transcribed from a complex promoter including several transcriptional start sites (TSSs), one of them specifically induced in cells becoming heterocysts (Rajagopalan and Callahan, 2010; Mitschke *et al*., 2011). HetR is a DNA-binding protein that positively regulates its own expression and the expression of many genes, including negative elements essential for the maintenance of the pattern of heterocysts along the filaments, such as *patS* (Yoon and Golden, 1998), *hetN* (Callahan and Buikema, 2001) and *patX* (Elhai and Khudyakov, 2018). The binding of the pentapeptide RGSGR, which is part of the sequence of PatS, blocks the DNA-binding activity of HetR (Huang *et al*., 2004) and signals its degradation (Risser and Callahan, 2009). HetR is an autoproteolytic protein (Zhou *et al*., 1998) that seems also to be degraded by the HetF protease (Wong and Meeks, 2001; Risser and Callahan, 2008). A *hetF* mutant accumulates a large amount of HetR, has an altered morphology with enlarged cells and does not differentiate heterocysts (Risser and Callahan, 2008), suggesting HetR turnover is required for differentiation. However, since *hetF* transcription seems unpatterned along the filaments (Risser and Callahan, 2008), the question of how HetF may contribute to the differential accumulation of active HetR in heterocysts versus vegetative cells remains open.

Small non-coding RNAs (sRNAs) are important regulators of gene expression in bacteria, including cyanobacteria (Kopf and Hess, 2015; Muro-Pastor and Hess, 2020). Most frequently they bind to their target mRNAs close to the ribosome binding site, therefore inhibiting translation, but a large diversity of alternative regulatory mechanisms operated by sRNAs is known (Wagner and Romby, 2015). They are involved in the regulation of diverse processes related to response to stress or adaptation to a changing environment, including cell differentiation. In *Bacillus* several sRNAs have been identified that are sporulation-specific and with expression restricted to the forespore (Silvaggi *et al*., 2006; Schmalisch *et al*., 2010; Marchais *et al*., 2011; Mars *et al*., 2016). In heterocyst-forming cyanobacteria, sRNAs and antisense RNAs with heterocyst-specific expression have been identified (Ionescu *et al*., 2010; Brenes-Álvarez *et al*., 2016; Brenes-Álvarez *et al*., 2019). A heterocyst-specific antisense RNA is involved in carbon fixation shutdown in heterocysts (Olmedo-Verd *et al*., 2019).

NsiR1 is an sRNA of about 60 nucleotides encoded in 12 tandem copies in the region upstream of *hetF* in *Nostoc* sp. PCC 7120 (Ionescu *et al*., 2010) (Fig. 1A). NsiR1 is expressed specifically in cells becoming heterocysts before any sign of morphological differentiation is observed (Muro-Pastor, 2014). Transcription of the different copies of NsiR1 requires HetR (Ionescu *et al*., 2010) and takes place from individual promoters that contain the DIF1 sequence motif associated with heterocyst-specific transcription (Mitschke *et al*., 2011; Brenes-Álvarez *et al*., 2019). It is currently unknown whether the transcription of NsiR1 is regulated by HetR directly or indirectly. The central seven copies *nsiR1.3* to *nsiR1.9* (hereafter referred to as *nsiR1.4*) are identical whereas the other copies, located at the periphery of the array, have some sequence divergence. According to the position of the TSS of *hetF* (Ionescu *et al*., 2010; Mitschke *et al*., 2011), *nsiR1.1* overlaps the *hetF* 5′ UTR. NsiR1.1 could, therefore, function as an antisense of the *hetF* transcript. As the seven identical copies are very similar to *nsiR1.1*, they could also potentially regulate the expression of *hetF*. In this work, we show that *Nostoc* strains constitutively overexpressing *nsiR1.1* or *nsiR1.4* have distinct phenotypes related to heterocyst differentiation, indicating that they both regulate *hetF*, although they could have additional non-overlapping targets.

Results

NsiR1 interacts with the 5′ UTR of hetF

NsiR1.1 is transcribed antisense to the 257 nt long 5′ UTR of *hetF* in *Nostoc* sp. PCC 7120 and this arrangement is conserved among genomes of heterocystous cyanobacteria (Ionescu *et al*., 2010) (Figs S1 and S2),

therefore, we hypothesized that NsiR1.1 might regulate *hetF* expression. If NsiR1 binding to *hetF* 5′ UTR regulates *hetF* expression, we would expect different accumulation of HetF protein in heterocysts versus vegetative cells because *nsiR1* transcription is heterocyst-specific (Ionescu *et al*., 2010; Brenes-Álvarez *et al*., 2019). To test this hypothesis, we constructed *Nostoc* sp. PCC 7120 strains that express, under the control of a constitutive promoter (*rnpB*), a translational fusion of the *hetF* 5′ UTR plus 60 nucleotides of its coding region to *gfpmut2* (plasmid pMBA93, see a scheme in Fig. 1A). A strain expressing *gfpmut2* and its original 5′UTR from the *rnpB* promoter was constructed as a control (plasmid pMBA92, Fig. 1A). As expected, the strain bearing the control plasmid pMBA92 showed similar green fluorescence in vegetative cells or heterocysts (Fig. 1B). However, the strain bearing pMBA93 showed detectable green fluorescence in vegetative cells but not in heterocysts (Fig. 1B), suggesting heterocyst-specific downregulation at the level of the *hetF* 5′ UTR.

We then tested the interaction of NsiR1.1 with *hetF* 5′ UTR using a heterologous reporter assay in *E. coli* (Urban and Vogel, 2007; Corcoran *et al*., 2012) (see Materials and Methods section for details) in which the 5′ UTR of *hetF* plus 60 nucleotides of its coding region was fused to the gene sf*gfp* encoding superfolder GFP (sfGFP). Because identical copies (*nsiR1.3* to *nsiR1.9*) are very similar to *nsiR1.1* (Ionescu *et al*., 2010), we also tested the interaction between the *hetF* 5′ UTR and the sRNA encoded in one of the identical copies, namely *nsiR1.4* (Fig. S3). We therefore co-expressed in *E. coli* the *hetF*:: sf*gfp* fusion together with NsiR1.1, NsiR1.4 or a control non-sense RNA corresponding to the unrelated terminator of the *rrnB* gene of *E. coli* (transcribed from plasmid pJV300, see Materials and Methods section). For this experiment, both the *hetF*::sf*gfp* translational fusion and the different RNAs are transcribed constitutively. Cells carrying the *hetF*::sf*gfp* translational fusion showed significant fluorescence, indicating that the translational initiation region of *hetF* was functional in *E. coli* (Fig. 1C and D). The fluorescence of cells carrying the *hetF*::sf*gfp* fusion decreased about 70% or 40% of the control when the fusion was co-expressed with NsiR1.1 or NsiR1.4 respectively, suggesting that each of these versions of NsiR1 was able to interact with the 5′ UTR of *hetF* (Fig. 1C and D) when transcribed in a heterologous system.

