

# Nitrogen-regulated antisense transcription in the adaptation to nitrogen deficiency in *Nostoc* sp. PCC 7120

Manuel Brenes-Álvarez <sup>1</sup>, Agustín Vioque  and Alicia M. Muro-Pastor <sup>\*</sup>

Instituto de Bioquímica Vegetal y Fotosíntesis, Consejo Superior de Investigaciones Científicas and Universidad de Sevilla, Américo Vespucio 49, 41092 Sevilla, Spain

<sup>1</sup>Present address: Genetics and Experimental Bioinformatics, Faculty of Biology, University of Freiburg, 79104 Freiburg, Germany.

<sup>\*</sup>To whom correspondence should be addressed: Email: [alicia@ibvf.csic.es](mailto:alicia@ibvf.csic.es)

Edited By: Marena Wilson-Pham

## Abstract

Transcriptomic analyses using high-throughput methods have revealed abundant antisense transcription in bacteria. Antisense transcription is often due to the overlap of mRNAs with long 5' or 3' regions that extend beyond the coding sequence. In addition, antisense RNAs that do not contain any coding sequence are also observed. *Nostoc* sp. PCC 7120 is a filamentous cyanobacterium that, under nitrogen limitation, behaves as a multicellular organism with division of labor among two different cell types that depend on each other, the vegetative CO<sub>2</sub>-fixing cells and the nitrogen-fixing heterocysts. The differentiation of heterocysts depends on the global nitrogen regulator NtcA and requires the specific regulator HetR. To identify antisense RNAs potentially involved in heterocyst differentiation, we assembled the *Nostoc* transcriptome using RNA-seq analysis of cells subjected to nitrogen limitation (9 or 24 h after nitrogen removal) in combination with a genome-wide set of transcriptional start sites and a prediction of transcriptional terminators. Our analysis resulted in the definition of a transcriptional map that includes >4,000 transcripts, 65% of which contain regions in antisense orientation to other transcripts. In addition to overlapping mRNAs, we identified nitrogen-regulated noncoding antisense RNAs transcribed from NtcA- or HetR-dependent promoters. As an example of this last category, we further analyzed an antisense (*as\_gltA*) of the gene-encoding citrate synthase and showed that transcription of *as\_gltA* takes place specifically in heterocysts. Since the overexpression of *as\_gltA* reduces citrate synthase activity, this antisense RNA could eventually contribute to the metabolic remodeling that occurs during the differentiation of vegetative cells into heterocysts.

**Keywords:** *Anabaena*, antisense RNA, citrate synthase, heterocyst, posttranscriptional regulation

## Significance Statement

The transcription of bacterial genomes extends beyond the regions classically defined as genes. Antisense transcription of the complementary strand of coding regions is very frequently observed, not only because of an overlap between transcripts that are produced in adjacent genes, but also as autonomous, nonprotein-coding transcripts, whose expression is sometimes regulated in response to environmental changes. The consequences of antisense transcription are still a matter of debate, but accumulating evidence points to possible regulatory roles. In this context, the observation that antisense transcription takes place from nitrogen-regulated promoters controlled by the cyanobacterial regulators NtcA or HetR suggests that these antisense transcripts may play a role in adaptation to nitrogen stress and/or the differentiation of functional heterocysts.

## Introduction

Cyanobacteria are oxygen-producing photosynthetic organisms. In addition to light, which has a daily cycle of availability, nutrients that are critical for cyanobacterial growth, such as nitrogen, are also subject to fluctuations in natural, especially aquatic, environments, imposing the need for complex regulatory circuits to adapt quickly and accurately to changing conditions. Adaptation to nitrogen deficiency in both unicellular and filamentous cyanobacteria is under the control of the global nitrogen regulator NtcA (1) and involves the use of alternative nitrogen sources, as well as the recycling of cellular components (e.g. the phycobiliproteins) as a source of amino acids (2). In addition,

some filamentous cyanobacteria can fix atmospheric nitrogen in specialized cells called heterocysts. Heterocysts differentiate from certain vegetative cells along the filaments in a process that involves structural and metabolic transformations to provide the microaerobic environment required for the function of nitrogenase, the enzyme that performs N<sub>2</sub> fixation (3–5). Many of the gene expression changes associated with heterocyst differentiation are under the control of the HetR regulator (6). Genes whose expression is activated by HetR are induced specifically in cells that are becoming heterocysts. Genes exhibiting NtcA- or HetR-dependent expression constitute the NtcA and HetR regulons, respectively.

**Competing Interest:** The authors declare no competing interest.

**Received:** February 13, 2023. **Revised:** May 24, 2023. **Accepted:** May 30, 2023

© The Author(s) 2023. Published by Oxford University Press on behalf of National Academy of Sciences. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

A previous genome-wide analysis of transcriptional start sites (TSSs) in *Nostoc* sp. PCC 7120 revealed two main categories of nitrogen-regulated transcripts (7). The DEF (deficiency) category includes transcripts that are similarly regulated in the wild-type (WT) and *hetR* mutants, which cannot differentiate heterocysts. Although, in some cases, NtcA might be indirectly responsible for the observed regulation of DEF genes, most promoters in this category contain NtcA-binding sites, and therefore, the corresponding genes constitute the directly regulated NtcA regulon. The second category of nitrogen-regulated promoters, the DIF (differentiation) category, includes transcripts that require HetR for induction and, therefore, are specifically related to heterocyst differentiation. Two groups of genes can be defined in this category (E-DIF and L-DIF) based on their temporal pattern of induction during the early (E) or late (L) stages of differentiation, respectively (6). Most genes in the E-DIF category are transcribed from promoters with conserved sequence motifs (DIF1 or DIF2 motifs), associated with their expression specifically in cells that are initiating differentiation as heterocysts (6, 7).

Previous transcriptomic analyses carried out in both unicellular and filamentous cyanobacteria have shown abundant transcription of antisense RNAs (asRNAs) (8). In the case of *Nostoc* sp. PCC 7120, ~30% of TSSs determined using differential RNA-seq (dRNA-seq) (7) were assigned to the antisense category (aTSS), consistent with the abundance of antisense transcripts in RNA-seq data sets (9). In previous work, we have characterized two antisense RNAs with heterocyst-specific transcription, *as\_glpX* (10) and *NsiR1* (11). Here, we aim at the global analysis of nitrogen-regulated antisense transcription in *Nostoc*, with a focus on asRNAs whose expression is restricted to heterocysts. The identification of asRNAs that are under control of the regulators involved in heterocyst differentiation, and specifically transcribed during the transformation of a vegetative cell into a mature heterocyst, points to a participation of these asRNAs in the posttranscriptional regulation of this unique cell differentiation process.

## Results

### Nitrogen-regulated transcriptome in *Nostoc* sp. PCC 7120

We sequenced the transcriptome of *Nostoc* cells growing in the presence of ammonium (time 0) or subjected to nitrogen deprivation for 9 or 24 h (time 9 and time 24), two time points suitable for the analysis of transcriptional changes associated with the early and late stages of heterocyst differentiation, respectively. RNA from two biological replicas of each condition was sequenced to a depth of 17–21 million paired-end reads each sample. A transcriptome was assembled as detailed in Materials and Methods in combination with a previously available data set of transcription start sites obtained by dRNA-seq (7) and a prediction of transcription terminators. 4,039 transcripts were predicted, 3,127 of them associated with TSS previously defined using cells grown in the presence of ammonium or after 8 h of combined nitrogen deprivation (7) (Data Set S1).

Data Set S1 contains comparisons of the expression of each transcriptional unit (TU) at 9 or 24 h after nitrogen removal versus time 0. Consistent with previous analysis of the response of *Nostoc* sp. PCC 7120 to nitrogen deficiency, transcripts corresponding to NtcA-regulated genes in the DEF category (e.g. *nrrA*, *ntcB*) (7), or HetR-dependent, heterocyst-specific genes in the DIF category (e.g. *het*, *dev*, *nif* genes) (7) are among the transcripts induced more than eight-fold after 9 h of nitrogen starvation [i.e.  $\log_2$  (FC

9–0) > 3]. Also, in agreement with previous transcriptomic studies (6), genes expressed during early stages of heterocyst differentiation (such as *hep* genes involved in polysaccharide deposition) show higher expression at 9 h than at 24 h after nitrogen removal. In contrast, genes whose expression occurs in mature heterocysts (such as *nif* genes that encode nitrogenase and associated functions) show higher expression at 24 h than at 9 h after nitrogen removal.

