PLOS COMPUTATIONAL BIOLOGY

RESEARCH ARTICLE

Terminal heterocyst differentiation in the *Anabaena patA* mutant as a result of post-transcriptional modifications and molecular leakage

Pau Casanova-Ferrer[1,2], **Saúl Ares**[1,2]*, **Javier Muñoz-García**[1]*

1 Grupo Interdisciplinar de Sistemas Complejos (GISC) and Departamento de Matemáticas, Universidad Carlos III de Madrid, Leganés, Spain, **2** Centro Nacional de Biotecnologia (CNB), CSIC, Madrid, Spain

* saul.ares@csic.es (SA); jamunoz@math.uc3m.es (JM-G)

Abstract

The *Anabaena* genus is a model organism of filamentous cyanobacteria whose vegetative cells can differentiate under nitrogen-limited conditions into a type of cell called heterocyst. These heterocysts lose the possibility to divide and are necessary for the colony because they can fix and share environmental nitrogen. In order to distribute the nitrogen efficiently, heterocysts are arranged to form a quasi-regular pattern whose features are maintained as the filament grows. Recent efforts have allowed advances in the understanding of the interactions and genetic mechanisms underlying this dynamic pattern. However, the main role of the *patA* and *hetF* genes are yet to be clarified; in particular, the *patA* mutant forms heterocysts almost exclusively in the terminal cells of the filament. In this work, we investigate the function of these genes and provide a theoretical model that explains how they interact within the broader genetic network, reproducing their knock-out phenotypes in several genetic backgrounds, including a nearly uniform concentration of HetR along the filament for the *patA* mutant. Our results suggest a role of *hetF* and *patA* in a post-transcriptional modification of HetR which is essential for its regulatory function. In addition, the existence of molecular leakage out of the filament in its boundary cells is enough to explain the preferential appearance of terminal heterocysts, without any need for a distinct regulatory pathway.

Author summary

Understanding multicellular pattern formation is key for the study of both natural and synthetic developmental processes. Arguably one of the simplest model systems for this is the filamentous cyanobacterium *Anabaena*, that in conditions of nitrogen deprivation undergoes a dynamical differentiation process that differentiates roughly one in every ten cells into nitrogen-fixing heterocysts, in a quasi-regular pattern that is maintained as the filament keeps growing. One of the most characteristic mutations affecting this process forms heterocysts mostly constrained to the terminal cells of the filament. We have used

Citation: Casanova-Ferrer P, Ares S, Muñoz-García J (2022) Terminal heterocyst differentiation in the *Anabaena patA* mutant as a result of post-transcriptional modifications and molecular leakage. PLoS Comput Biol 18(8): e1010359. https://doi.org/10.1371/journal.pcbi.1010359

Editor: Tobias Bollenbach, Universitat zu Koln, GERMANY

Received: January 17, 2022

Accepted: July 5, 2022

Published: August 15, 2022

Data Availability Statement: We have stated that the code of our model is available at: https://github.com/PauCasanova/AnabaenaPLOS.

Funding: This research was supported by MCIN/AEI/10.13039/501100011033/ and FEDER Una manera de hacer Europa through grant no. FIS2016-78313-P to S.A and the associated, and MCIN/AEI/10.13039/501100011033 through grant BADS, no. PID2019-109320GB-100, to S.A. and J. M-G. The Spanish MICINN has also funded the "Severo Ochoa" Centers of Excellence to CNB, SEV

2017-0712. P.C-F is hired through the FPI contract BES-2017-079755 associated to the grant FIS2016-78313-P. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

experimental observations to propose a mathematical model of heterocyst differentiation able to reproduce this striking phenotype. The model extends our understanding of the regulations in this pattern-forming system and makes several predictions on molecular interactions. Importantly, a key aspect is the boundary condition at the filament's ends: inhibitors of differentiation should be able to leak out of the filament, or otherwise the terminal cells would not differentiate. This highlights, in a very clear example, the importance of considering physical constraints in developmental processes.

Introduction

The biology of cyanobacteria has been the subject of intensive work during the last decades. A great number of studies have focused on a specific type of cyanobacteria that forms colonies consisting of one-dimensional filaments, the strain PCC 7120 of the genus *Anabaena*, to such an extent that this has become a model organism in the field [1–3]. Under nitrogen-rich conditions, these filaments are composed only of vegetative cells carrying photosynthesis. However, as a response to different environmental stresses, vegetative cells can differentiate into specialized cell types that can fix atmospheric nitrogen. These individual conversions allow the survival of the colony and represent a paradigmatic example of a prokaryotic life form with differentiated cell types. Furthermore, bacteria and archaea are the only organisms capable of fixing atmospheric nitrogen, which makes them crucial for the viability of all living beings on Earth. The process used during nitrogen fixation in cyanobacteria is catalyzed by nitrogenase, and this enzyme is easily degraded by oxygen. In order to avoid this degradation, some cyanobacteria have developed a mechanism to protect nitrogenase from the oxygen produced by vegetative cells. This mechanism is the specialization of some cells into a form denominated heterocysts which accumulates nitrogenase and does not carry out photosynthesis. When external nitrogen sources are scarce, heterocysts appear in a quasi-regular pattern, with intervals of roughly ten vegetative cells between heterocysts. Fixed nitrogen produced by the heterocysts reaches the vegetative cells and sustains their growth. Reciprocally, nutrients produced by photosynthesis in vegetative cells are also shared to maintain the production of nitrogenous compounds in heterocysts, which require high energy consumption [1, 4]. Upon differentiation, heterocysts lose the possibility to undergo cell division, but vegetative cells continue dividing, producing filament growth and increasing the distance between consecutive heterocysts. Due to this, new heterocysts differentiate in the middle of the intervals between previously existing heterocysts in order to not diminish the supply of nitrogen to distant vegetative cells. In this way, the dynamic process of differentiation allows the overall pattern of heterocyst location to conserve its properties over time.