Taken together, the results obtained in *Nostoc* sp. PCC 7120 and *E. coli* suggested a post-transcriptional negative regulation exerted by NsiR1 on *hetF* expression. The overlapping of NsiR1.1 and the *hetF* 5′ UTR (Fig. 1A) extends between positions 20 and 78 of the UTR, 180 nucleotides upstream the start codon. To analyse the possible regulatory mechanism, we carried out *in vitro* footprinting assays of the interaction between *in vitro*
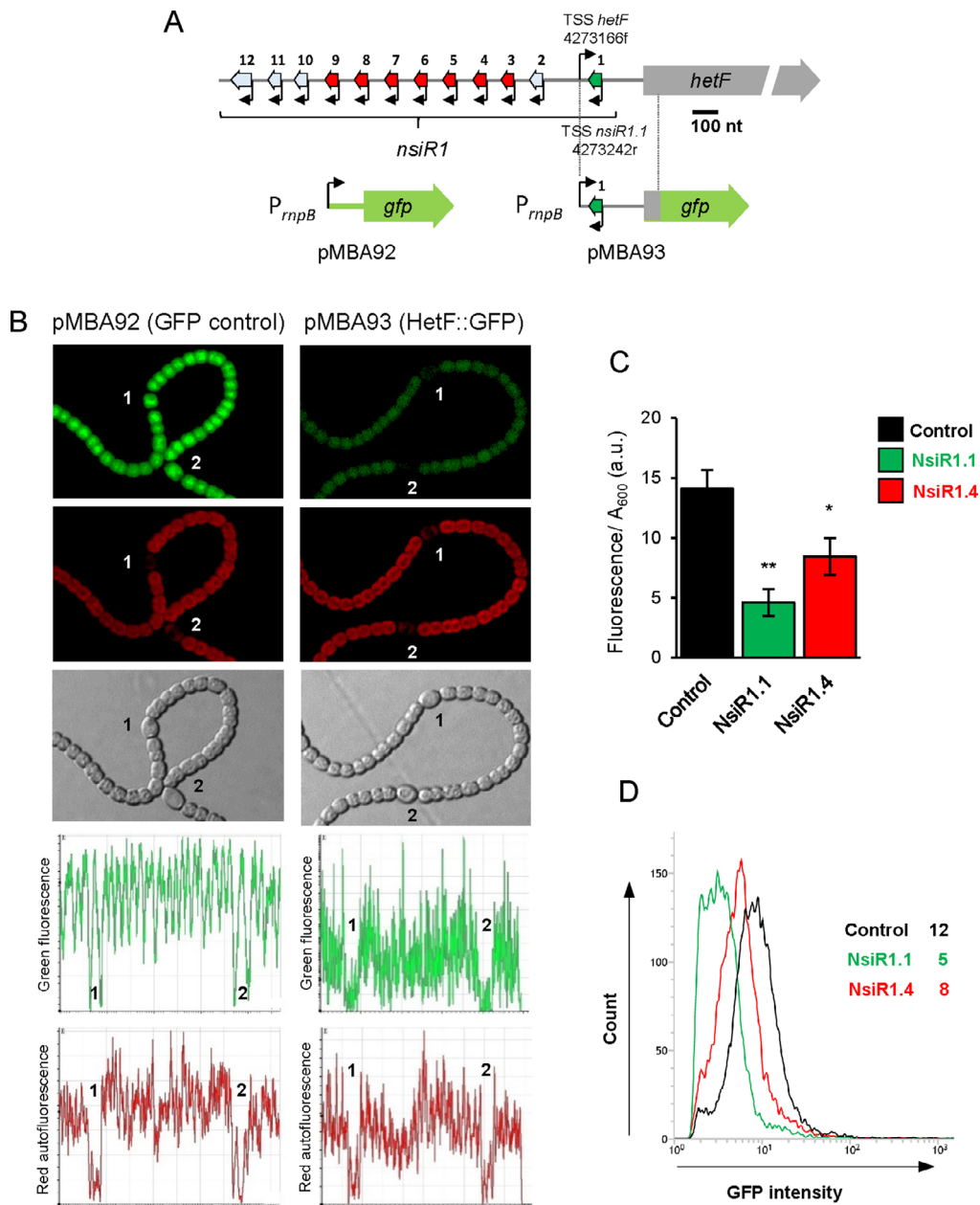
Fig 1. NsiR1 interacts with *hetF* 5′ UTR.
A. Schematic representation of the genomic region encoding NsiR1 in *Nostoc* sp. PCC 7120 and constructs in plasmids pMBA92 and pMBA93. Bent arrows indicate the transcriptional start sites. Identical repeats of *nsiR1* (3–9, e.g. *nsiR1.4*) are shown as red arrows. Repeats of *nsiR1* with some sequence divergence (1, 2, 10, 11 and 12) are shown in light blue except repeat 1 (*nsiR1.1*), antisense to *hetF*, that is shown in green. B. Top, confocal fluorescence images (green channel, red channel and bright field) of filaments from *Nostoc* strains bearing plasmids pMBA92 (control, *gfp* gene expressed from the *rnpB* promoter) or pMBA93 (translational fusion, expressed from the *rnpB* promoter, of the *hetF* 5′ UTR to the *gfp* gene) growing on top of nitrogen-free medium. Bottom, quantification of the signals for the green (GFP) and red (autofluorescence) channels. The loss of autofluorescence from photosynthetic pigments is a characteristic feature of heterocyst differentiation. Heterocysts are indicated with numbers. C. GFP fluorescence of *E. coli* cultures bearing different combinations of a plasmid expressing 5′ UTR *hetF*::sf*gfp* with plasmids expressing NsiR1.1, NsiR1.4 or an unrelated control RNA (pJV300, see Materials and Methods section for details). The data are represented as the mean and standard deviation of results from cultures corresponding to four independent colonies after normalization for cell density and subtraction of background fluorescence in cells bearing pXG0. *T*-test *p*-value <0.05*; <0.01**. D. FACS-based reporter assays. The plot shows the distribution of fluorescence in 10000 *E. coli* cells bearing plasmids expressing 5′ UTR *hetF*:sf*gfp* and the unrelated control RNA (black), NsiR1.1 (green) or NsiR1.4 (red). The inset shows the average fluorescence of each sample.

transcribed NsiR1.1 and *hetF* 5′ UTR (Fig. 2). ^{32}P-labelled *hetF* 5′ UTR (a fragment extending from the TSS at position −257 to +60 with respect to the start of the coding sequence) was incubated with unlabelled NsiR1.1 and probed with RNase T1, RNase A, or lead(II) acetate (Fig. 2A). A footprint was detected between positions

23 and 81 of *hetF* 5′ UTR (Fig. 2A, red bar), in agreement with the perfect antisense sequence complementary to NsiR1.1. In addition to this footprint, we also detected some structural changes when *hetF* 5′ UTR was combined with NsiR1.1; a region between positions 98 and 104 of *hetF* 5′ UTR became more sensitive to RNase T1,
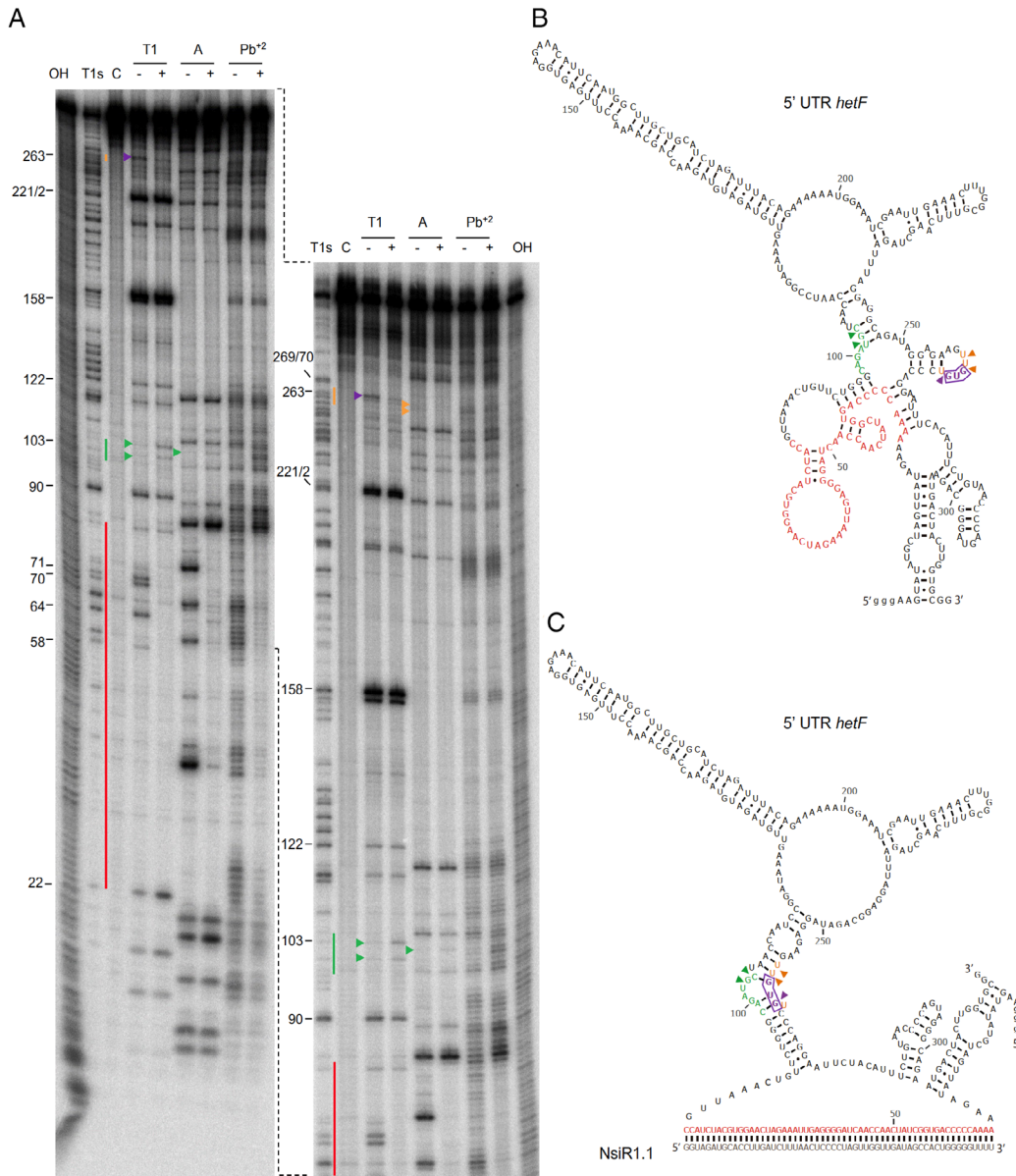

Fig 2. *In vitro* footprinting assay of the interaction between NsiR1.1 and the *hetF* 5′ UTR.
A. 5′-labelled *hetF* 5′ UTR RNA was treated with RNase T1, RNase A, or lead(II) acetate (Pb[+2]) in the absence (−) or presence (+) of NsiR1.1. Samples were loaded on two gels that were run to different lengths. Vertical bars indicate the NsiR1.1 binding site (red), a region with increased sensitivity to the probes in the presence of NsR1.1 (green) and the region including the start codon (orange). Nucleotides with differential sensitivity to RNase T1 or RNase A in the presence or absence of NsiR1.1 are indicated with coloured triangles. Nucleotide positions of *hetF* 5′ UTR are shown on the left. C, untreated control; OH, alkaline ladder; T1s, RNase T1 sequencing ladder. B, C. Models of the secondary structure of *hetF* 5′ UTR from *Nostoc* sp. PCC 7120 without (B) or with (C) bound NsiR1.1. Nucleotides corresponding to the areas highlighted by vertical bars in the gels are coloured in the same way. Nucleotides with differential sensitivity to RNase T1 or RNase A outside the NsiR1 binding site are indicated with coloured triangles. The start codon is in purple with a purple frame. The three non-encoded guanosines added at the 5′-end of *hetF* 5′UTR for efficient transcription are shown in lowercase.