It should be noted that although this transcriptome was assembled on the basis of RNA from cells growing under only 3 different conditions, the assembled transcriptome contains transcripts covering 75% of all *Nostoc* sp. PCC 7120 genes (4,614 genes of the total of 6,137 genes annotated for this organism according to the NCBI annotation). Therefore, this transcriptome may represent a good overview of the overall transcriptome of our organism, including housekeeping genes expressed independently of environmental condition. We have analyzed some global parameters of the assembled transcriptome, such as the length distribution of transcripts, 5'UTRs and 3'UTRs (Fig. S1). One of the longest transcripts corresponds to the operon encoding ribosomal proteins. Also some operons essential for heterocyst differentiation, such as the *hep* operon or the *hgl* operon, are among the longest TUs in the *Nostoc* transcriptome (Data Set S1). The *Nostoc* transcriptome also contains long 5'UTRs and 3'UTRs (Fig. S1B and C). The presence of long 3'UTRs could be explained by the absence of Rho-termination factors in cyanobacteria (12, 13). To test this hypothesis, we compared the length distribution of 3'UTRs of *Nostoc* and *Synechocystis* sp. PCC 6803, a unicellular cyanobacterium, with the length of 3'UTRs in two gram-negative bacteria encoding homologs of Rho termination factor, *Campylobacter jejuni* 81116 (14) and *Helicobacter pylori* 26695 (15, 16), for which available transcriptomic data allow a similar analysis (17). According to these comparisons, significantly longer 3'UTRs were observed in the two tested cyanobacterial transcriptomes than in the other two transcriptomes analyzed (Fig. S1C). However, the most striking feature was the significant increase in the length of 5'UTRs in the cyanobacterial transcriptomes (Fig. S1B). In all four organisms analyzed, their transcriptomes showed a preference for 5'UTRs with an average length of 30 nucleotides. Because long 5'UTR can be involved in regulation, we also compared the length of TUs corresponding to previously identified HetR-dependent genes (6) with the rest of TUs in the data set. In fact, the length of 5'UTRs in HetR-regulated transcripts is significantly higher than in non-HetR-regulated transcripts (Fig. S2).

### Nitrogen-regulated antisense transcription in *Nostoc* sp. PCC 7120

The presence of long 5'UTRs and 3'UTRs in the *Nostoc* transcriptome led us to analyze the possible antisense regulation exerted by overlapping protein-coding mRNAs during the response to nitrogen deprivation. All TUs that overlapped with a TU of the other strand for >50 nucleotides were classified as asRNAs (Data Set S1). These represent 2,620 transcripts (65%), which is consistent with the prevalence of antisense transcription previously described for cyanobacteria (8, 18). Most asRNAs were, in fact, protein-coding mRNAs that overlapped (1757). However, we also found 863 antisense transcripts that were not predicted to encode any protein.

Among the TUs classified as asRNAs, we selected for further analysis those showing  $\log_2$  (FC 9-0) above 3. In 23 cases, the asRNA is an mRNA containing at least one coding sequence (Table 1). Among these nitrogen-regulated overlapping mRNAs,

**Table 1.** Selected nitrogen-regulated mRNAs antisense to another mRNA.

TU	Nitrogen-regulated transcript			Regulation <sup>a</sup>	Overlapped transcript	
	log <sub>2</sub> (FC 9-0)	log <sub>2</sub> (FC 24-0)	Gene(s)		TU	Gene(s)
TU01301	7.28	6.04	<i>asl2052</i>	NtcA	TU01300	<i>ggt</i>
TU03408	6.23	4.75	<i>alr4984</i>	E-DIF	TU03409	<i>susA</i>
TU02503	6.18	3.76	<i>all3772</i>	E-DIF	TU02502	<i>alr3771</i>
TU00203	5.42	7.24	<i>hesF</i>	L-DIF	TU00204	<i>prfB</i> to <i>all0268</i>
TU00037	5.27	3.50	<i>all0059</i>		TU00036	<i>alr0056-alr0058</i>
TU02896	5.15	3.59	<i>nrrA</i>	NtcA	TU02895	<i>alr4310</i>
TU01911	4.82	5.80	<i>all2965</i>	L-DIF	TU01912	<i>alr2966-alr2967</i>
TU01656	4.65	2.96	<i>all2571</i>	E-DIF	TU01655	<i>petR-petP</i>
TU00909	4.56	5.50	<i>nifV1</i> to <i>alr1410</i>		TU00910	<i>asl1412-all1411</i>
TU01443	4.33	4.49	<i>asl2299</i>		TU01442	<i>alr2298</i>
TU00901	4.30	4.47	<i>all1395-asl1394</i>	NtcA	TU00899	<i>alr1390</i> to <i>PCC7120DELTA_RS32405</i>
TU02336	4.01	3.59	<i>alr3554</i>		TU02337	<i>all3556-ilvB</i>
TU02841	3.86	2.58	<i>PCC7120DELTA_RS22895</i>		TU02840	<i>argJ2</i>
TU02078	3.81	2.08	<i>alr3195</i>	E-DIF	TU02079	<i>asl3196</i>
					TU02080	<i>all3197</i>
TU01767	3.78	3.22	<i>all2747</i>		TU01765	<i>alr2745-alr2746</i>
TU02125	3.75	2.59	<i>asr3279</i>	NtcA	TU02124	<i>all3278</i>
TU01093	3.71	2.86	<i>asr1734</i>		TU01094	<i>all1736</i>
TU00968	3.64	2.72	<i>patD</i>		TU00967	<i>all1549</i>
TU02430	3.59	1.80	<i>alr3676</i>	E-DIF	TU02429	<i>all3675</i>
					TU02431	<i>all3677</i>
TU00398	3.43	2.74	<i>alr0627</i>		TU00399	<i>asl0628</i>
TU01671	3.39	2.88	<i>alr2582</i>	L-DIF	TU01673	<i>fecC1</i> to <i>fecE1</i>
			<i>fecB1</i>			
TU03256	3.23	2.22	<i>alr4800</i>		TU03257	<i>rpsF-all4801</i>
TU01496	3.17	1.96	<i>ald</i>		TU01497	<i>all2358</i> to <i>phnE2</i>

<sup>a</sup>NtcA indicates that the nitrogen-regulated gene is preceded by a previously described NtcA-regulated promoter (7, 60). E-DIF or L-DIF indicates that the nitrogen-regulated gene appears in one of the previously defined groups of HetR-dependent genes (early or late, respectively) (6).

we found genes whose expression depends on HetR or NtcA. Eight mRNAs include genes that belong to one of the two previously described groups of HetR-dependent transcripts (named E-DIF and L-DIF according to their early or late expression, respectively) (6). Most overlapping mRNAs in the E-DIF group, such as those of *alr4984*, *all2571*, *alr3195*, or *alr3676*, are transcribed from promoters bearing a DIF1 motif associated with heterocyst-specific expression (6, 7). In some other cases, nitrogen-regulated overlapping mRNAs are transcribed from promoters containing NtcA-binding sites, such as those of *asl2052*, *nrrA*, or *all1395* (7). We noticed a prevalence of tail-to-tail disposition, but there are also examples of head-to-head disposition. As an illustration of tail-to-tail or head-to-head disposition, two examples involving antisense HetR-dependent TUs (TU03408, *as\_susA* and TU00968, *as\_all1549-patD*) are shown in Fig. S3 together with TU01524, previously described *as\_fraE* (TU01524), which corresponds to overlapping transcription of the 3'UTR of *fraF* mRNA (19) but with log<sub>2</sub> (FC 9-0) = 1.95 falls below the threshold log<sub>2</sub> (FC 9-0) > 3 applied above.

In contrast to the above-described overlapping mRNAs, in 12 cases, antisense transcripts with log<sub>2</sub> (FC 9-0) > 3 correspond to TUs that do not contain any annotated coding sequence. These noncoding antisense transcripts are produced from TSSs annotated as antisense TSS (aTSS) or nTSS (noncoding TSS) (7) (Table 2). Transcription from some of these TSS is driven by promoters bearing DIF1 motifs (*as\_gltA*, *as\_hglD*, *as\_acsF*) or NtcA-binding sites (*as\_pstS2*, *as\_pknC*, *as\_fdxB*, *as\_leuA*) (7). All three asRNAs transcribed from DIF1-containing promoters belong to the E-DIF group of HetR-dependent transcripts (6). Two nitrogen-regulated noncoding asRNAs previously described appeared in our data set with log<sub>2</sub> (FC 9-0) below 3. Data for the regions corresponding to *as\_glpX* [TU00652; log<sub>2</sub> (FC 9-0) = 2.27] (10) and to *as\_anacyp40* [TU03461; log<sub>2</sub> (FC 9-0) = 1.98] (6, 20),

both transcribed from DIF1-containing HetR-dependent promoters, are shown in Fig. S4.