There are a large number of processes involved in the regulation of heterocyst pattern formation and its maintenance. In addition to nitrogen levels and other environmental aspects, many genes and transcription factors play a role [5]. It has been shown that, when nitrogen stress is perceived, the transcription regulator *ntcA* is important to trigger heterocyst differentiation [6, 7], by directly or indirectly controlling the expression of several genes including *hetR* [8]. The expressions of *ntcA* and *hetR* are mutually dependent, although the latter seems to be sufficient for heterocyst development [9]. Thus, positive auto-regulation of *hetR* is required for differentiation and is particularly significant in developing heterocysts [10, 11]. *hetR* expression is the main positive regulatory factor in heterocyst development [9, 10] and this gene is epistatic to the others involved in heterocyst differentiation [12].

The gene *patS* is a negative regulator of *hetR*, suppressing differentiation when overexpressed and inducing multiple contiguous heterocysts, the so-called Mch phenotype, when deleted [13, 14]. The expression of *patS* produces a short peptide PatS, predicted to be formed by 13 or 17 amino acids, which contains a carboxyl-terminal that prevents DNA binding activity of HetR [15, 16] and inhibits differentiation when added to culture medium [13]. The expression of *patS* in small groups of cells was shown to diminish the levels of HetR in adjacent cells [17], suggesting that a PatS-dependent signal can be trafficked along the filament [18]. Recent studies have also proven a redundancy in the inhibitory signal of *patS* through the gene *patX* [19, 20]. Despite acting in an analogous way to *patS*, *patX* seems to play a secondary role complementing the main signal produced by *patS*. As shown in [20], the Δ*patX* single mutant does not present an altered phenotype while the Δ*patX*Δ*patS* double mutant displays a much higher percentage of heterocysts than the Δ*patS* single mutant.

Although lack of *patS* expression initially produces a pattern with groups of contiguous heterocysts and short intervals of vegetative cells between those clusters, this pattern tends to recover the characteristics of a wild type-like pattern later on [14]. This suggests the presence of additional patterning signals operating long after nitrogen deprivation. The most relevant player that leads to this late inhibitory effect is the *hetN* gene, expressed only in heterocysts. Similarly to *patS*, the product codified by *hetN* also contains an ERGSGR sequence of aminoacids (Glutamate-Arginine-Glycine-Serine-Glycine-Arginine), raising the possibility that an ERGSGR-containing peptide derived from the full protein goes from cell to cell [21, 22]. However, in contrast to the *patS* mutant phenotype, the *hetN* mutant phenotype has a heterocyst pattern similar to the wild type at the initial stages of nitrogen depletion and a multiple contiguous heterocyst phenotype after 48 hours [23]. This indicates that *hetN* expression is activated later than that of *patS*. Additional proof of the inhibitory function of *patS* and *hetN* is that, when both genes are suppressed, almost all cells eventually differentiate, causing lethal levels of heterocysts [24].

Experimental results, such as the previously described and more recent ones with transcriptional information [25], have allowed advancing in the understanding of the mechanisms and interactions between *hetR*, *patS*, and *hetN* that give rise to the appearance and maintenance of heterocyst patterns. However, in addition to these, other transcription factors such as *patA* and *hetF* have been shown to play an important role at the early steps of differentiation, regulating the transcriptional activity of *hetR* [26, 27]. All this complex network of interactions, where other heterocyst-related genes, such as *hetC*, *hetP*, *hetL*, *patN*, and *hetZ*, also play a role, has made the complete understanding of heterocyst differentiation a challenge during the last two decades.

In this work, we investigate the function of *patA* and *hetF* together with their interactions with the main genes responsible for heterocyst pattern formation. In the next section we review the main experimental results regarding these genes. Based on these findings we propose a theory combining genetic, metabolic, and morphological aspects. The proposed mathematical model reproduces the diverse experimental phenotypes and explains the main function of both *patA* and *hetF* in the gene-regulatory network of heterocyst differentiation.