RNase A and lead(II) acetate (Fig. 2A, green bar) and the region including the start codon became less accessible (Fig. 2A, orange bar). These results suggest that upon binding of NsiR1 there is a structural reorganization of the 5′ UTR of *hetF* resulting in reduced access to the initiation codon (Fig. 2B and C) that could explain the reduction of HetF expression in heterocysts.

Construction of strains with altered levels of different versions of NsiR1.

In order to analyse possible phenotypic effects caused by altered levels of NsiR1, we have constructed strains with constitutive high levels of NsiR1.1 (OE_NsiR1.1) or NsiR1.4 (OE_NsiR1.4) transcribed from the strong *trc* promoter to the *T1* terminator of the *rrnB* gene of *E. coli* (Fig. 3A). In addition, because we have introduced a NsiI restriction site (6 bp) at the 5′-end of *nsiR1.1* or *nsiR1.4* to facilitate cloning, this 6 nt tag allows the discrimination, based on its length, of molecules of NsiR1.1 and NsiR1.4 transcribed from their native promoters from molecules expressed from the *trc* promoter in the constructs introduced in *Nostoc* sp. PCC 7120.

The location of *nsiR1.1* in the 5′ UTR of *hetF* precludes the deletion of all copies of *nsiR1* without removing a segment of *hetF* 5′ UTR. As an alternative strategy to deplete NsiR1 without altering the genomic region upstream of *hetF*, we have followed a strategy similar to that in Brenes-Álvarez *et al*. (2020) and generated a *Nostoc* strain (OE_as_NsiR1) that constitutively expresses the sequence complementary to *nsiR1.4* from the *trc* promoter (Fig. 3A). This antisense RNA is expected to act as a sponge neutralizing the accumulation of *nsiR1*.

Using probes specific for NsiR1.1 or NsiR1.4 we have verified by Northern blot the constitutive expression of NsiR1.1 and NsiR1.4 under the control of the *trc* promoter in strains OE_NsiR1.1 and OE_NsiR1.4 respectively (Fig. 3B). In addition, Northern blot hybridization also showed that, in the OE_C strain, signals corresponding to hybridization to the probe for NsiR1.4 (corresponding to transcription of endogenous *nsiR1.3* to *nsiR1.9*) were stronger than those corresponding to *nsiR1.1*, as expected from the different number of identical copies of *nsiR1.4* (seven copies) versus *nsiR1.1* (single copy). Transcripts corresponding to NsiR1.1 and NsiR1.4 whose transcription did not stop at their own Rho-independent terminators but instead were terminated at the *T1* terminator were also observed above the 100 nt long size marker (indicated with black bars in Fig. 3B). Regarding the OE_as_NsiR1 strain, Northern blot hybridization showed that the accumulation of NsiR1.1 and NsiR1.3 to NsiR1.9 was diminished in this strain, suggesting effective removal of endogenous native NsiR1 by overexpression of the

antisense sponge (Fig. 3B). Both probes used for hybridization included the *T1* terminator, present also in as_NsiR1 (Fig. 3A), leading to its detection in the OE_as_NsiR1 strain (Fig. 3B).

We were unable to quantify *hetF* expression by Northern blot or RT-PCR, suggesting that, similar to the observations previously made in *Nostoc punctiforme* (Wong and Meeks, 2001), the *hetF* transcript is likely unstable or accumulates in small quantity in *Nostoc* sp. PCC 7120. We have instead analysed the accumulation of *hetF* transcripts and their putative interaction with the different versions of NsiR1 by primer extension assays (Fig. 3C). The 5′-ends of *hetF* transcripts were analysed by retrotranscription using an oligonucleotide complementary to *hetF* 5′ UTR. Extension products corresponding to the TSS of *hetF*, position 4273166f in the *Nostoc* sp. PCC 7120 genome (Ionescu et al., 2010; Mitschke et al., 2011), were in fact observed using RNA samples isolated from wild-type strain or the *hetR* mutant (Fig. S4). In addition, shorter HetR-dependent extension products were also observed around position 4173242, corresponding to the TSS of *nsiR1.1* (Fig. S4). Since, according to dRNA-Seq (Mitschke et al., 2011), there are no additional TSS for *hetF* in this region, these shorter extension products could be the result of the termination of retrotranscription at processing sites or regions with strong secondary structure.