Taken together, these observations suggest that both the NtcA-regulated responses (not specifically linked to heterocyst differentiation) and the HetR-regulated responses (specifically associated with heterocyst differentiation) include abundant transcription of asRNAs, either because of overlap between mRNAs located in a head-to-head or tail-to-tail disposition or due to the transcription of *bona fide*, noncoding asRNAs.

### Noncoding antisense transcripts in the NtcA and HetR regulons

Observed overlap of antisense protein-encoding mRNAs with long 3'UTRs could be the result of improper transcription termination. However, the presence of noncoding antisense RNAs with strict transcriptional regulation strongly suggests a regulatory function for this type of RNAs. We have selected for further validation nitrogen-regulated TU with log<sub>2</sub> (FC 9-0) > 3 that fall in the category of noncoding asRNAs (i.e. are not mRNAs) and are transcribed antisense to genes with a previous annotation (*as\_pstS2*, *as\_hglD*, *as\_fdxB*, *as\_gltA*, *as\_acsF*, *as\_leuA*, *as\_pknC*; Table 2). These asRNAs were transcribed antisense to genes that belong to different functional categories, from amino acid biosynthesis (*leuA*) to regulation (*pknC*, protein kinase).

Figure 1A shows RNA-seq data for four selected asRNAs transcribed from NtcA-regulated promoters, after 9 or 24 h of nitrogen deprivation. Their transcription was verified by primer extension using oligonucleotides located at least 100 nucleotides downstream of the associated aTSS (7) (see Materials and Methods for details). Products of the expected size (from the position of the oligonucleotide used in the primer extension reaction to the position of the aTSS) were observed (Fig. 1B). In the four cases, the

**Table 2.** Nitrogen-regulated noncoding RNAs antisense to an mRNA.

Nitrogen-regulated transcript					Overlapped transcript	
TU	Name	Log <sub>2</sub> (FC 9-0)	Log <sub>2</sub> (FC 24-0)	Regulation <sup>a</sup>	TU	Gene(s)
TU02147		5.11	5.09		TU02146	all3305
TU03089	<i>as_pstS2</i>	4.51	3.60	NtcA	TU03088	<i>pstS2-pstC2-pstA2-pstB2</i>
TU03269	<i>as_pknC</i>	4.43	4.50	NtcA	TU03268	<i>pknC</i>
TU00166	<i>as_gltA</i>	4.43	3.70	E-DIF	TU00167	<i>gltA</i>
TU03638	<i>as_hglD</i>	4.40	3.42	E-DIF	TU03636	<i>asr5350-hglEF-hglG-hglD-hglC-hglA-hglB</i>
TU00092		4.25	4.33	L-DIF	TU00091	<i>alr0130</i>
TU01725		4.05	3.53		TU01724	<i>all2700</i>
TU01622	<i>as_fdxB</i>	4.03	3.42	(NtcA)	TU01621	<i>fdxB</i>
TU02297	<i>as_leuA</i>	3.79	2.87	NtcA	TU02296	<i>leuA</i>
TU01075		3.79	3.48		TU01074	<i>alr1712</i>
TU00785		3.56	2.96		TU01076	<i>alr1713-asr1714-alr1715</i>
TU00867	<i>as_acsF</i>	3.02	1.97	E-DIF	TU00784	<i>alr1240-alr1241-alr1242</i>
					TU00866	<i>acsF</i>

<sup>a</sup>NtcA indicates that the nitrogen-regulated gene is preceded by a previously described NtcA-regulated promoter (7) or, in the case of TU01622, a promoter with an NtcA-binding site identified by ocular inspection (NtcA). E-DIF or L-DIF indicates that the nitrogen-regulated gene appears in one of the previously defined groups of HetR-dependent genes (early or late, respectively) (6).

expression of inducible asRNAs did not take place in the *ntcA* mutant strain, but was still observed in the *hetR* mutant strain, confirming that these asRNAs belong to the NtcA regulon and the DEF category, according to (7). Comparison of promoters upstream of the corresponding aTSS shows the presence of putative NtcA-binding sites in a position compatible with direct activation by NtcA (the well-described NtcA-regulated promoter of *glnA* is included for comparison; Fig. 1C).

Figure 2A shows RNA-seq data of three selected asRNAs associated with aTSS in the DIF category, therefore potentially linked to heterocyst differentiation (*as\_hglD*, *as\_gltA*, and *as\_acsF*). Again, we verified nitrogen-regulated transcription from the corresponding aTSS along a time course of nitrogen deprivation by primer extension, obtaining products of the expected sizes (Fig. 2B). Consistent with their classification as E-DIF (6), expression of all three was already induced 3 h after nitrogen removal. In the three cases, the comparison of the promoters upstream of the corresponding HetR-dependent aTSS (7) reveals the presence of DIF1 motifs, associated with heterocyst-specific expression (the DIF1-containing promoter of *hetR* is included for comparison; Fig. 2C).

### A heterocyst-specific asRNA in the *gltA* gene

One of the nitrogen-regulated noncoding asRNAs verified above, *as\_gltA*, TU00166, is transcribed from a TSS at position 238644r, on the antisense strand of *gltA*, encoding citrate synthase (Fig. 3). This TSS is located at position +474 relative to the start codon of *gltA* (Fig. 3A). The accumulation of the corresponding transcript is induced in a HetR-dependent manner, as demonstrated by primer extension (Fig. 3B). The 3' end of *as\_gltA* was identified by 3'RACE. The most abundant sequenced products ended between positions 238399 and 238424, but less abundant, longer products extending to position 238278 were also detected (not shown). The accumulation of *as\_gltA* was verified by Northern blot, which revealed, in agreement with the RACE results, an abundant band slightly above 200 nt and another less abundant band, above 300 nt, as well as some signals above the 400 nt size marker. No expression of the *as\_gltA* RNA was observed in the *hetR* mutant (Fig. 3B and C). The size of the predicted transcript TU00166 (866 bp) suggests that the major band detected by Northern blot might be generated by processing TU00166.

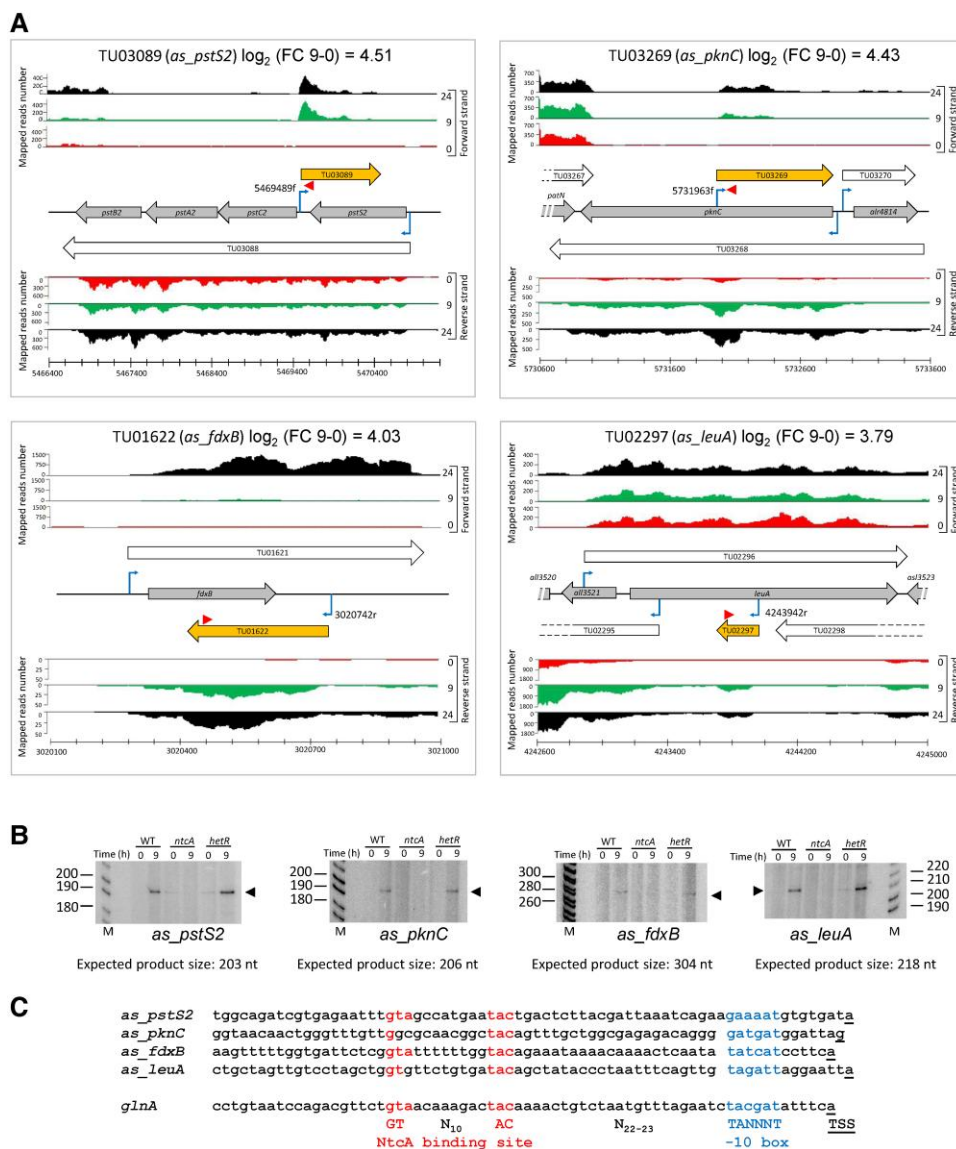
The presence of a DIF1 motif-containing promoter (Fig. 2) and the HetR dependence of *as\_gltA* transcription (Fig. 3B and C)