Experimental evidence about *patA* and *hetF*

This two genes seem to have a clear role in pattern formation given the heterocyst distribution on the filament is heavily distorted in their deletion mutants. A *patA* mutation suppresses heterocyst differentiation primarily in the intercalary cells, in most cases forming only a pair of heterocysts at both ends of the filament. The fact that *patA* seems to be required for the differentiation of intercalary heterocysts but not for terminal heterocysts have made some authors

think that a different differentiation process in which *patA* is not involved could occur depending on cell position [12]. Surprisingly, even though rare intercalary heterocysts are formed in strains lacking the *patA* gene, high levels of HetR (higher than in the wild type) are measured 18 hours after removing combined nitrogen [26, 28]. Regarding the connection between *patA* and the master regulator in heterocyst differentiation, *hetR*, multiple contiguous heterocysts appear when *hetR* is overexpressed [29]. However, this differentiation is mostly suppressed in the *patA* mutant, for which the same *patA* phenotype with only terminal heterocysts is obtained even under nitrogen starvation. At the same time, in [30] it was observed that the *patA* transcription is greatly reduced in strains for which the expression of *hetR* is blocked. Therefore, the activation of *patA* expression seems to be directly upregulated by HetR.

More insights into the functional relationship between *patA* and other genes involved in heterocyst differentiation are presented in [12]. This work studies the connections between *patA* and *hetR*, *patS*, and *hetN*, analyzing the single, double, and triple mutant phenotypes. For the *patA* mutant background with *hetN* inactivated, they obtain a phenotype indistinguishable from the *patA* mutant with single terminal heterocysts at 24 hours post-induction. However, an increasing number of contiguous heterocysts are formed mostly at the ends of the filament after that time. In the case of the Δ*patA* Δ*patS* double mutant, its phenotype is identical to the Δ*patS* single mutant during short times after induction. This seems to imply that a functional *patS* gene is required to obtain a *patA* phenotype. However, after 48 hours the average distance between heterocysts for the Δ*patA*Δ*patS* double mutant is larger than in the Δ*patS* single mutant and the wild type.

Based on the previous results, it is suggested that PatA might reduce the efficiency of the inhibitory function of both PatS and HetN. This effect could be achieved in two general ways. The first one would be that PatA interacts with PatS and HetN to reduce its inhibitory potential through a post-transcriptional modification that could be forced degradation, a conformation change, or sequestration. On the other hand, PatA could also interact with HetR to protect it by reducing its sensitivity to inhibition. Nevertheless, given that the Δ*patA*Δ*patS* and the Δ*patA*Δ*patS*Δ*hetN* mutants do not present the same phenotype of the Δ*patS* and the Δ*patS*Δ*hetN* mutants, PatA must have another effect besides the protection of HetR to the inhibition through PatS and HetN. The Δ*patA*-like phenotype obtained for an isolated allele of *hetR* made the authors in [12] suggest that *patA* might also promote differentiation independently from its effects on *patS* and *hetN*.

A study by Risser and Callahan [26] shows that the deletion of *hetF* in *Anabaena* sp. PCC 7120 produces enlarged vegetative cells (with a morphology similar to *patA* overexpression) that do not differentiate into heterocysts even after several days on nitrogen starvation. When *hetF* is overexpressed, vegetative cells become significantly smaller than those in the wild type, and multiple contiguous heterocysts (the Mch phenotype) is induced 24h after nitrogen stepdown.

Their deletion mutants (and the double mutant) present similar high levels of HetR. However, the addition of an ectopic functional HetF reverts the phenotype of all these mutants to the wild type. Furthermore, the regulatory effect of those genes on *hetR* seems to be post-transcriptional. Both *patA* and *hetR* are necessary for the aberrant cell morphology of the Δ*hetF* mutant and the addition of extra copies of *hetF* can functionally bypass the deletion of *patA* without requiring a direct interaction of PatA with *hetF*. Finally *hetR* self-regulation and *patS* upregulation through HetR depend on *hetF*. These results lead the authors of [26] to suggest the existence of an activation process of HetR controlled by *hetF* which induces the *hetR* regulatory function.