Similar to the observation made in the wild-type strain (Fig. S4), extension products that would correspond to the full-length 5′ UTR of *hetF* were observed using RNA from the control strain OE_C as well as from strains OE_NsiR1.4 and OE_as_NsiR1 (indicated with an asterisk in Fig. 3C) but not in strain OE_NsiR1.1 (see below). Extension products that stopped at the position corresponding to the TSS of the chromosomal *nsiR1.1* were observed using RNA isolated from the OE_C, OE_NsiR1.1 and OE_NsiR1.4 strains when the cells were subjected to nitrogen deficiency (Fig. 3C, black triangle). In the OE_as_NsiR1 strain, overexpressing a sponge that is expected to remove endogenous NsiR1, these stops were only observed after 12 h of nitrogen deficiency and in reduced amounts. Retrotranscriptase stops when it encounters an obstacle, for example, regions of strong secondary structure or a double-stranded RNA such as an sRNA-mRNA duplex (Jagodnik et al., 2017). Because of the location of the additional stops observed, one possible explanation is that they are produced by binding of NsiR1. In fact, and in contrast to the other three strains included in the experiment, no full-length extension products that corresponded to the 5′-end of *hetF* were observed using RNA isolated from strain OE_NsiR1.1. Instead, a strong stop around the start position of *nsiR1.1* was observed irrespective of nitrogen removal (Fig. 3C, green triangle), suggesting that the presence of NsiR1.1 expressed from the constitutive *trc*
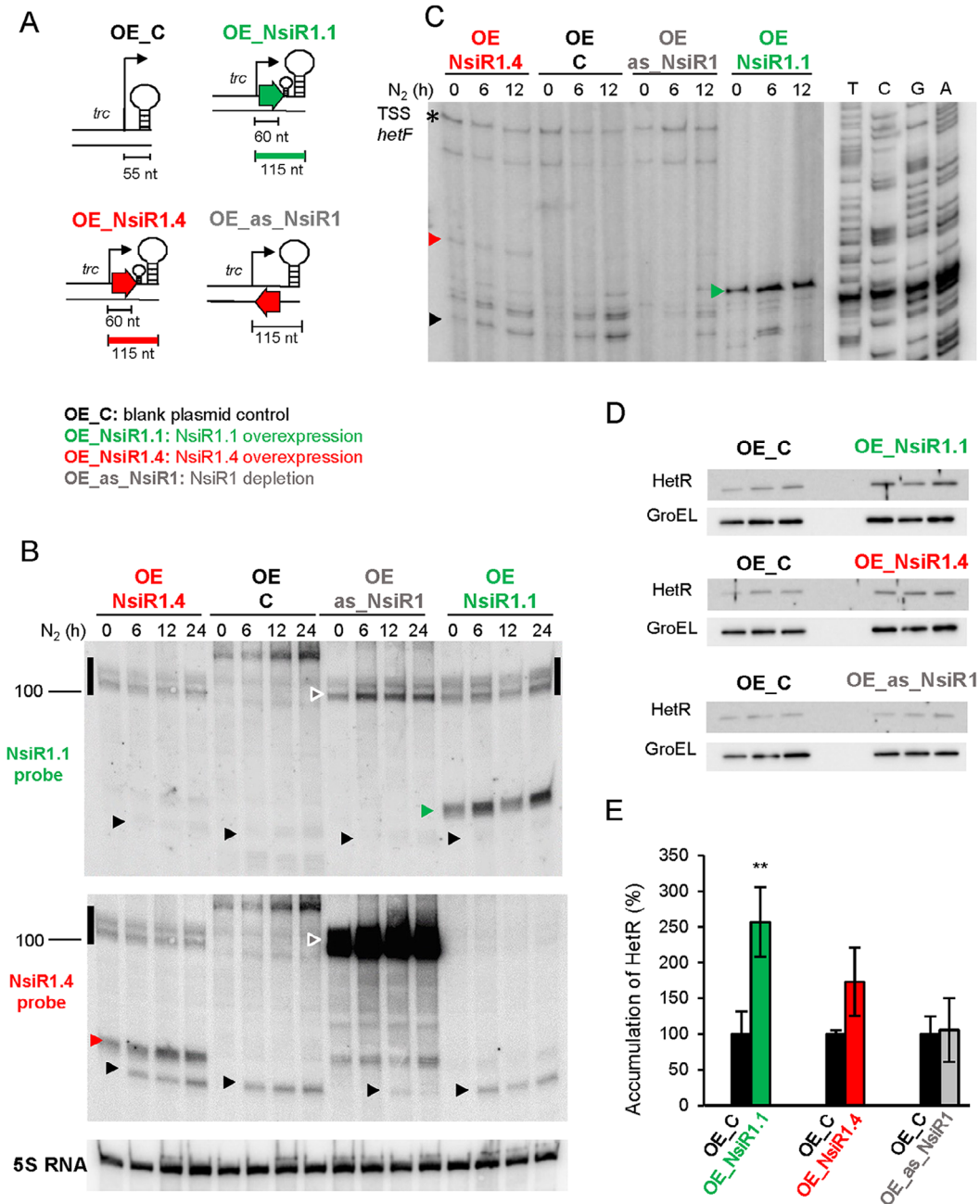
Fig 3. *Nostoc* sp. PCC 7120 strains with altered levels of NsiR1.
A. Scheme of DNA fragments cloned in the plasmids used to generate the different strains. Transcription start site (bent arrows), Rho-independent terminator of NsiR1 (small stem-loop), *T1* terminator (large stem-loop), *trc* promoter, strand-specific probes for NsiR1.1 (green line) and for NsiR1.4 (red line) are indicated. B. Northern blots using RNA extracted from the different strains at different times after nitrogen removal hybridized with probes for NsiR1.1 (top), NsiR1.4 (middle) and 5S rRNA (bottom) as a loading control. Endogenous NsiR1 (black triangles) and the three RNA species constitutively expressed from the *trc* promoter, NsiR1.1 (green triangle), NsiR1.4 (red triangle) and as_NsiR1 (grey triangles), are indicated. Black bars indicate transcripts whose transcription did not stop at the Rho-independent terminators of *nsiR1*. C. Primer extension analysis of *hetF* transcripts using RNA isolated from the mutant strains at different time points after combined nitrogen removal. The position of the TSS of *hetF* (4273166f) is indicated with an asterisk. The position corresponding to the TSS of *nsiR1.1* (4273242r) is indicated with a black triangle. Additional stops provoked by constitutive overexpression of NsiR1.1 (green triangle) and NsiR1.4 (red triangle) from the *trc* promoter are indicated. On the right, a sequencing reaction performed with the same oligonucleotide used for primer extension. D. Accumulation of HetR in the mutant strains 18 h after combined nitrogen removal. Western blots were carried out with proteins extracted from three independent cultures of each strain using antibodies against HetR or GroEL (used as a loading control). E. Quantification of Western blot shown in (D). The data are presented as the mean ± standard deviation of the HetR signal normalized to the GroEL signal in the three independent clones of each strain shown in (D). The mean of HetR accumulation of the OE_C strain in each gel is used as 100%. **T-test *p*-value <0.01.
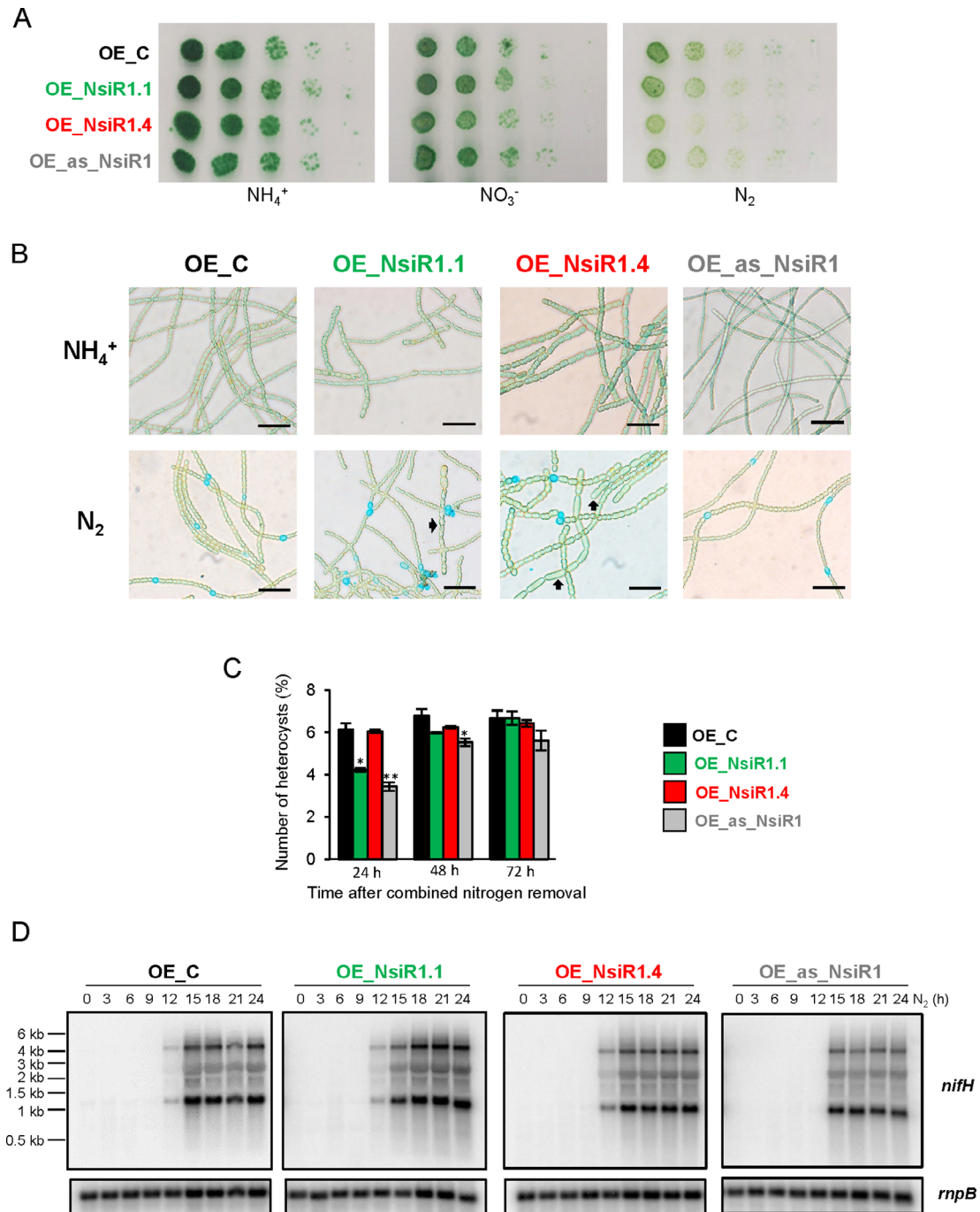
Fig 4. The phenotype of OE_C, OE_NsiR1.1, OE_NsiR1.4 and OE_as_NsiR1 strains.
A. Growth on solid media. Cells from liquid cultures of the indicated strains grown in the presence of nitrate were collected and resuspended in $BG11_0$ at an $OD_{750} = 0.17$. Fivefold serial dilutions were prepared and 10 μl of each dilution plated on $BG11_0$ plates containing ammonium (NH_4^+), nitrate (NO_3^-) or lacking combined nitrogen (N_2). Pictures were taken after 10 days of incubation at 30 °C. B. Bright-field images of filaments of OE_C, OE_NsiR1.1, OE_NsiR1.4 and OE_as_NsiR1 strains growing in liquid media containing ammonium (NH_4^+) or growing in liquid media without combined nitrogen 24 h after nitrogen removal (N_2). Filaments with enlarged irregular cells are pointed by black arrows. Heterocysts were stained with Alcian Blue. Scale bars, 25 μm. C. Frequency of heterocysts in OE_C, OE_NsiR1.1, OE_NsiR1.4 and OE_as_NsiR1. The data are presented as the mean and standard deviation of two cultures from two independent colonies. *T*-test *p*-value <0.05*; <0.01**. D. Northern blots with RNA isolated from the indicated strains at different time points (indicated in hours) after nitrogen removal and hybridized with probes for *nifH* (upper panels) and *rnpB* (bottom panels) as a loading control. Sizes are indicated on the left in kilobases.

promoter was responsible for such stops. Similarly, stops independent of the nitrogen condition but weaker than those observed with RNA isolated from strain OE_NsiR1.1 were also observed with RNA isolated from the OE_NsiR1.4 strain (Fig. 3C, red triangle). Taken together, these observations indicate that the HetR-dependent NsiR1.1 sRNA (and possibly NsiR1.4) interacts with the 5′ UTR of *hetF* in *Nostoc* sp. PCC 7120.

Based on data from fluorescence microscopy in *Nostoc* strains bearing pMBA93, the observations made in *E. coli*, and *in vitro* footprinting assays (Figs 1 and 2), NsiR1.1 and, to a lesser extent, NsiR1.4 could exert a negative regulation of *hetF* expression. Because it is known that a *hetF* mutant accumulates more HetR protein than the WT strain (Risser and Callahan, 2008), we quantified the accumulation of HetR in the OE_C, OE_NsiR1.1, OE_NsiR1.4 and OE_as_NsiR1 strains 18 h after combined nitrogen removal as an indication of HetF levels (Fig. 3D and E). Using three independent clones of each strain, we determined that strain OE_NsiR1.1 accumulated about twice more HetR protein than the OE_C strain, which could be a consequence of reduced HetF amounts in OE_NsiR1.1. OE_NsiR1.4 also seems to accumulate more HetR protein than the OE_C strain, but in this case, perhaps due to the effects of NsiR1.4 on HetF accumulation being weaker, the results were not quite statistically significant ($p = 0.057$). In order to analyse whether the increased amounts of HetR protein could be instead due to increased transcription of *hetR* in the overexpressing strains, we measured the accumulation of *hetR* transcripts in all four strains analysed above (OE_C, OE_NsiR1.1, OE_NsiR1.4 and OE_as_NsiR1) (Fig. S5). After 18 h of nitrogen deficiency, the amount of *hetR* transcript in the OE_NsiR1.1 strain was not above the amounts observed in the control strain (OE_C), indicating that the increased amount of HetR in strain OE_NsiR1.1 was not the consequence of increased transcription of the *hetR* gene.