suggest that the expression of *as\_gltA* is heterocyst specific. To confirm this hypothesis, we constructed a *Nostoc* strain that carries the *gfpmut2* gene under control of the *as\_gltA* promoter (positions 238804 to 238632, sequence –160 to +13 with respect to the TSS; see pMBA75 in Fig. 3A). We also constructed a strain that carries the *gfpmut2* gene under the control of the *gltA* promoter (positions 237627 to 237874, sequence –200 to +47 with respect to TSS; see pMBA117 in Fig. 3A). GFP expression was analyzed by confocal fluorescence microscopy in both strains (Fig. 3D). Quantification of GFP fluorescence along the filament demonstrated that transcription from the promoter of *as\_gltA* took place specifically in heterocysts and could be detected in cells that had only initiated the process of differentiation, even before characteristic signs of differentiation such as reduced red autofluorescence or increased size were observed (see cells indicated with asterisks in the left panels of Fig. 3D and E). Early expression of *as\_gltA* in cells that are only initiating differentiation as heterocysts is consistent with the expression pattern observed for genes that contain DIF1-containing promoters such as *NsiR1* (21) or *hetR* (6). In contrast, there was a similar level of GFP expression from the *gltA* promoter along the filament in both vegetative cells and heterocysts (right panels in Fig. 3D and E).

### Effect of *as\_gltA* on the expression of citrate synthase

We hypothesize that transcription of *as\_gltA* might regulate the accumulation or translation of *gltA* mRNA and, therefore, the amount of citrate synthase. Because *as\_gltA* is transcribed specifically in heterocysts (Fig. 3D and E), any effect on the accumulation of *gltA* would be restricted to these cells, which represent a small percentage of the cells in the filaments. Therefore, to analyze possible effects of *as\_gltA*, we have constructed a strain of *Nostoc* that overexpresses *as\_gltA* from a constitutive promoter in all cells of the filament (OE\_ *as\_gltA*) and a strain that overexpresses an antisense RNA of *as\_gltA* (OE\_ *as\_as\_gltA*). The 245 nt fragment defined as the main product of *as\_gltA* transcription by Northern blot (Fig. 3C) and 3'RACE (positions 238644r to 238399r) was cloned between the *trc* promoter and the strong T1 transcriptional terminator of *rmB* of *Escherichia coli*, and the resulting plasmid (pMBA104) was introduced in *Nostoc*, generating strain OE\_ *as\_gltA*. To remove *as\_gltA* without altering its genomic region, we cloned the same fragment in reverse orientation downstream of the *trc* promoter to generate a

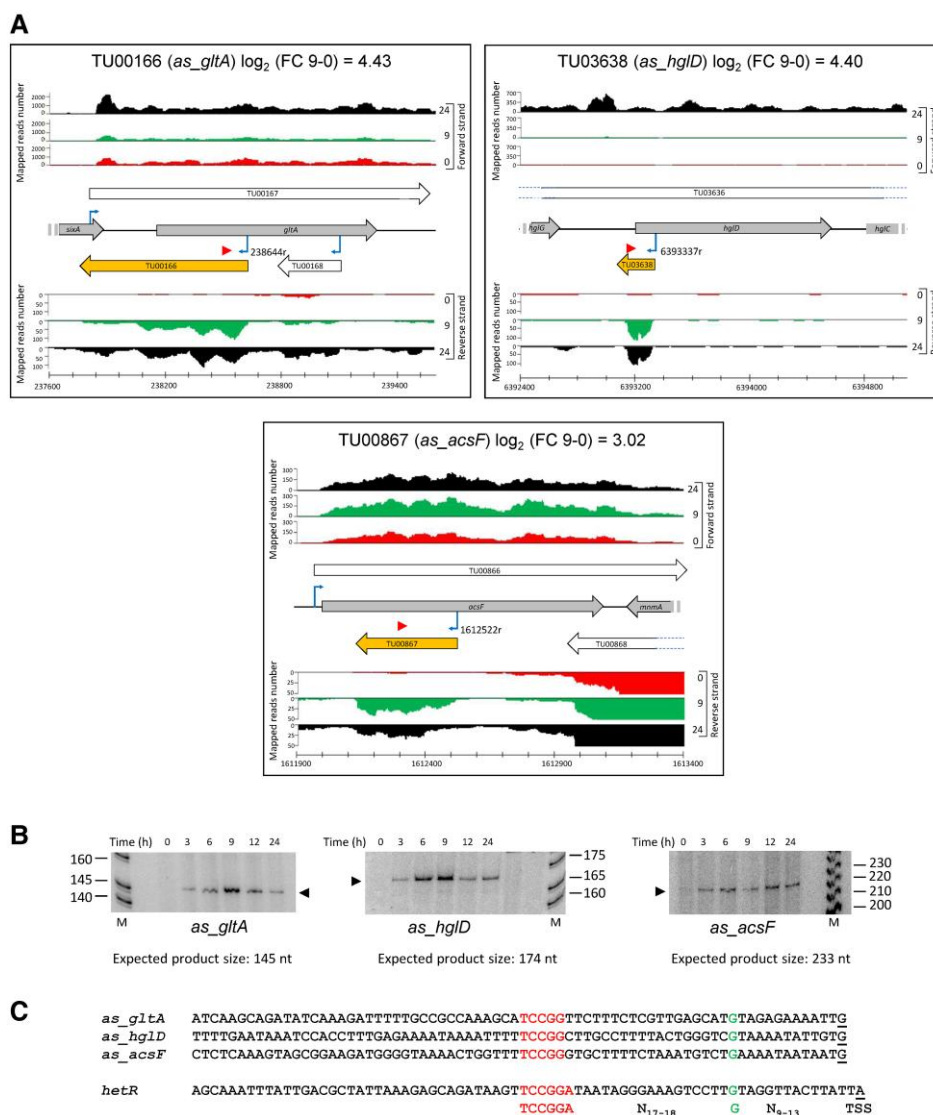




**Fig. 1.** Examples of NtcA-dependent nitrogen-regulated asRNAs. A) Distribution of reads mapped in relevant regions of RNA obtained from cells grown in the presence of  $\text{NH}_4^+$  (red), or after 9 (green) or 24 h (black) in the absence of combined nitrogen. Annotated ORFs are represented by gray arrows. TUs are represented by white arrows with their corresponding identification. Nitrogen-regulated antisense TUs are colored orange. Previously identified TSSs are indicated by bent blue arrows. Genomic coordinates denote the position on the *Nostoc* sp. PCC 7120 chromosome. The scale indicates the number of mapped reads per nucleotide position. B) Primer extension analysis of the asRNAs shown in A using RNA extracted from the cells of WT, *ntcA* mutant, or *hetR* mutant grown in the presence of  $\text{NH}_4^+$  (0), or after 9 h (9) in the absence of combined nitrogen. The positions of the primers used are indicated by red triangles in A. Black triangles point to the primer extension products. The expected sizes of the extension products are included under each panel. Relevant size markers (M) are indicated. C) Sequences upstream of the TSS of the four analyzed asRNAs, compared with the sequence upstream of the NtcA-regulated TSS of *glnA* and the consensus for NtcA-activated promoters. The conserved NtcA-binding motif is colored red. The -10 box is highlighted in blue, and the TSS is underlined.