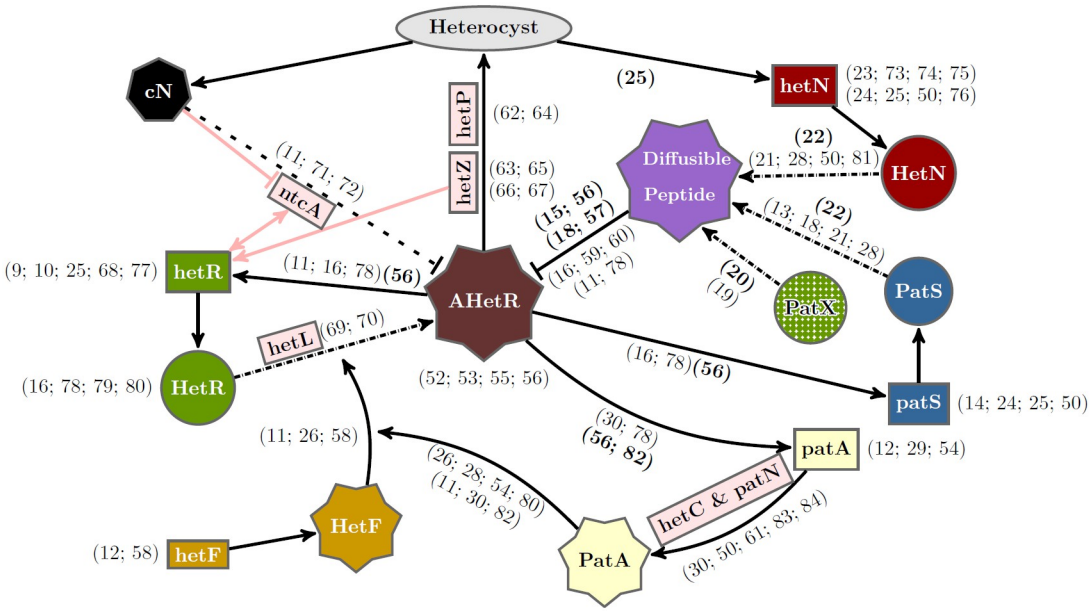

Fig 1. Gene regulatory network of heterocyst differentiation. The main elements involved and their interactions for heterocyst differentiation are depicted schematically together with other relevant genes (light red rectangles) not included in our model. Rectangles, circles, and polyhedric forms represent genes, inactive proteins, and active products, respectively. AHetR stands for the active form of HetR. The ellipse represents differentiation into a heterocyst. Arrows with solid lines represent interactions between elements. Arrows with dashed-dotted lines represent post-transcriptional changes. The dashed line represents a simplified interaction of nitrogen sensing through *ntcA*. References justifying these interactions are incorporated. Regular and bold formatted references indicate phenotypically inferred and observed molecular interactions, respectively.

https://doi.org/10.1371/journal.pcbi.1010359.g001

Materials and methods

Mathematical model

A number of works have presented mathematical models of heterocyst differentiation [31–44]. Here we have used the interactions explained below in the Gene regulatory network section and depicted in Fig 1 to formulate our model schematically presented in Fig 2. Based on these interactions we formulated a set of differential equations for the evolution of the species involved in heterocyst differentiation. Assuming that protein interactions are much faster than the production of those proteins, equilibrium states were considered for the reactions with a shorter timescale. Applying these simplifications we get a more manageable mathematical model to describe the temporal evolution of the concentration of the main protein monomers. Thus, the temporal evolution equation for the main species are

$$\frac{dR_j}{dt} = \beta_R + g(R_j, A_j, I_j, G_i)\rho_R - \alpha_R R_j(1 + 2\mu R_j) \quad (1)$$

$$\frac{dA_j}{dt} = g(R_j, A_j, I_j, G_i)\rho_A - \alpha_A A_j \quad (2)$$

$$\frac{dS_j}{dt} = (1 - \delta_{Hc,j})g(R_j, A_j, I_j, G_i)\rho_S - 2c_S S_j - \alpha_S S_j \quad (3)$$

$$\frac{dN_j}{dt} = \delta_{Hc,j}\rho_N - 2c_N N_j - \alpha_N N_j \quad (4)$$

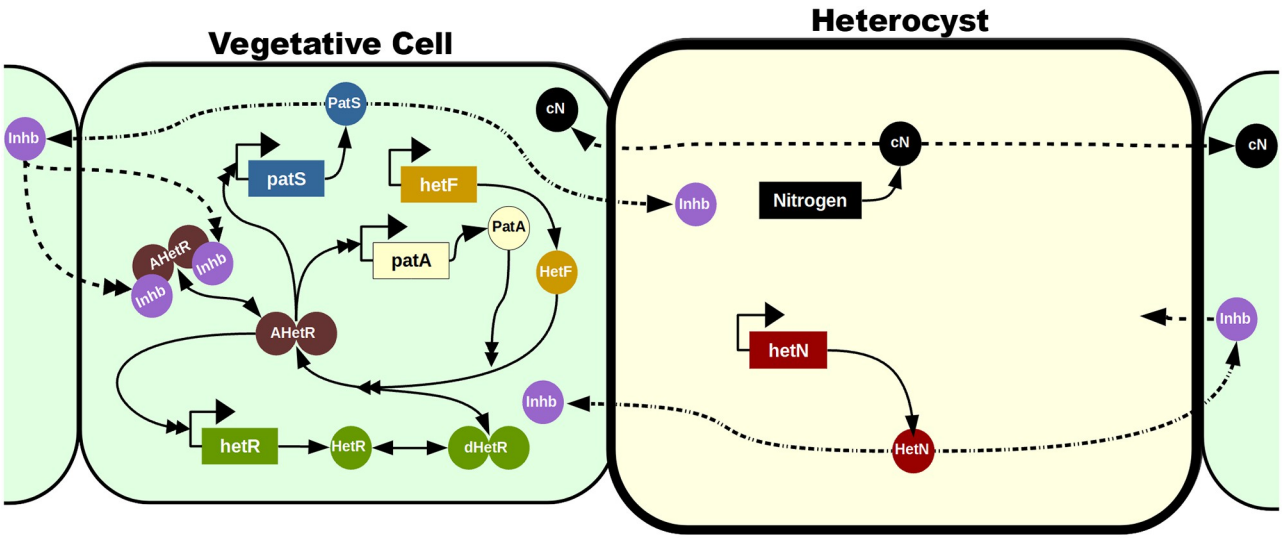

Fig 2. Mechanistic model. The vegetative cells are represented with a soft green background while the heterocyst has a soft yellow background and a thicker cell wall. Genes are represented in rectangles and proteic elements with circles. The dimers are represented with two attached circles and can be inactivated (in green), activated (in brown), and activated and inhibited (in brown with two attached purple inhibitors). Solid lines represent production (with only one simple arrowhead), transformations (with a simple arrowhead in both ends), and interactions (with a double arrowhead). Dashed lines represent inter-cellular traffic and dashed-doted lines represent a transformation when exported to a neighboring cell.