Altered heterocyst differentiation in strains OE_NsiR1.1, OE_NsiR1.4 and OE_as_NsiR1

In order to analyse the effects of altered NsiR1 expression at the phenotypic level, we tested the growth of the OE_C, OE_NsiR1.1, OE_NsiR1.4 and OE_as_NsiR1 strains in solid media containing different nitrogen sources (Fig. 4A). All four strains grew similarly in the presence of ammonium, however, strain OE_NsiR1.4 grew worse in the absence of combined nitrogen and a bit slower in the presence of nitrate than the rest of strains. In liquid cultures, strain OE_NsiR1.4 showed slight phenotypic differences (larger cell size) in the presence of ammonium. The differences were more dramatic after combined nitrogen removal (Fig. 4B). After nitrogen

removal, strain OE_NsiR1.4 differentiated cells that were stained by Alcian blue (indicative of the presence of heterocyst polysaccharides) but some filaments had enlarged cells, which resembled the phenotype previously described for the *hetF* mutant (Risser and Callahan, 2008). OE_NsiR1.1 also contained some filaments with larger cells under nitrogen deprivation. The frequency of heterocysts in filaments of strains with altered levels of NsiR1 was determined taking samples at different time points after nitrogen removal and staining with Alcian blue. Whereas all four strains reached a similar frequency of heterocyst after 72 h of nitrogen deprivation, different frequencies were observed at an earlier time point (24 h) indicating a delay in the differentiation of heterocysts in strains OE_NsiR1.1 and OE_as_NsiR1 (Fig. 4C). This delay was longer in OE_as_NsiR1, the only strain with a statistically significant difference in frequency of heterocyst after 48 h (Fig. 4C).

Nitrogen fixation is carried out by nitrogenase, and expression of the *nifHDK* genes can be taken as an indication of heterocyst maturity. We, therefore, analysed the timing of expression of *nifH* during heterocyst differentiation upon nitrogen removal in the different strains (Fig. 4D). OE_NsiR1.1 and OE_NsiR1.4 showed no difference with respect to the control, however, OE_as_NsiR1, depleted of NsiR1, showed a delay in the accumulation of *nifH* mRNA, suggesting a delay in the differentiation of mature heterocysts.

Discussion

Heterocyst differentiation is a complex process that requires precise regulatory mechanisms in order to establish exclusive transcriptional programs for different cell types (vegetative cells, immature pro-heterocysts and mature heterocysts). Heterocyst differentiation is triggered by mutual induction of expression of the transcriptional regulator NtcA and the master regulator HetR in vegetative cells becoming heterocysts (Black *et al*., 1993; Muro-Pastor *et al*., 2002; Olmedo-Verd *et al*., 2006). The accumulation of active HetR in pro-heterocysts seems to trigger the downstream cascade. NsiR1 was the first HetR-dependent sRNA identified in *Nostoc* sp. PCC 7120 (Ionescu *et al*., 2010), but it is also well conserved in non-heterocyst-forming filamentous cyanobacteria (Brenes-Álvarez *et al*., 2016). However, while most non-heterocystous strains contain a single copy of NsiR1, heterocystous strains have multiple tandem copies with slight sequence divergence (Figs S1, S2 and S6). The increase in the number of copies of NsiR1, which correlates with the evolution of the ability to differentiate heterocysts, and their sequence divergence, could allow some degree of functional specialization by regulating a variety of targets.

According to the position of the TSS of *hetF* in *Nostoc* sp. PCC 7120 (Ionescu *et al*., 2010; Mitschke *et al*., 2011), the *nsiR1* copy closest to *hetF* (*nsiR1.1*) could function as an antisense of *hetF* 5′ UTR. This function of NsiR1.1 may be conserved in heterocyst-forming cyanobacteria because most of them have a *nsiR1* homologue in the upstream region of *hetF* in an antisense orientation (Figs S1 and S2).

Coexpression experiments in *E. coli* indicate that NsiR1.1, and to a lower extent NsiR1.4, can repress the expression of *hetF* (Fig. 1C and D). A translational fusion of *hetF* 5′ UTR to GFP had reduced expression in heterocysts (Fig. 1B), as expected if its translation is repressed by NsiR1, which is expressed at an early stage during heterocyst development (Muro-Pastor, 2014). In addition, primer extension analysis in *Nostoc* sp. PCC 7120 also suggested a putative interaction of NsiR1.1 with *hetF* 5′ UTR based on the stop of retrotranscription (Fig. S4). Several observations point to post-transcriptional regulation of translation rather than transcriptional interference as the mechanism for *hetF* regulation by NsiR1. First, negative regulation of *hetF* expression is observed in the *E. coli* system, in which NsiR1 and the *hetF* 5′ UTR are constitutively transcribed from *E. coli* promoters in separate plasmids. Second, the observation that NsiR1.1, which perfectly matches the *hetF* 5′ UTR, and NsiR1.4, which contains several nucleotide changes, repress *hetF* translation to a different extent (Fig. 1C), is indicative of post-transcriptional regulation by perfect (NsiR1.1) or imperfect (NsiR1.4) base pairing. Finally, the total quantity of *hetF* mRNA did not significantly decrease when NsiR1 was present (Fig. S4), indicating that NsiR1 did not affect *hetF* transcription.

The interaction site of NsiR1 in the *hetF* 5′ UTR was located relatively far upstream from the *hetF* start codon, precluding repression of translation by direct interaction of NsiR1 with the translation initiation region and pointing instead to a translational regulation based on a possible conformational change of the *hetF* 5′ UTR. *In vitro* footprinting data suggest that the *hetF* 5′ UTR is highly structured and the interaction with NsiR1.1 provokes a conformational change in which the region of the start codon of *hetF* becomes less accessible (Fig. 2), providing a mechanistic explanation for the negative regulation of *hetF* expression by NsiR1.1.

In order to study the function of NsiR1, we created strains of *Nostoc* sp. PCC 7120 with altered levels of different versions of NsiR1 (Fig. 3). The overexpression of NsiR1.1 or NsiR1.4 under the control of the *trc* promoter was irrespective of the nitrogen source and strong enough to overpass the endogenous expression of NsiR1, which was induced upon nitrogen removal. As expected, the OE_as_NsiR1 strain had reduced amounts

of NsiR1, implying that the interaction between NsiR1 and its antisense results in its degradation (Fig. 3B). Primer extension assays with RNA extracted from strains with altered levels of different versions of NsiR1 confirmed the interaction of NsiR1.1 and NsiR1.4 with *hetF* 5′ UTR (Fig. 3C). We were unable to quantify changes in *hetF* expression in our mutant strains, likely because the expression of *hetF* is low and the transcripts are unstable (Wong and Meeks, 2001; Risser and Callahan, 2008). Instead, because one of the features described for a *hetF* mutant is the accumulation of higher levels of HetR protein (Risser and Callahan, 2008), we analysed HetR protein levels as an indication of HetF function. In fact, strains OE_NsiR1.1 and, to a lesser extent, OE_NsiR1.4, had higher amounts of HetR, as expected for a strain with a reduced amount of HetF (Fig. 3D and E). This higher amount of HetR was not due to increased *hetR* transcription in strains OE_NsiR1.1 and OE_NsiR1.4 (Fig. S5). Finally, OE_NsiR1.1 had a phenotype partially similar in some respects to that described for a *hetF* mutant. A *hetF* mutant grows poorly in media without combined nitrogen, has enlarged cells and cannot differentiate heterocysts (Risser and Callahan, 2008). OE_NsiR1.1 showed some filaments with larger cells in liquid cultures growing without combined nitrogen (Fig. 4). Taken together, all these *in vivo* and *in vitro* results strongly support a role for NsiR1 in reducing the amount of HetF in cells becoming heterocysts. Proper HetR turnover seems to be required for heterocyst differentiation, since *hetF* mutants accumulating HetR are unable to differentiate (Risser and Callahan, 2008). Similarly, strains bearing several alleles of *hetR* that result in reduced turnover are also unable to differentiate (Risser and Callahan, 2007). Specific NsiR1 expression in developing heterocysts would reduce HetF accumulation resulting in an increased amount of HetR in these cells, favouring expression of HetR-dependent transcripts and progression of differentiation (Fig. 5). Finally, despite the small difference in sequence between NsiR1.1 and NsiR1.4 (Fig. S3), strains OE_NsiR1.1 and OE_NsiR1.4 had slightly different phenotypes. They had some growth differences in solid media with different nitrogen sources (Fig. 4A). In fact, the observation that the effect of OE_NsiR1.4 is more important than OE_NsiR1.1, while NsiR1.4 is less efficient in repressing *hetF* expression, suggests that NsiR1.1 and NsiR1.4 might have additional non-overlapping targets related to heterocyst differentiation. On the other hand, the observation that the strain OE_as_NsiR1, depleted of NsiR1, showed delayed expression of *nifH* transcripts encoding nitrogenase indicates NsiR1 influences the overall timing of heterocyst maturation.