perfect antisense that would act as a sponge, neutralizing *as\_gltA*. The resulting plasmid (pMBA105) was introduced in *Nostoc*, generating strain OE\_*as\_as\_gltA*. We have previously used this strategy to deplete different transcripts in *Nostoc* (22, 23). pMBA51, a plasmid that does not have an insert between the *trc* promoter and the T1 terminator (10), was also introduced in *Nostoc* as a control, generating strain OE\_C (Fig. 4A). We first analyzed the expression of *as\_gltA* by Northern blot in the three strains after 18 h of culture in medium lacking combined nitrogen (Fig. 4B). As expected, strain OE\_*as\_gltA* accumulated a much higher amount of *as\_gltA* than the control strain OE\_C, while strain OE\_*as\_as\_gltA*, which expresses an antisense RNA

to *as\_gltA*, had a reduced amount of *as\_gltA*. No significant growth defect was detected in strains OE\_*as\_gltA* and OE\_*as\_as\_gltA* after combined nitrogen depletion (not shown). We next studied the effect of the changes in the amount of *as\_gltA* on the accumulation of *gltA* mRNA by Northern blot (Fig. 4C). Overexpression of *as\_gltA* resulted in a 30% reduction in the *gltA* mRNA signal with respect to the control strain, while depletion of *as\_gltA* had no significant effect. Finally, we measured citrate synthase activity in the three strains (Fig. 4D). Overexpression of *as\_gltA* resulted in a 45% reduction in citrate synthase activity, while depletion of *as\_gltA* had no significant effect.



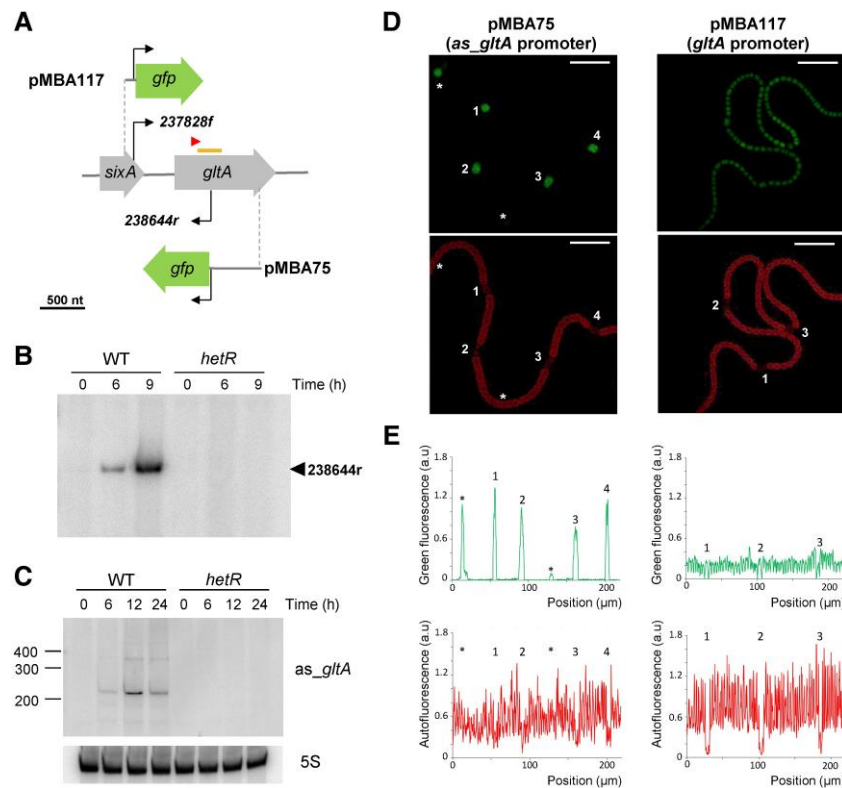
**Fig. 2.** Examples of HetR-dependent nitrogen-regulated asRNAs. A) Distribution of reads mapped in relevant regions of RNA obtained from cells grown in the presence of  $\text{NH}_4^+$  (red), or after 9 (green) or 24 h in the absence of combined nitrogen. All labels are as in Fig. 1. For *as\_acsF*, the scale of the reverse strand was adjusted for better visualization of the reads mapped to the antisense RNA. B) Primer extension analysis of asRNAs shown in A using RNA extracted from cells of WT strain grown in the presence of  $\text{NH}_4^+$  (0), or after 3 to 24 h in the absence of combined nitrogen. The positions of the primers used are indicated by red triangles in A. Black triangles point to the primer extension products. The expected sizes of the extension products are included under each panel. Relevant size markers (M) are indicated. C) Sequences upstream of the TSS of the three analyzed asRNAs compared with the sequence upstream of the heterocyst-specific TSS of *hetR*. The conserved DIF1 motif is highlighted in red. A conserved G is highlighted in green, and the TSS is underlined.

## Discussion

Photosynthetic microorganisms face unique challenges that make their survival dependent on sophisticated regulatory circuits. In the case of heterocystous cyanobacteria, adaptation to nitrogen limitation involves a cell differentiation process that completely transforms the physiology of the filament as a whole to survive using atmospheric  $\text{N}_2$  as the only nitrogen source. In this context, regulatory mechanisms orchestrated by noncoding RNAs have already been described that participate in the differentiation of certain cells into heterocysts. Differential accumulation in heterocysts of trans-acting sRNAs NsiR1, NsiR3, or NsiR4 modulates the differentiation process (11, 23) or the activities of several enzymes (24, 25), specifically in developing heterocysts. In this work, we provide the first transcriptional map of a heterocystous cyanobacterium along the heterocyst differentiation

process. The boundaries of the transcripts are based on conventional RNA-seq data from this work, an available TSS data set (7), and a prediction of transcriptional terminators.

Antisense transcription is widespread in all organisms. A significant proportion of the transcripts identified in bacteria that have been analyzed using high-throughput methodologies are in antisense orientation, from *E. coli* (26) to *Staphylococcus* (27) or *Prochlorococcus*, a cyanobacterium with a streamlined genome (28). By using a combination of RNA-seq with a previously available TSS data set (7) and a prediction of Rho-independent transcriptional terminators, in this work, we provide a transcriptomic map of *Nostoc* sp. PCC 7120, and show that in this heterocystous cyanobacterium, similar to the observations made for other bacteria, a large number of TUs (65%) overlap another TU in the antisense orientation. We have focused on nitrogen-regulated antisense transcripts as a potential source of regulation in the



**Fig. 3.** Expression pattern of the *gltA* and *as\_gltA* promoters in *Nostoc* filaments. A) Scheme of the promoter fusions to GFP in plasmids pMBA75 and pMBA117. B) Primer extension analysis of *as\_gltA* using RNA extracted from cells of WT or *hetR* mutant grown in the presence of  $\text{NH}_4^+$  (0), or after 9 h in the absence of combined nitrogen. C) Northern blot of *as\_gltA* with RNA extracted from cells of WT or *hetR* mutant grown in the presence of  $\text{NH}_4^+$  (0), or after 6 to 24 h in the absence of combined nitrogen. As a loading control, the filter was hybridized to a probe for 5S RNA. D) Representative confocal fluorescence images of *Nostoc* filaments carrying the *gfp* gene under the control of the *as\_gltA* promoter (plasmid pMBA75, left) or the *gltA* promoter (plasmid pMBA117, right) and growing on top of medium lacking any source of combined nitrogen are shown for the green channel (GFP fluorescence) and the red channel (autofluorescence). Numbers indicate heterocysts, and asterisks indicate immature heterocysts. Scale bars 20  $\mu\text{m}$ . E) Quantification of the signals for the green and red channels in the images shown in D. Left, pMBA75; right, pMBA117. Numbers and asterisks as in D.

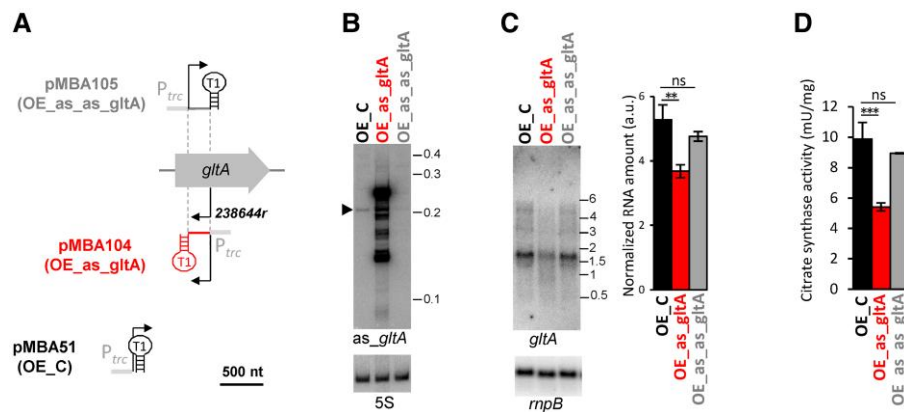
context of adaptation to nitrogen deficiency and, specifically, in the differentiation of heterocysts.