https://doi.org/10.1371/journal.pcbi.1010359.g002

$$\frac{dI_j}{dt} = c_S(S_{j-1} + S_{j+1}) + c_N(N_{j-1} + N_{j+1}) + d_I(I_{j-1} - 2I_j + I_{j+1}) - \alpha_I I_j \tag{5}$$

$$\frac{dG_j}{dt} = \delta_{Hc,j}\rho_G + d_G(G_{j-1} - 2G_j + G_{j+1}) - \alpha_G G_j \tag{6}$$

$$g(R_j, A_j, I_j, G_i) = \frac{F_R R_j^2\left(1 + \frac{A_j}{\tau_A}\right)}{1 + F_R R_j^2\left(1 + \frac{A_j}{\tau_A}\right) + \frac{I_j^2}{K_d^2} + \frac{G_j}{K_G}}, \tag{7}$$

where the subindex j indicates that the variable refers to the cell number j in the filament (being then $j - 1$ and $j + 1$ its neighboring cells) and t denotes time. The concentration of the protein monomers are represented by R_j, A_j, S_j, and N_j, which stand for the concentration of HetR, PatA, the addition of both PatS and PatX, and HetN, respectively, in the cell j. We consider the same functional mobile form of the inhibitor for all the inhibitory genes considered: *patS*, *hetN* and *patX*. This inhibitor is the hexapeptide ERGSGR [22], represented by I_j in our model. We consider that the hexapeptide is produced as a modification of PatS, HetN or PatX at the cell membrane, with characteristic rates for each protein. The product of the modification is exported to the neighbor cell. Finally, G_j represents the concentration of fixed nitrogen in cell j. The factor $\delta_{Hc,j}$ specifies whether the cell is vegetative or a heterocyst: its value is 1 if cell j is a heterocyst, 0 if it is vegetative. This value is used as a switch for the production of both HetN and fixed nitrogen which are only produced in heterocysts and the production of PatS which is only produced in heterocysts.

Assuming that HetF is produced at a constant basal rate and degraded linearly, we have simplified the model considering an equilibrium concentration of HetF (F_{eq}) following

$$\frac{dF_j}{dt} = \beta_F - \alpha_F F_j = 0,$$

from where we get a constant concentration of HetF at any cell: $F_{eq} = \beta_F/\alpha_F$. Given the experimental observation that vegetative cells in *hetF* mutants have morphological phenotypes, it is probable that *hetF* has an additional role in cell morphology regulation. However, since we are focused on studying heterocyst differentiation, regulation of both HetR and HetF in heterocysts is irrelevant for our purposes and is not considered in the model.

For simplicity, we have considered basal production only for HetR monomers and HetF proteins with rates β_R and β_F respectively. The maximum regulated production rates are represented by ρ_R, ρ_A, ρ_S, ρ_N, and ρ_G, for HetR, PatA, both PatS and PatX, HetN, and fixed nitrogen, respectively. The linear degradation and dilution rates are α_R, α_F, α_A, α_S, α_N, α_I, and α_G, since dilution and degradation act in the same direction, these constants combine the effects of these two mechanisms. The active transport rates or diffusion between adjacent cells are c_S, c_N, d_I, and d_G. We assume the border cells at the filament's ends leak both the inhibitor and the fixed nitrogen to the exterior at a lower rate than the communication between neighboring cells. Since molecular trafficking between neighboring cells may depend on elements sitting on both cell membranes, something not possible in the border cells, we have reasoned that it might be possible that leakage to the exterior is less efficient than intercellular trafficking. To allow for this possibility, we have modeled this multiplying the rates d_I, d_G for molecular trafficking between the exterior and the first and last cells of the filament by a factor d_{border}. The exact value of this factor does not have a qualitative effect on the model but can affect the number of terminal heterocysts in some conditions. The resulting equations for I and G in the first and last cells of the filament are:

$$\frac{dI_1}{dt} = c_S S_2 + c_N N_2 - d_{border} d_I I_1 + d_I(I_2 - I_1) - \alpha_I I_1 \tag{8}$$

$$\frac{dG_1}{dt} = \delta_{Hc,1}\rho_G - d_{border} d_G G_1 + d_G(G_2 - G_1) - \alpha_G G_1 \tag{9}$$

$$\frac{dI_N}{dt} = c_S S_{N-1} + c_N N_{N-1} - d_{border} d_I I_N + d_I(I_{N-1} - I_N) - \alpha_I I_N \tag{10}$$

$$\frac{dG_N}{dt} = \delta_{Hc,N}\rho_G - d_{border} d_G G_N + d_G(G_{N-1} - G_N) - \alpha_G G_N. \tag{11}$$