Further study of additional potential targets for NsiR1, and the characterization of additional heterocyst-specific
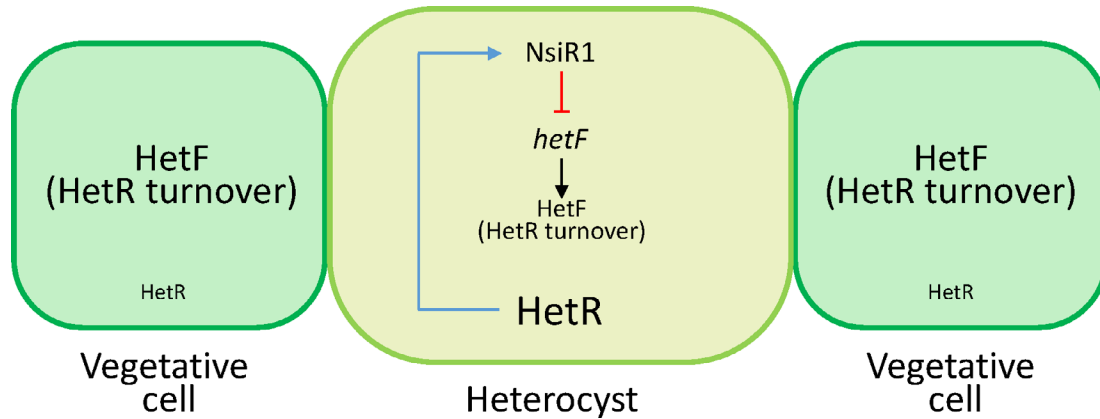
Fig 5.   Model of the function of NsiR1 in the regulation of heterocyst differentiation. Positive regulation is indicated by a blue arrow. Negative post-transcriptional regulation by NsiR1 is indicated by a red line with a blunt end. Different sizes indicate the relative amount of HetF and HetR in vegetative cells versus heterocysts.

non-coding RNAs (Brenes-Álvarez *et al.*, 2019) could reveal the importance of these types of molecules in the regulation of this developmental process.

Materials and methods

Strains and growth conditions

For Northern blot and primer extension analysis, cultures of wild-type, *hetR* mutant 216 (Buikema and Haselkorn, 1991) and different derivative strains of *Nostoc* sp. PCC 7120 (Table S1) were bubbled with an air/CO_2 mixture (1% vol/vol) and grown photoautotrophically at 30 °C in BG11 medium (Rippka *et al.*, 1979) containing ferric citrate instead of ammonium ferric citrate, lacking $NaNO_3$ but containing 6 mM NH_4Cl, 10 mM $NaHCO_3$ and 12 mM *N*-tris (hydroxymethyl) methyl-2-aminoethanesulfonic acid-NaOH buffer (pH 7.5) ($BG11_0C$ + 6 mM NH_4Cl). Nitrogen deficiency was induced by filtering, washing and resuspending cells in nitrogen-free-medium with 10 mM $NaHCO_3$ ($BG11_0C$). Chlorophyll concentration of the cultures was measured as described (MacKinney, 1941). For Western blot analysis, cultures of different derivative strains of *Nostoc* sp. PCC 7120 (Table S1) were grown in flasks at 30 °C in liquid BG11 medium for 1 week. Nitrogen deficiency was induced as described above. OE_C, OE_NsiR1.1, OE_NsiR1.4 and OE_as_NsiR1 strains were grown in the presence of streptomycin (Sm) and spectinomycin (Sp), 2 μg ml^{-1} each (liquid medium) or 5 μg ml^{-1} each (solid medium). *Nostoc* strains bearing Nm^R plasmids (pMBA92, pMBA93 or pMBA94) were grown in the presence of neomycin (Nm), 5 μg ml^{-1} (liquid medium) or 25 μg ml^{-1} (solid medium). *E. coli* strains were grown in LB medium, supplemented with appropriate antibiotics (Sambrook and Russell, 2001).

Construction of Nostoc strains with altered levels of NsiR1

Plasmids and oligonucleotides used in this work are described in Tables S2 and S3. The sequences of all fragments amplified by PCR were entirely verified by sequencing. We have used pMBA37 (Olmedo-Verd *et al*., 2019) as a backbone for overexpressing different copies of *nsiR1* or an antisense to *nsiR1*. pMBA37 contains the *trc* promoter and the *T1* terminator from *E. coli rrnB* gene. The *trc* promoter from pMBA37 was amplified using oligonucleotides 573 and 574. This product and oligonucleotide 573 were later used for PCRs with oligonucleotides 757, 758 or 512 generating fragments in which the *trc* promoter was fused respectively to *nsiR1.1* (fragment 1), *nsiR1.4* (fragment 2) or *as_nsiR1* (fragment 3). After digestion of fragments 1–3 with ClaI and XhoI in the sites provided by oligonucleotides 573, 757, 758 and 512, these fragments were cloned into ClaI-XhoI digested pMBA37 vector, rendering respectively pMBA77, pMBA78 and pMBA42 (Table S2). pMBA77, pMBA78 and pMBA42 were introduced in *Nostoc* sp. PCC 7120 wild type by conjugation (Elhai and Wolk, 1988) generating strains OE_NsiR1.1, OE_NsiR1.4 and OE_as_NsiR1 respectively (Table S1).

Construction of Nostoc strain with 5′UTR of hetF fused to gfp

We have used pMBA91, a derivative of pMBA20 (Olmedo-Verd *et al*., 2019), as a backbone for the constitutive overexpression from *rnpB* promoter of a translational fusion to GFP of the 5′UTR of *hetF*. As a control, the *gfpmut2* gene was cloned with its own 5′UTR. pMBA91 is similar to pMBA20 but with an Nm resistance gene instead of an SmSp resistance cassette. The Nm

resistance gene from pRL278 (Black *et al.*, 1993) was amplified using oligonucleotides 892 + 893. After digestion with BamHI, this product was cloned in BamHI-digested pMBA20, rendering pMBA91 (Table S2). The *hetF* 5′UTR was amplified using as template genomic DNA and oligonucleotides 555 + 895. *gfpmut2* gene was obtained by PCR using as template pSAM270 (Brenes-Álvarez *et al.*, 2019) and oligonucleotides 896 + 894. A PCR product with a translational fusion between the *hetF* 5′UTR plus sequences encoding the first 20 amino acids of HetF and *gfpmut2* was amplified using as templates the two above-described PCR fragments and oligonucleotides 555 + 894. The PCR product was digested with NsiI and XhoI and cloned in NsiI-XhoI digested pMBA91, rendering pMBA93. The 5′UTR of *gfpmut2* was used as control. The *gfpmut2* gene plus its 5′UTR were amplified using as template pSAM270 (Brenes-Álvarez *et al.*, 2019) and oligonucleotides 894 + 899. The PCR product was digested with NsiI and XhoI and cloned in NsiI-XhoI digested pMBA91, rendering pMBA92 (Table S2). pMBA92 and pMBA93 were introduced in *Nostoc* sp. PCC 7120 by conjugation (Elhai and Wolk, 1988). Sequences of inserts in plasmids expressing GFP translational fusions in *Nostoc* sp. PCC 7120 are shown in Table S4.

In vitro *synthesis and labelling of RNA*

RNA transcripts were generated *in vitro* with MEGAscript High Yield Transcription Kit (AM1333, Ambion). The DNA templates for the transcription of *hetF* 5′UTR, NsiR1.1 and NsiR1.4 were generated by PCR with a primer that includes a T7 promoter sequence upstream the 5′-end of the coded RNA, and a primer matching the 3′-end of the RNA (Table S3). The templates used for these PCR amplifications were genomic DNA (for *hetF* 5′UTR), or only the overlapping oligonucleotides (for NsiR1.1 and NsiR1.4). The sequences of the RNAs transcribed *in vitro* are shown in Table S5. Three (*hetF*) or one (*nsiR1*) non-encoded guanosines were added at the 5′-end of each template for efficient T7 transcription. After *in vitro* transcription, RNAs were treated with DNase I and purified by phenol, phenol/chloroform and chloroform extractions, ethanol-precipitated at −20 °C, and washed with 70% ethanol. 100 pmol of RNAs were labelled at the 5′-end with [γ^{32}P]ATP and polynucleotide kinase and purified in a 6% polyacrylamide gel as described (Álvarez-Escribano *et al.*, 2018).

In vitro *structure probing and footprinting*

For structure probing of NsiR1.1 and NsiR1.4, 0.1 pmol of labelled NsiR1.1 and NsiR1.4 RNAs in a 7 μl volume were denatured for 1 min at 95 °C and chilled on ice for

5 min. For the footprinting of NsiR1.1 on *hetF* 5′UTR, 0.1 pmol of labelled *hetF* 5′UTR RNA was mixed in 7 μl with water (−) or 1 pmol NsiR1.1 (+), denatured for 1 min at 95 °C and chilled on ice for 5 min. After denaturing and chilling steps, we added to all samples 1 μl of 1 mg ml^{-1} yeast RNA (Ambion AM7118) and 1 μl of 10× structure buffer (Ambion). The samples were incubated further for 15 min at 37 °C. RNase T1, RNase A and lead(II) acetate treatment were performed as described (Álvarez-Escribano *et al.*, 2018). Alkaline and RNase T1 G ladders were generated as described (Álvarez-Escribano *et al.*, 2018). All samples were run on 6% polyacrylamide, 7 M urea gels and bands visualized with a Cyclone Storage Phosphor System.