Two major categories of asRNAs can be considered. On the one hand, mRNAs that are transcribed in a convergent (tail to tail) or, less frequently, in a divergent (head to head) disposition may overlap because their long 5' or 3'UTRs invade the transcriptional space of adjacent genes. In these cases, the consequences of nitrogen-regulated transcription of a given transcript go beyond the expression of the gene(s) included, potentially affecting the expression of oppositely oriented adjacent genes via antisense transcription (some examples are shown in Fig. S3). In contrast to overlapping mRNAs, a second group of asRNAs, sometimes referred to in the literature as *cis*-antisense or *bona fide* asRNAs, includes noncoding RNAs transcribed antisense to an mRNA (see examples in Figs. 1, 2, and S4). In this work, we have validated several of these asRNAs that are transcribed from NtcA- or HetR-dependent promoters (Figs. 1 and 2). These observations suggest that the two transcriptional regulators that control both the response to nitrogen limitation and the differentiation of heterocysts could carry out posttranscriptional regulation mediated by nitrogen-regulated asRNAs.

Overlapping asRNAs could be particularly abundant in cyanobacteria as a consequence of the absence of the Rho termination factor, which has been involved in the repression of pervasive, including antisense, transcription (29, 30). Because Rho protein is not found in the cyanobacterial clade (12, 13), transcriptional termination is based on the structure of the nascent mRNA, which

usually produces longer 3'UTRs in cyanobacteria (Fig. S1C). We observed a clear preference for 5'UTRs of around 20–40 nucleotides in the four transcriptomes tested (Fig. S1B), which may reflect an optimization for translation (16). The cyanobacterial transcriptomes analyzed also showed longer 5'UTRs, with some 5'UTRs longer than 500 nucleotides (Fig. S1B). Long 5'UTRs may have regulatory functions. Our observation that HetR-regulated genes have longer 5'UTRs than nonregulated genes (Fig. S2) supports this possibility. Long 5'UTRs can also generate overlapping head-to-head mRNAs. For example, the presence of an 800-nt long 5'UTR in *conR* mRNA (TU00137), a gene essential for diazotrophic growth (31), overlaps in a head-to-head disposition with *alr0188* (GDP-mannose pyrophosphorylase). The commonly accepted concept of operon, a genomic arrangement in which genes involved in the same metabolic pathway appear clustered together to facilitate coupled regulation, may be extended to antisense regulation. The genomic arrangement of some overlapping antisense transcripts may also be used for coupled regulation.

The physiological role of antisense transcription remains a matter of debate (32, 33). Pervasive, nonregulated, antisense transcription points to a global function in RNA processing (34) or transcription-coupled DNA repair (35). In contrast, asRNAs whose expression responds to environmental cues and hence becomes differentially expressed under certain conditions or, as exemplified here by heterocyst-specific asRNAs, in a specific cell type, are suggestive of a functional role in the adaptation to certain



**Fig. 4.** Effect of *as\_gltA* overexpression or depletion on *gltA* expression and citrate synthase activity. A) Scheme of the DNA fragments cloned in the plasmids constructed to generate *Nostoc* strains with altered levels of *as\_gltA* RNA. Transcriptional start sites (bent arrows), T1 terminators (stem loops), and *trc* promoters are indicated. B) Northern blot analysis of the accumulation of *as\_gltA* transcript using RNA extracted 18 h after nitrogen removal from the control strain (OE\_C), the strain overexpressing *as\_gltA* (OE\_as\_gltA), and the strain depleted of *as\_gltA* (OE\_as\_as\_gltA). 5S RNA was used as a loading control. C) Left, Northern blot analysis of the accumulation of *gltA* transcript using RNA extracted 18 h after nitrogen removal from the control strain (OE\_C), the strain overexpressing *as\_gltA* (OE\_as\_gltA), and the strain depleted of *as\_gltA* (OE\_as\_as\_gltA). *rnpB* was used as a loading control. Right, quantification of *gltA* mRNA in the three strains. Average and standard deviation of the quantification of three biological replicates. D) Citrate synthase activity in crude extracts prepared 18 h after nitrogen removal from the control strain (OE\_C), the strain overexpressing *as\_gltA* (OE\_as\_gltA), and the strain depleted of *as\_gltA* (OE\_as\_as\_gltA). Average and standard deviation of the quantification of three biological replicates (ns, not significant; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ , Student's *t* test).

environmental changes. In both *Listeria monocytogenes* and *Staphylococcus aureus*, a subset of asRNAs are transcribed from SigB-dependent promoters (27, 36), suggesting a role under specific conditions. Here, we describe several asRNAs whose expression responds to nitrogen availability and is regulated in a NtcA- or HetR-dependent manner. In fact, some of the asRNAs described here belong to one of the previously defined groups of HetR-dependent transcripts (6) and exhibit expression profiles that match those of well-known genes involved in heterocyst differentiation and function, again suggesting a functional relevance in the heterocyst differentiation process.

Here, we show that *as\_gltA* is HetR dependent and is expressed specifically in heterocysts. Its overexpression reduces the amount of *gltA* mRNA and citrate synthase activity. Therefore, our results suggest that one specific metabolic adaptation of the heterocyst is the reduction of citrate synthase activity, resulting in reduced flow through the tricarboxylic acid (TCA) cycle. We can envision several consequences of such a reduced flow. The two substrates of citrate synthase, oxalacetate and acetyl-CoA would be more available for alternative pathways upon inhibition of this enzymatic activity in the heterocyst. Heterocysts synthesize cyanophycin, an aspartate–arginine polymer (37). Reduction of flow through the TCA cycle would favor the channeling of oxalacetate toward aspartate synthesis. Similarly, a greater pool of acetyl-CoA would be available for heterocyst-specific biosynthetic pathways, including envelope glycolipids (5). In fact, partial repression of citrate synthase activity, resulting in increased Acetyl-CoA pools, has been shown to be a successful strategy to increase carbon partitioning and biofuel production in unicellular cyanobacteria (38). Therefore, the *as\_gltA* described here would participate in the metabolic remodeling of heterocysts, similar to the previously described heterocyst-specific asRNA to the *glpX* gene (Fig. S4) that contributes to the shutdown of CO<sub>2</sub> fixation in heterocysts by negatively regulating sedoheptulose-1,7 bisphosphatase (10). Additionally, a heterocyst-specific asRNA to gene *alr5059* (6) (Fig. S4) could modulate the accumulation of its product, recently described as *anaCyp40*, a cyclophilin that regulates the assembly of the photosystem and the association of the phycobilisome

(20). One particular circumstance that hinders the identification of HetR-dependent, heterocyst-specific RNAs, including asRNAs, is the fact that they are transcribed only in the cells that are differentiating (usually <10% of the cells in the filaments) but not in vegetative cells. In this context, the observation that a heterocyst-specific asRNA appears in our transcriptomic data for whole filaments with a strong change of expression implies very strong induction at the single cell level (i.e. in developing heterocysts). Taken together, these observations suggest that regulatory mechanisms orchestrated by asRNAs may represent a major strategy to regulate metabolic remodeling in heterocysts.

The specific mechanisms used by asRNAs in the regulation of gene expression range from co-degradation of RNA duplexes by RNases to regulation of mRNA traducibility (8, 32, 33). It should be mentioned that the output of antisense interactions might be either positive or negative. In fact, previous work in *Synechocystis* sp. PCC6803 showed that both modes of regulation can be functionally relevant (39–42). We can speculate about the possible regulatory output of the transcription of some of the nitrogen-regulated asRNAs described here. For example, heterocyst-specific transcription of TU03408 (*as\_susA*, Fig. S3) could contribute to the down-regulation of *susA* expression in heterocysts, consistent with the observed down-regulation of the expression of the *susA* promoter in these cells (43). Similarly, nitrogen-regulated transcription of TU00968 (*as\_all1549-patD*, Fig. S3) could affect the expression of *all1549*, involved in ppGpp metabolism and required for heterocyst differentiation (44). The negative effects of overlapping mRNAs would be consistent with those previously described for *as\_furA* (45) or hypothesized for *as\_fraE*, which is transcribed as an extension of the *fraF* mRNA, and represents an example of two convergent genes with related functions but having opposite effects (19). This arrangement represents a paradigm in RNA-mediated regulation, the excludon concept (46). Dedicated RNAs could also constitute a mechanism to regulate one particular gene in an operon. As an example, one of the validated NtcA-regulated asRNAs (TU03089, *as\_pstS2*, Fig. 1) could regulate an ABC transporter for phosphate whose four subunits (PstB2, PstA2, PstC2, and PstS2) are cotranscribed in a single



mRNA (TU03088). While transcription of this mRNA shows little variation upon nitrogen removal, transcription of *as\_pstS2*, which overlaps only the first gene in the transcript, encoding the substrate binding subunit, is strongly induced upon nitrogen removal, suggesting that this asRNA could regulate a constitutively expressed transporter by affecting only the expression of the subunit involved in substrate recognition.