The μ parameter refers to the nonlinear degradation of HetR mediated through its dimerization and could be further expressed as a function of the rates of binding (k_b) and unbinding (k_u) of monomers to form dimers and the degradation rates for both monomers (α_R) and dimers (α_d) of HetR as

$$\mu = \frac{k_b}{\alpha_R}\left(1 - \frac{k_u}{k_u + \alpha_d}\right).$$

We can simplify the system assuming that the promoter regulation through HetR is equivalent for *hetR*, *patA* and *patS*. This regulation is modeled using the factor $g(R_j, A_j, I_j, G_i)$. Eq 7 represents the equilibrium state of the processes of dimerization, activation, and inhibition of HetR. To obtain this expression we have considered the following equilibrium constants: K_R for the dimerization of HetR, K_F for the activation of the HetR dimers by HetF, τ_A for the activation mediated through PatA, and K_d and K_G for the inhibition through the hexapeptide and the fixed nitrogen respectively. Assuming an equilibrium concentration of HetF, the

expression for this regulatory term is reduced to

$$g(R_j, A_j, I_j, G_i) = \frac{\frac{F_{eq}}{K_F}\frac{R_j^2}{K_R}\left(1+\frac{A_j}{\tau_A}\right)}{1+\frac{F_{eq}}{K_F}\frac{R_j^2}{K_R}\left(1+\frac{A_j}{\tau_A}\right)+\frac{I_j^2}{K_d^2}+\frac{G_j}{K_G}},$$

where we have defined

$$F_R \equiv \frac{F_{eq}}{K_F K_R^2}$$

in order to obtain the simplified version in Eq 7. Our results are robust to the details of this function: for instance, assuming regulation by HetR tetramers by substituting R_j^2 by R_j^4 allows for a similar fit to experimental data. Also, allowing details like affinities or exponents to be different for the different species regulated by HetR would only increase the number of free parameters of the model, so for simplicity and to reduce the risk of overfitting we have made the choice of using the same function in all cases.

Model implementation

We have implemented a code in an object-oriented platform to model both the biochemical interactions which give rise to heterocyst differentiation and filament growth. Each cell of the filament has its own variables representing the cellular size and concentration of considered species. The dynamical equations for the concentrations of ERGSGR inhibitor and fixed nitrogen in each cell are coupled with its adjacent neighbors. The resulting set of equations that controls the filament evolution is the noisy extension of the deterministic system Eqs 1–7. This system of equations has been expanded to the Langevin dynamics in the Itô interpretation [45]. This expansion adds an stochastic term of the form $\omega_i^x(t)\sqrt{P_i^x + |D_i^x|}$ for each cell i and species x, where P_i^x and D_i^x are the sum of production (synthesis) terms and the sum of degradation terms respectively and $\omega_i^x(t)$ is an uncorrelated Gaussian white noise [46]. This noise has zero mean and variance $\langle \omega_i^x(t)\omega_j^x(t')\rangle = \Omega_\Phi \delta_{ij}\delta(t-t')$ and models the intrinsic fluctuations in the genetic dynamics.

To differentiate into a heterocyst, a vegetative cell has to accumulate up a certain level of HetR. This has been implemented with an integration of the value of HetR concentration over time for each vegetative cell, once the value of R_j is above a threshold, T_R^j.

This threshold is cell-specific, being drawn from a Gaussian distribution with mean T_R and variance $T_R^2\Omega_\Phi$. If at any point the value of R_j drops below T_R^j, the integral is reset to zero. Otherwise, if the integral ever reaches a value M_R^j, also a cell-specific Gaussian distributed parameter with mean M_R and variance $M_R^2\Omega_\Phi$, the vegetative cell differentiates into a heterocyst. Given the integrative nature of this process, we have set a minimum time T_{min} necessary to avoid unrealistic sudden differentiation due to spikes in HetR production and properly reflect the extensive biological changes required to obtain a mature functional heterocyst. For simplicity, we have also used Ω_Φ to parametrize variability of T_R and M_R, to avoid defining too many noise-related parameters.

Vegetative cell growth was modeled by a stochastic differential equation for each cell:

$$\frac{d\Lambda_i}{dt} = \lambda\left[1+\omega_i^\Lambda(t)\right], \tag{12}$$

where Λ_i is the size of cell i, λ is a constant growth rate, and $\omega_i^\Lambda(t)$ is an uncorrelated Gaussian white noise with zero mean and variance $\langle \omega_i^\Lambda(t)\omega_j^\Lambda(t')\rangle = \Omega_\Lambda \delta_{ij}\delta(t-t')$, which models the

intrinsic fluctuations in the growth process. Starting from an initial size, each cell evolves following Eq 12 up to a maximum size M_Λ^i, which is a noisy value drawn for each cell from a Gaussian distribution of mean M_Λ and variance $M_\Lambda^2\Omega_\Lambda$. We have used Ω_Λ to parametrize this variance for simplicity, again to avoid having too many parameters describing noisy magnitudes.

When this size M_Λ^i is reached, the vegetative cell divides, producing two new vegetative cells with one-half of its current size and identical protein concentrations. Heterocysts follow the same growth, but once they have reached their maximum size, M_Λ^i, they do not divide and stop growing.