RNA isolation, Northern blot analysis and primer extension assays

Total RNA was isolated using hot phenol as described (Mohamed and Jansson, 1989) with modifications (Brenes-Álvarez *et al.*, 2016). Samples for Northern blot hybridization of NsiR1 were separated in 10% urea-polyacrylamide gels as described (Steglich *et al.*, 2008) (7 μg of total RNA). Samples for Northern blot hybridization of longer transcripts (*hetR* or *nifH*) were separated in 1% agarose denaturing formaldehyde gels (10 μg of total RNA) and transferred to Hybond-N+ membrane (GE Healthcare) with 20× SSC buffer. Strand-specific ^{32}P-labelled probes were prepared with Taq DNA polymerase using a PCR fragment as a template and the oligonucleotides specified in Table S3 in a reaction with [α-^{32}P]dCTP and one single oligonucleotide as a primer (corresponding to the complementary strand of the sRNA or mRNA to be detected). PCR fragments for NsiR1.1 and NsiR1.4 probes were amplified with oligonucleotides 506 + 573 and pMBA77 (NsiR1.1) or pMBA78 (NsiR1.4) as templates. PCR fragments for *hetR* and *nifH* probes were amplified from genomic DNA using oligonucleotides hetR8 + hetR9 and 787 + 788 respectively. Hybridization to *rnpB* (Vioque, 1992) or 5S rRNA probes was used as a loading and transfer control. Primer extension of 5′ ends of *hetF* was performed as previously described (Muro-Pastor *et al.*, 1999) using 10 μg of total RNA and oligonucleotide 105 labelled with [γ-^{32}P]ATP.

Western blot analysis

For HetR Western blot analysis, crude extracts were prepared using glass beads. Cells from 25 ml cultures of OE_C, OE_NsiR1.1, OE_NsiR1.4 and OE_as_NsiR1 subjected to 18 h of nitrogen deficiency were harvested by filtration, washed with TE50/100 buffer (50 mM Tris pH 8, 100 mM EDTA) and resuspended in 500 μl of

resuspension buffer (50 mM Tris–HCl pH 8.0, 2 mM 2-mercaptoethanol) containing a protease inhibitor cocktail (Roche). The cellular suspension was mixed with glass beads (SIGMA, 200 μm) in an Eppendorf tube and subjected to 7 cycles of 1 min vortexing plus 1 min of cooling on ice. Cell extract was separated from cell debris and unbroken cells by centrifugation (3 min at 3000g at 4 °C). The soluble fraction was obtained by centrifugation of the crude extract at 16.000g for 30 min at 4 °C. The protein concentration was determined by the Lowry procedure (Markwell *et al.*, 1978). Forty micrograms of proteins were fractionated on 10% SDS-PAGE. Antibodies against *Nostoc* sp. PCC 7120 HetR (Valladares *et al.*, 2016) and *E. coli* GroEL (Sigma-Aldrich) were used.

Reporter assay for *in vivo* verification of targets

We used the GFP reporter system described in Urban and Vogel (2007) with the sfGFP and plasmid pXG10-SF (Corcoran *et al.*, 2012) for the experimental verification of targets in *E. coli*. The 5′UTR of *hetF* from the TSS at position −257 (4273166f) plus 60 nucleotides within its coding region was amplified using oligonucleotides 555 + 557. The GTG start codon was changed to ATG using overlapping PCRs and oligonucleotides 555, 556, 557 and 558. The information about TSS was taken from Mitschke and colleagues (2011). To express NsiR1.1 and NsiR1.4 in *E. coli*, the genes were amplified from genomic DNA using primers 559 (5′ phosphorylated) + 561 or 560 (5′ phosphorylated) + 562 respectively. These PCR products were digested with XbaI and fused to a plasmid backbone that was amplified from pZE12-luc with primers PLlacOB and PLlacOD and digested with XbaI, rendering pMML3 and pMML4 (Table S2). Different combinations of plasmids were introduced in *E. coli* DH5α and GFP fluorescence measurements were carried out by either flow cytometry or with a microplate reader (Varioskan) using liquid cultures from individual colonies, as previously described (Wright *et al.*, 2013). Plasmid pJV300 (Table S2) was used as a control. Sequences of inserts from plasmids used in the heterologous reporter system are shown in Table S6.

Fluorescence microscopy

Fluorescence of *Nostoc* sp. PCC 7120 filaments carrying plasmid pMBA92, pMBA93 or pMBA94 (Table S2) growing on top of solidified nitrogen-free medium was analysed and quantified as described (Muro-Pastor, 2014) using a Leica HCX PLAN-APO 63× 1.4 NA oil immersion objective attached to a Leica TCS SP2 laser-scanning confocal microscope. Samples were excited at 488 nm by an argon-ion laser and the fluorescent emission was monitored by collection across windows of 500–538 nm (GFP) and 630–700 nm (photosynthetic pigment autofluorescence).

Filaments were stained with Alcian blue as described (Olmedo-Verd *et al.*, 2005) and visualized using an Olympus BX60 microscope. The frequency of heterocysts was calculated as the number of cells stained by Alcian blue per at least 500 individual cells counted. All results are expressed as the average of two biological replicates ± SD.

Computational and statistical methods

To detect *nsiR1* homologues in the upstream region of *hetF*, the HetF sequence of *Nostoc* sp. PCC 7120 was used in a blastp search in IMG (Markowitz *et al.*, 2014) against 100 cyanobacterial genomes, those used in Shih and colleagues (2013), excluding the genomes corresponding to the fast-evolving *Prochlorococcus* clade. Fifty HetF homologues were identified. We then used one of the identical repeats of *nsiR1* in *Nostoc* sp. PCC 7120 as input in a blastn search to predict putative *nsiR1* homologues in the region (4 kb) upstream each of the 50 predicted *hetF* homologues. Sequences of *nsiR1* homologues that were not in the upstream region of *hetF* were taken from Brenes-Álvarez and colleagues (2016). All the NsiR1 homologues identified are shown aligned in Fig. S6. The Student's *t*-test was used to determine statistical significance. The number of biological samples can be found in the corresponding figure legends. Models of RNA secondary structures were generated with Mfold (Zuker, 2003) implementing restrictions according to the information obtained from *in vitro* structure probing.

Author Contributions

M.B.-A., A.V., and A.M.M.-P. designed research; M.B.-A., M.M., A.V., and A.M.M.-P. performed research; M.B.-A., A.V., and A.M.M.-P. analysed data; and M.B.-A., A.V., and A.M.M.-P. wrote the manuscript.

References

Álvarez-Escribano, I., Vioque, A., and Muro-Pastor, A.M. (2018) NsrR1, a nitrogen stress-repressed sRNA, contributes to the regulation of *nblA* in *Nostoc* sp. PCC 7120. *Front Microbiol* **9**: 2267.

Black, T.A., Cai, Y., and Wolk, C.P. (1993) Spatial expression and autoregulation of *hetR*, a gene involved in the control of heterocyst development in *Anabaena*. *Mol Microbiol* **9**: 77–84.

Brenes-Álvarez, M., Mitschke, J., Olmedo-Verd, E., Georg, J., Hess, W.R., Vioque, A., and Muro-Pastor, A.M. (2019) Elements of the heterocyst-specific transcriptome unravelled by co-expression analysis in *Nostoc* sp. PCC 7120. *Environ Microbiol* **21**: 2544–2558.

Brenes-Álvarez, M., Olmedo-Verd, E., Vioque, A., and Muro-Pastor, A.M. (2016) Identification of conserved and potentially regulatory small RNAs in heterocystous cyanobacteria. *Front Microbiol* **7**: 48.

Brenes-Álvarez, M., Vioque, A., and Muro-Pastor, A.M. (2020) The integrity of the cell wall and its remodeling during heterocyst differentiation are regulated by phylogenetically conserved small RNA Yfr1 in *Nostoc* sp. strain PCC 7120. *mBio* **1**: e02599-02519.

Buikema, W.J., and Haselkorn, R. (1991) Characterization of a gene controlling heterocyst differentiation in the cyanobacterium *Anabaena* 7120. *Genes Dev* **5**: 321–330.

Callahan, S.M., and Buikema, W.J. (2001) The role of HetN in maintenance of the heterocyst pattern in *Anabaena* sp. PCC 7120. *Mol Microbiol* **40**: 941–950.

Corcoran, C.P., Podkaminski, D., Papenfort, K., Urban, J.H., Hinton, J.C., and Vogel, J. (2012) Superfolder GFP reporters validate diverse new mRNA targets of the classic porin regulator, MicF RNA. *Mol Microbiol* **84**: 428–445.

Elhai, J., and Khudyakov, I. (2018) Ancient association of cyanobacterial multicellularity with the regulator HetR and an RGSGR pentapeptide-containing protein (PatX). *Mol Microbiol* **110**: 931–954.

Elhai, J., and Wolk, C.P. (1988) Conjugal transfer of DNA to cyanobacteria. *Methods Enzymol* **167**: 747–754.