The morphological and metabolic transformation of certain cells of a cyanobacterial filament into a heterocyst requires different gene expression programs in adjacent cells. One mechanism behind such expression patterns is the occurrence of heterocyst-specific promoters, which have long been described for most genes involved in differentiation (4, 7), and also contributes to the proper transcription of housekeeping genes whose expression patterns must be modulated during differentiation, such as *sigA*, encoding the major sigma factor (47). In this context, the existence of noncoding transcripts, both sRNAs and asRNAs, with cell-specific expression could also contribute to the onset of cell-specific expression programs.

In this work, we have shown that there are several asRNAs with heterocyst-specific regulated expression that could be involved in modulating gene expression in heterocysts. In fact, we demonstrate that an antisense to *gltA*, which is strongly induced in developing heterocysts, regulates citrate synthase expression. asRNAs constitute an additional layer of regulation that seems to be involved in heterocyst differentiation. Further study of the catalog of heterocyst-specific asRNAs described here would reveal details on how heterocysts achieve differential gene expression beyond transcriptional regulation by protein factors.

## Materials and methods

### Strains and growth conditions

The different *Nostoc* sp. PCC 7120 strains used in this work (Table S1) were grown in BG11 medium (48) as detailed in Supplementary Materials and Methods. *Escherichia coli* strains (Table S1) were grown in LB medium, supplemented with appropriate antibiotics (49).

### Construction of *Nostoc* sp. PCC 7120 derivative strains

Details on the construction of plasmids (Tables S2 and S3) and strains are given in the Supplementary Materials and Methods.

### RNA preparation, library processing, and RNA-seq analysis

Total RNA was isolated using hot phenol as described (50) with modifications (51). RNA samples were treated with turbo DNase (Invitrogen), and strand-specific libraries were prepared using Illumina Stranded TOTAL RNA preparation RIBO-ZERO PLUS kit and sequenced on the Illumina platform NextSeq500 at the Genomics Core Facility of Cabimer (Seville, Spain). Raw RNA-Seq data can be accessed in the GEO database under accession number GSE212705. Sequencing reads were mapped with HISAT2 (52) and the prediction of TUs was carried out by ANNOgesic (17). Fragments associated with TUs were counted using HTseq (53) and differential expression analysis was carried out using *edgeR* (54) and *limma* (55) package in R. Previously published data were used for the analysis of TUs in *Synechocystis* sp. PCC 6803 (56), *C. jejuni* 81116 (14), and *H. pylori* 26695 (15, 16).

Further details about the bioinformatic methods are given in the Supplementary Materials and Methods.

### Northern blot analysis, 3'RACE and primer extension assays

Strand-specific <sup>32</sup>P probe labeling and Northern blots were performed as detailed in the Supplementary Materials and Methods. The RNA used as a probe for *as\_gltA* was transcribed in vitro (MEGAscript Kit; Life Technologies, #AM1330) from PCR-generated DNA templates using the primers indicated in the Table S2. 3'RACE (10) and primer extension analysis (57) were performed as previously described using the oligonucleotides specified in Table S2.

### Fluorescence microscopy

Details of fluorescence quantification are given in the Supplementary Materials and Methods.

### Citrate synthase assay

The crude extracts were prepared in 50 mM Tris-HCl pH 8 buffer by mechanically disrupting the cells using glass beads. Protein concentration was determined using the Bradford method (58) and citrate synthase activity was determined using a 5',5'-dithiobis-(2-nitrobenzoate) colorimetric assay (59). Details about the composition of the reaction mixtures are given in the Supplementary Materials and Methods.

### Statistical methods

The Wilcoxon–Mann–Whitney test and Student's t test were used to determine statistical significance. The use of each test and the number of biological samples can be found in the figure legends.

### Acknowledgments

The authors thank Alicia Orea (IBVF, CSIC-Universidad de Sevilla) for her help with microscopy. We acknowledge support of the publication fee by the CSIC Open Access Publication Support Initiative through its Unit of Information Resources for Research (URICI).

### Supplementary material

Supplementary material is available at PNAS Nexus online.

### Funding

This work was supported by grant PID2019-105526GB-I00, from Agencia Estatal de Investigación (AEI/10.13039/501100011033), Ministerio de Ciencia e Innovación, Spain, to A.M.M.P. and grant US-1379643 from Consejería de Universidad, Investigación e Innovación, Junta de Andalucía, Spain and European Regional Development Fund to A.V.

### Author contributions

M.B.A., A.V., and A.M.M.P. designed the research; M.B.A., A.V., and A.M.M.P. performed the research; M.B.A., A.V., and A.M.M.P. analyzed the data; and M.B.A., A.V., and A.M.M.P. wrote the paper.

### Data availability

All data are included in the manuscript and/or supporting information. Raw RNA-Seq data can be accessed in the GEO database under accession number GSE212705.