The effect of the noise amplitude can be observed in S1 Fig. For low noise in cellular growth Ω_Λ, oscillations appear in the time evolution of the mean number of vegetative cells between heterocysts; this effect is due to the effect of nearly synchronized cell divisions, and was discussed with detail in [42]. As this noise increases, the system loses synchronization between cells and therefore both the mean vegetative interval and the percentage of heterocysts are more stabilized to roughly the mean value that one observes in the less noisy regimes.

The genetic regulatory noise Ω_Φ affects the heterocyst pattern more than the noise in cell growth. This is because a certain genetic noise is required for the pattern formation. Therefore reducing the noise does not seem to affect the long term stability of the pattern but it delays its inception to much later on. Alternatively increasing the noise does not modify the early pattern but it produces considerably more heterocysts which reduces the overall length of the intervals. Finally if one increases the two noises by the same rate one observes that the growth noise Ω_Λ effect dominates for the low overall noise regime while the regulatory noise Ω_Φ effect is the dominant one.

The code to simulate the model is available at: https://github.com/PauCasanova/AnabaenaPLOS.

Parameter estimation

The parameter values employed in simulation results shown in this work can be found in S1 Table. In order to obtain the parameter values that best fit the experimental data, a custom simulated annealing algorithm [47] was employed parameters above the double line in S1 Table, employing as initial values for the optimization those in the model of [42] when an equivalent parameter exists. The last six parameters (K_d, λ, M_Λ, T_R, M_R, T_{min}) where manually set. The value of the affinity of ERGSGR to HetR, represented by K_d, was taken from [15]. Additionally, we choose both the mean maximum cell size to $M_\Lambda = 4\ \mu$m and the cellular growth rate to $\mu = 0.08\mu m \cdot h^{-1}$ in agreement with filament growth data [48]. Finally, the values of T_R, M_R and T_{min} are fixed beforehand in order to obtain a commitment time of around 8–9h [1]. The selection of T_R would be dependent on the typical concentration of HetR on a vegetative cell that is differentiating into a heterocyst. However, given that this experimental information is not available, we set a value taken directly from [42], and adjust the rest of the parameters around it. If one changes this value and runs again the optimization algorithm one obtains a new set of parameters, for which always the concentration of HetR for the *patA* mutant is uniformly maintained under the threshold, but close enough to allow stochastic formation of heterocysts due to fluctuations. Once the value of T_R is set, M_R is the equivalent of maintaining the HetR concentration T_R over 12h, which is the maximum commitment time reported, and T_{min} is set to 5h to ensure that the concentration is at least this time over T_R and therefore eliminates the excessive relevance that sudden bursts could have on an integrative decision process, given the extensive morphological changes that the differentiation to heterocysts requires.

We have used a a custom simulated annealing algorithm that selected a set of data that minimizes the following energy function:

$$E \equiv \sum_{O}[w_1 \cdot \Delta_{KS}(hist_{Exp}, hist_{Sim})^2 + w_2 \cdot ((m_{Exp} - m_{Sim})^2 + (p_{Exp} - p_{Sim})^2)] + \\ + \sum_{N}[w_3 \cdot \Delta_{KS}(hist_{Exp}, hist_{Sim})^2 + w_4 \cdot p_{Sim}] + \delta_{WT,NoHc}([\bar{HetR}]^{WT}_{Sim} - T_R)^2 \quad (13)$$

where $\Delta_{KS}(hist_{Exp}, hist_{Sim})$ is the Kolmogorov-Smirnov distance [49] between the experimental and simulation histograms, m is the mean distance between heterocysts and p is the mean percentage of heterocysts. These quantities are compared between experiments and simulations. The comparison with [50] is done in the summatory with O which are the datasets of the wild type, Δ*patS* and Δ*hetN* for 24h, 48h and 72h. While the comparison with [12] is done in the summatory N which are the datasets of the Δ*patA* and Δ*patA*Δ*hetN* for 24h, 48h, 72h and 96h. The inclusion of both the means and the distributions of vegetative length intervals allows the algorithm to set the noise in order to imitate the variance of the experimental data. Finally, $\delta_{WT,NoHc}$ is a factor that is 0 except if the wild type does not present heterocysts where 1. Given that all the other energy terms suppose heterocyst formation we included a term that compares the HetR concentration for the wild type with the threshold value T_R in order to drive the system towards a set of parameters compatible with heterocyst formation. The set of parameters presented in this paper has been obtained using the weights (w_1, w_2, w_3, w_4) = (1000, 10, 500, 1), which after some tests produced the best results.

To simulate loss-of-function conditions we have considered the production rates equal to zero, except for of the *patS* loss-of-function, where we have reduced the production rate of *patS* by 90%. The remaining 10% represents the redundant effect still present through the expression of *patX* [19, 20].

Sensitivity analysis

To assess the robustness of our results, we have performed a sensitivity analysis following the approach in [42]. We calculate the sensitivity of the model to a given parameter X by evaluating the observable Y at two points, the wild type value X_0 and the perturbation $X = X_0 + \Delta X$. Using the resulting change in the observable, ΔY, we calculate the sensitivity S_{YX} of the observable to the parameter as:

$$S_{YX} = \frac{\Delta Y / Y}{\Delta X / X}. \quad (14)$$

In S2 Fig we show the sensitivity of the mean distance between contiguous heterocysts to changes of 10% in the parameters. The results are qualitatively similar to our previous work [42], however, we find that the extension of the model has made it even more robust to variations in individual parameters. The largest sensitivity is found when modifying HetR production and degradation, followed by inhibitor degradation, the strength of the inhibition, and PatS production. For all other parameters, the relative changes in mean interval length are much smaller than the relative change in the parameter.