Flores, E., and Herrero, A. (2010) Compartmentalized function through cell differentiation in filamentous cyanobacteria. *Nat Rev Microbiol* **8**: 39–50.

Flores, E., Picossi, S., Valladares, A., and Herrero, A. (2019) Transcriptional regulation of development in heterocyst-forming cyanobacteria. *Biochim Biophys Acta Gene Regul Mech* **1862**: 673–684.

Herrero, A., Muro-Pastor, A.M., and Flores, E. (2001) Nitrogen control in cyanobacteria. *J Bacteriol* **183**: 411–425.

Herrero, A., Stavans, J., and Flores, E. (2016) The multicellular nature of filamentous heterocyst-forming cyanobacteria. *FEMS Microbiol Rev* **40**: 831–854.

Huang, X., Dong, Y., and Zhao, J. (2004) HetR homodimer is a DNA-binding protein required for heterocyst differentiation, and the DNA-binding activity is inhibited by PatS. *Proc Natl Acad Sci U S A* **101**: 4848–4853.

Ionescu, D., Voβ, B., Oren, A., Hess, W.R., and Muro-Pastor, A.M. (2010) Heterocyst-specific transcription of NsiR1, a non-coding RNA encoded in a tandem array of direct repeats in cyanobacteria. *J Mol Biol* **398**: 177–188.

Jagodnik, J., Brosse, A., Le Lam, T.N., Chiaruttini, C., and Guillier, M. (2017) Mechanistic study of base-pairing small regulatory RNAs in bacteria. *Methods* **117**: 67–76.

Kopf, M., and Hess, W.R. (2015) Regulatory RNAs in photosynthetic cyanobacteria. *FEMS Microbiol Rev* **39**: 301–315.

MacKinney, G. (1941) Absorption of light by chlorophyll solutions. *J Biol Chem* **140**: 315–322.

Marchais, A., Duperrier, S., Durand, S., Gautheret, D., and Stragier, P. (2011) CsfG, a sporulation-specific, small noncoding RNA highly conserved in endospore formers. *RNA Biol* **8**: 358–364.

Markowitz, V.M., Chen, I.M., Palaniappan, K., Chu, K., Szeto, E., Pillay, M., *et al*. (2014) IMG 4 version of the integrated microbial genomes comparative analysis system. *Nucleic Acids Res* **42**: D560–D567.

Markwell, M.A., Haas, S.M., Bieber, L.L., and Tolbert, N.E. (1978) A modification of the Lowry procedure to simplify protein determination in membrane and lipoprotein samples. *Anal Biochem* **87**: 206–2010.

Mars, R.A., Nicolas, P., Denham, E.L., and van Dijl, J.M. (2016) Regulatory RNAs in *Bacillus subtilis*: a Gram-positive perspective on bacterial RNA-mediated regulation of gene expression. *Microbiol Mol Biol Rev* **80**: 1029–1057.

Mitschke, J., Vioque, A., Haas, F., Hess, W.R., and Muro-Pastor, A.M. (2011) Dynamics of transcriptional start site selection during nitrogen stress-induced cell differentiation in *Anabaena* sp. PCC 7120. *Proc Natl Acad Sci U S A* **108**: 20130–20135.

Mohamed, A., and Jansson, C. (1989) Influence of light on accumulation of photosynthesis-specific transcripts in the cyanobacterium *Synechocystis* 6803. *Plant Mol Biol* **13**: 693–700.

Muro-Pastor, A.M. (2014) The heterocyst-specific NsiR1 small RNA is an early marker of cell differentiation in cyanobacterial filaments. *MBio* **5**: e01079-01014.

Muro-Pastor, A.M., and Hess, W.R. (2012) Heterocyst differentiation: from single mutants to global approaches. *Trends Microbiol* **20**: 548–557.

Muro-Pastor, A.M., and Hess, W.R. (2020) Regulatory RNA at the crossroads of carbon and nitrogen metabolism in photosynthetic cyanobacteria. *Biochim Biophys Acta Gene Regul Mech* **1863**: 194477.

Muro-Pastor, A.M., Valladares, A., Flores, E., and Herrero, A. (1999) The *hetC* gene is a direct target of the NtcA transcriptional regulator in cyanobacterial heterocyst development. *J Bacteriol* **181**: 6664–6669.

Muro-Pastor, A.M., Valladares, A., Flores, E., and Herrero, A. (2002) Mutual dependence of the expression of the cell differentiation regulatory protein HetR and the global nitrogen regulator NtcA during heterocyst development. *Mol Microbiol* **44**: 1377–1385.

Olmedo-Verd, E., Brenes-Álvarez, M., Vioque, A., and Muro-Pastor, A.M. (2019) A heterocyst-specific antisense RNA contributes to metabolic reprogramming in *Nostoc* sp. PCC 7120. *Plant Cell Physiol* **60**: 1646–1655.

Olmedo-Verd, E., Flores, E., Herrero, A., and Muro-Pastor, A.M. (2005) HetR-dependent and -independent expression of heterocyst-related genes in an *Anabaena* strain overproducing the NtcA transcription factor. *J Bacteriol* **187**: 1985–1991.

Olmedo-Verd, E., Muro-Pastor, A.M., Flores, E., and Herrero, A. (2006) Localized induction of the *ntcA* regulatory gene in developing heterocysts of *Anabaena* sp. strain PCC 7120. *J Bacteriol* **188**: 6694–6699.

Rajagopalan, R., and Callahan, S.M. (2010) Temporal and spatial regulation of the four transcription start sites of *hetR* from *Anabaena* sp. strain PCC 7120. *J Bacteriol* **192**: 1088–1096.

Rippka, R., Deruelles, J., Waterbury, J.B., Herdman, M., and Stanier, R.Y. (1979) Generic assignments, strain histories and properties of pure cultures of cyanobacteria. *Microbiology* **111**: 1–61.

Risser, D.D., and Callahan, S.M. (2007) Mutagenesis of *hetR* reveals amino acids necessary for HetR function in the heterocystous cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* **189**: 2460–2467.

Risser, D.D., and Callahan, S.M. (2008) HetF and PatA control levels of HetR in *Anabaena* sp. strain PCC 7120. *J Bacteriol* **190**: 7645–7654.

Risser, D.D., and Callahan, S.M. (2009) Genetic and cytological evidence that heterocyst patterning is regulated by inhibitor gradients that promote activator decay. *Proc Natl Acad Sci U S A* **106**: 19884–19888.

Sambrook, J., and Russell, D.W. (2001) *Molecular Cloning: A Laboratory Manual*. New York: Cold Spring Harbor Laboratory.

Schmalisch, M., Maiques, E., Nikolov, L., Camp, A.H., Chevreux, B., Muffler, A., *et al*. (2010) Small genes under sporulation control in the *Bacillus subtilis* genome. *J Bacteriol* **192**: 5402–5412.

Shih, P.M., Wu, D., Latifi, A., Axen, S.D., Fewer, D.P., Talla, E., *et al*. (2013) Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. *Proc Natl Acad Sci U S A* **110**: 1053–1058.

Silvaggi, J.M., Perkins, J.B., and Losick, R. (2006) Genes for small, noncoding RNAs under sporulation control in *Bacillus subtilis*. *J Bacteriol* **188**: 532–541.

Steglich, C., Futschik, M.E., Lindell, D., Voss, B., Chisholm, S.W., and Hess, W.R. (2008) The challenge of regulation in a minimal photoautotroph: non-coding RNAs in *Prochlorococcus*. *PLoS Genet* **4**: e1000173.

Urban, J.H., and Vogel, J. (2007) Translational control and target recognition by *Escherichia coli* small RNAs in vivo. *Nucleic Acids Res* **35**: 1018–1037.

Valladares, A., Flores, E., and Herrero, A. (2016) The heterocyst differentiation transcriptional regulator HetR of the filamentous cyanobacterium *Anabaena* forms tetramers and can be regulated by phosphorylation. *Mol Microbiol* **99**: 808–819.

Vioque, A. (1992) Analysis of the gene encoding the RNA subunit of ribonuclease P from cyanobacteria. *Nucleic Acids Res* **20**: 6331–6337.

Wagner, E.G.H., and Romby, P. (2015) Small RNAs in bacteria and archaea: who they are, what they do, and how they do it. *Adv Genet* **90**: 133–208.

Wong, F.C., and Meeks, J.C. (2001) The *hetF* gene product is essential to heterocyst differentiation and affects HetR function in the cyanobacterium *Nostoc punctiforme*. *J Bacteriol* **183**: 2654–2661.

Wright, P.R., Richter, A.S., Papenfort, K., Mann, M., Vogel, J., Hess, W.R., *et al*. (2013) Comparative genomics boosts target prediction for bacterial small RNAs. *Proc Natl Acad Sci U S A* **110**: E3487–E3496.

Yoon, H.S., and Golden, J.W. (1998) Heterocyst pattern formation controlled by a diffusible peptide. *Science* **282**: 935–938.

Zhou, R., Wei, X., Jiang, N., Li, H., Dong, Y., Hsi, K.L., and Zhao, J. (1998) Evidence that HetR protein is an unusual serine-type protease. *Proc Natl Acad Sci U S A* **95**: 4959–4963.

Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* **31**: 3406–3415.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Appendix S1. Supporting pdf file, including supplementary Figs S1 to S6 and supplementary Tables S1 to S6.