## References

- 1 Herrero A, Muro-Pastor AM, Flores E. 2001. Nitrogen control in cyanobacteria. *J Bacteriol.* 183:411–425.
- 2 Grossman AR, Schaefer MR, Chiang GG, Collier JL. 1993. The phycobilisome, a light-harvesting complex responsive to environmental conditions. *Microbiol Rev.* 57:725–749.
- 3 Flores E, Herrero A. 2010. Compartmentalized function through cell differentiation in filamentous cyanobacteria. *Nat Rev Microbiol.* 8:39–50.
- 4 Muro-Pastor AM, Hess WR. 2012. Heterocyst differentiation: from single mutants to global approaches. *Trends Microbiol.* 20: 548–557.
- 5 Zeng X, Zhang CC. 2022. The making of a heterocyst in cyanobacteria. *Annu Rev Microbiol.* 76:597–618.
- 6 Brenes-Álvarez M, et al. 2019. Elements of the heterocyst-specific transcriptome unravelled by co-expression analysis in *Nostoc* sp. PCC 7120. *Environ Microbiol.* 21:2544–2558.
- 7 Mitschke J, Vioque A, Haas F, Hess WR, Muro-Pastor AM. 2011. Dynamics of transcriptional start site selection during nitrogen stress-induced cell differentiation in *Anabaena* sp. PCC7120. *Proc Natl Acad Sci USA.* 108:20130–20135.
- 8 Georg J, Hess WR. 2018. Widespread antisense transcription in prokaryotes. *Microbiol Spectrum.* 6. doi: [10.1128/microbiolspec.RWR-0029-2018](https://doi.org/10.1128/microbiolspec.RWR-0029-2018)
- 9 Flaherty BL, Van Nieuwerburgh F, Head SR, Golden JW. 2011. Directional RNA deep sequencing sheds new light on the transcriptional response of *Anabaena* sp. strain PCC 7120 to combined-nitrogen deprivation. *BMC Genomics.* 12:332.
- 10 Olmedo-Verd E, Brenes-Álvarez M, Vioque A, Muro-Pastor AM. 2019. A heterocyst-specific antisense RNA contributes to metabolic reprogramming in *Nostoc* sp. PCC 7120. *Plant Cell Physiol.* 60:1646–1655.
- 11 Brenes-Álvarez M, Minguet M, Vioque A, Muro-Pastor AM. 2020. NsiR1, a small RNA with multiple copies, modulates heterocyst differentiation in the cyanobacterium *Nostoc* sp. PCC 7120. *Environ Microbiol.* 22:3325–3338.
- 12 Johnson GE, Lalanne JB, Peters ML, Li GW. 2020. Functionally uncoupled transcription-translation in *Bacillus subtilis*. *Nature.* 585: 124–128.
- 13 Kelly CL, Taylor GM, Šatkutė A, Dekker L, Heap JT. 2019. Transcriptional terminators allow leak-free chromosomal integration of genetic constructs in cyanobacteria. *Microorganisms.* 7:263.
- 14 Dugar G, et al. 2013. High-resolution transcriptome maps reveal strain-specific regulatory features of multiple *Campylobacter jejuni* isolates. *PLoS Genet.* 9:e1003495.
- 15 Bischler T, Tan HS, Nieselt K, Sharma CM. 2015. Differential RNA-seq (dRNA-seq) for annotation of transcriptional start sites and small RNAs in *Helicobacter pylori*. *Methods.* 86:89–101.
- 16 Sharma CM, et al. 2010. The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature.* 464:250–255.
- 17 Yu SH, Vogel J, Förstner KU. 2018. ANNOgesic: a Swiss army knife for the RNA-seq based annotation of bacterial/archaeal genomes. *GigaScience.* 7:1–11.
- 18 Georg J, et al. 2009. Evidence for a major role of antisense RNAs in cyanobacterial gene regulation. *Mol Syst Biol.* 5:305.
- 19 Merino-Puerto V, Herrero A, Flores E. 2013. Cluster of genes that encode positive and negative elements influencing filament length in a heterocyst-forming cyanobacterium. *J Bacteriol.* 195: 3957–3966.
- 20 Yadav S, et al. 2022. Cyclophilin *anaCyp40* regulates photosystem assembly and phycobilisome association in a cyanobacterium. *Nat Commun.* 13:1690.
- 21 Muro-Pastor AM. 2014. The heterocyst-specific NsiR1 small RNA is an early marker of cell differentiation in cyanobacterial filaments. *mBio.* 5:e01079-14.
- 22 Brenes-Álvarez M, Vioque A, Muro-Pastor AM. 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.* 11:e02599-19.
- 23 Brenes-Álvarez M, Vioque A, Muro-Pastor AM. 2022. The heterocyst-specific small RNA NsiR1 regulates the commitment to differentiation in *Nostoc*. *Microbiol Spectrum.* 10:e0227421.
- 24 Brenes-Álvarez M, Olmedo-Verd E, Vioque A, Muro-Pastor AM. 2021. A nitrogen stress-inducible small RNA regulates CO<sub>2</sub> fixation in *Nostoc*. *Plant Physiol.* 187:787–798.
- 25 Álvarez-Escribano I, et al. 2021. NsiR3, a nitrogen stress-inducible small RNA, regulates proline oxidase expression in the cyanobacterium *Nostoc* sp. PCC 7120. *FEBS J.* 288:1614–1629.
- 26 Thomason MK, et al. 2015. Global transcriptional start site mapping using differential RNA sequencing reveals novel antisense RNAs in *Escherichia coli*. *J Bacteriol.* 197:18–28.
- 27 Lasa I, et al. 2011. Genome-wide antisense transcription drives mRNA processing in bacteria. *Proc Natl Acad Sci U S A.* 108: 20172–20177.
- 28 Voigt K, et al. 2014. Comparative transcriptomics of two environmentally relevant cyanobacteria reveals unexpected transcriptome diversity. *ISME J.* 8:2056–2068.
- 29 Bidnenko V, et al. 2017. Termination factor Rho: from the control of pervasive transcription to cell fate determination in *Bacillus subtilis*. *PLoS Genet.* 13:e1006909.
- 30 Peters JM, et al. 2012. Rho and NusG suppress pervasive antisense transcription in *Escherichia coli*. *Genes Dev.* 26:2621–2633.
- 31 Mella-Herrera RA, Neunuebel MR, Golden JW. 2011. *Anabaena* sp. strain PCC 7120 *conR* contains a LytR-CpsA-Psr domain, is developmentally regulated, and is essential for diazotrophic growth and heterocyst morphogenesis. *Microbiology.* 157:617–626.
- 32 Thomason MK, Storz G. 2010. Bacterial antisense RNAs: how many are there, and what are they doing? *Annu Rev Genet.* 44: 167–188.
- 33 Lejars M, Kobayashi A, Hajnsdorf E. 2019. Physiological roles of antisense RNAs in prokaryotes. *Biochimie.* 164:3–16.
- 34 Lasa I, Villanueva M. 2014. Overlapping transcription and bacterial RNA removal. *Proc Natl Acad Sci U S A.* 111:2868–2869.
- 35 Bharati BK, et al. 2022. Crucial role and mechanism of transcription-coupled DNA repair in bacteria. *Nature.* 604: 152–159.
- 36 Wurtzel O, et al. 2012. Comparative transcriptomics of pathogenic and non-pathogenic *Listeria* species. *Mol Syst Biol.* 8:583.
- 37 Ziegler K, et al. 1998. Molecular characterization of cyanophycin synthetase, the enzyme catalyzing the biosynthesis of the cyanobacterial reserve material multi-L-arginyl-poly-L-aspartate (cyanophycin). *Eur J Biochem.* 254:154–159.
- 38 Shabestary K, et al. 2018. Targeted repression of essential genes to arrest growth and increase carbon partitioning and biofuel titers in cyanobacteria. *ACS Synth Biol.* 7:1669–1675.
- 39 Klähn S, et al. 2021. Integrative analysis of the salt stress response in cyanobacteria. *Biol Direct.* 16:26.
- 40 Dühring U, Axmann IM, Hess WR, Wilde A. 2006. An internal antisense RNA regulates expression of the photosynthesis gene *isiA*. *Proc Natl Acad Sci U S A.* 103:7054–7058.

- 41 Sakurai I, et al. 2012. Positive regulation of *psbA* gene expression by cis-encoded antisense RNAs in *Synechocystis* sp. PCC 6803. *Plant Physiol.* 160:1000–1010.
- 42 Eisenhut M, et al. 2012. The antisense RNA *As1\_flv4* in the cyanobacterium *Synechocystis* sp. PCC 6803 prevents premature expression of the *flv4-2* operon upon shift in inorganic carbon supply. *J Biol Chem.* 287:33153–33162.
- 43 Curatti L, Giarocco L, Salerno GL. 2006. Sucrose synthase and RuBisCo expression is similarly regulated by the nitrogen source in the nitrogen-fixing cyanobacterium *Anabaena* sp. *Planta.* 223: 891–900.
- 44 Zhang SR, Lin GM, Chen WL, Wang L, Zhang CC. 2013. Pgpmp metabolism is involved in heterocyst development in the cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol.* 195:4536–4544.
- 45 Hernández JA, et al. 2006. Identification of a *furA* cis antisense RNA in the cyanobacterium *Anabaena* sp. PCC 7120. *J Mol Biol.* 355:325–334.
- 46 Sesto N, Wurtzel O, Archambaud C, Sorek R, Cossart P. 2013. The *excludon*: a new concept in bacterial antisense RNA-mediated gene regulation. *Nat Rev Microbiol.* 11:75–82.
- 47 Muro-Pastor AM, Brenes-Álvarez M, Vioque A. 2017. A combinatorial strategy of alternative promoter use during differentiation of a heterocystous cyanobacterium. *Environ Microbiol Rep.* 9: 449–458.
- 48 Rippka R, Deruelles J, Waterbury JB, Herdman M, Stanier RY. 1979. Generic assignments, strain stories and properties of pure cultures of cyanobacteria. *J Gen Microbiol.* 111:1–61.
- 49 Sambrook JF, Russell DW. 2001. *Molecular cloning: a laboratory manual*. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory.
- 50 Mohamed A, Jansson C. 1989. Influence of light on accumulation of photosynthesis-specific transcripts in the cyanobacterium *Synechocystis* 6803. *Plant Mol Biol.* 13:693–700.
- 51 Brenes-Álvarez M, Olmedo-Verd E, Vioque A, Muro-Pastor AM. 2016. Identification of conserved and potentially regulatory small RNAs in heterocystous cyanobacteria. *Front Microbiol.* 7:48.
- 52 Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol.* 37:907–915.
- 53 Anders S, Pyl PT, Huber W. 2015. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics.* 31: 166–169.
- 54 Robinson MD, McCarthy DJ, Smyth GK. 2010. Edger: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics.* 26:139–140.
- 55 Ritchie ME, et al. 2015. *Limma* powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43:e47.
- 56 Kopf M, et al. 2014. Comparative analysis of the primary transcriptome of *Synechocystis* sp. PCC 6803. *DNA Res.* 21:527–539.
- 57 Muro-Pastor AM, Valladares A, Flores E, Herrero A. 1999. The *hetC* gene is a direct target of the NtcA transcriptional regulator in cyanobacterial heterocyst development. *J Bacteriol.* 181: 6664–6669.
- 58 Bradford MM. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem.* 72:248–254.
- 59 Srere PA. 1969. Citrate synthase. *Methods Enzymol.* 13:3–11.
- 60 Muro-Pastor AM, Olmedo-Verd E, Flores E. 2006. All4312, an NtcA-regulated two-component response regulator in *Anabaena* sp. strain PCC 7120. *FEMS Microbiol Lett.* 256:171–177.