Boundary condition mechanism for the *patA* phenotype

A simple physical analogy of the importance of the boundary condition to understand the *patA* phenotype can be made using the following continuous reaction-diffusion model defined

on a filament of length L where the position is denoted by the coordinate $x \in [0, L]$:

$$\frac{\partial r(x,t)}{\partial t} = \beta_r + f(r,s)\rho_r - \alpha_r r \tag{15}$$

$$\frac{\partial s(x,t)}{\partial t} = f(r,s)\rho_s - \alpha_s s + d\frac{\partial^2 s}{\partial x^2}, \tag{16}$$

where r is the concentration of a non-diffusing activator (HetR), s the concentration of a diffusible inhibitor, β_r a basal production for the activator ($\beta_r > 0$), $f(r, s)$ a smooth regulatory function, monotonically increasing in r and decreasing in s, and the other symbols are parameters. These equations need to be supplemented with boundary conditions for s. If the inhibitor cannot diffuse across the filament's boundaries, the condition is:

$$\left.\frac{\partial s}{\partial x}\right|_{x=0,L} = 0. \tag{17}$$

With this condition, we assume that the system is such that a parameter regime exists where there is a stable homogeneous solution. However, if leakage of the inhibitor through the boundaries is possible, the relevant boundary condition is:

$$s|_{x=0,L} = 0. \tag{18}$$

With this condition and under the same parameter regime, there would be a gradient of inhibitor decreasing from the center to the boundaries of the filament, which in turn would produce a profile of activator with maxima in the boundaries, making them the favored location for heterocyst differentiation. This simple analogy explains the physical mechanism behind the boundary-induced pattern observed in the *patA* mutant and agrees qualitatively with the with the observation in the discrete model.

Statistics

In order to compare the results obtained from our numerical simulations with the experimental data from [12, 50] we have replicated the statistical analysis of these works. In both cases, all the data is aggregated for each experiment and then the averages and standard deviation are calculated between the experimental replicates. Thus, a certain amount of filaments have been aggregated in batches and then averaged to obtain the standard deviation. On the experimental side, in [50], for each strain, 300 cells, or 100 intervals, were counted in three or four independent experiments. In [12], the number of contiguous heterocysts at the ends of 50 filaments is averaged in three experiments. Alternatively, our data was obtained from 15 batches of 10 simulations for filaments with an initial size of 30 cells (consistent with experiments in [51]) which grow to have around 50, 100, 200 cells at 24h, 48h, and 72h respectively. Therefore, we present a bigger data sample than the one considered in [50] and of the same order of magnitude as the one in [12].

Results and discussion

Gene regulatory network

Using the information presented in the Introduction we propose a regulatory network, depicted in Fig 1, that includes the main genes involved in heterocyst differentiation. This genetic network modifies and expands a previous minimal model for the interaction of *hetR*,

patS, and *hetN* [42]. Fig 1 is very rich in information, and a complete discussion is beyond the scope of this work: for instance, we will make no further mention of references [52–65].

This proposal includes the novel gene *patX*. This gene has recently been described as a redundant gene for the inhibitory mechanism of *patS* [19, 20]. It has been also shown that, while its deletion in the wild type background does not alter considerably the phenotype, the double mutant Δ*patS patX* is lethal with an almost complete Mch phenotype [19, 20]. To take this into account in a simple way, we consider that the variable in our model for PatS represents the combined effects of PatS and PatX. The knock-out of *patS* will be modeled reducing 90% the value of the production rate for this variable; the remaining 10% represents the redundant effect of *patX* expression.

We consider the same functional mobile form of the inhibitor for all the inhibitory genes considered: *patS*, *hetN* and *patX*. This inhibitor is the hexapeptide ERGSGR [22]. We consider that the hexapeptide is produced as a modification of PatS, HetN or PatX at the cell membrane, with characteristic rates for each protein. The product of the modification is exported to the neighbor cell. This hexapeptide has shown to have a higher affinity with HetR than the pentapeptide originally proposed [15].

An important novel inclusion in the model we propose here is the requirement of a post-transcriptional transformation of HetR to act as a genetic regulator. This active form of HetR, which we term AHetR, is probably obtained through a phosphorylation process [66–68] and only the active fraction of HetR would contribute to heterocyst differentiation. This could explain the apparent paradox of a higher concentration of HetR with less heterocyst formation in both Δ*hetF* and Δ*patA*. The high concentration of HetR in these mutants would be explained through a higher turnover rate for the activated HetR protein [66]. A possible active form of HetR has been recently suggested in [69], where the authors present evidence for phosphorylation of HetR as crucial for its activity in *Nostoc* sp. PCC 7120. This phosphorylation is shown to require the presence of the Pkn22 kinase but no more information regarding the regulation of this phosphorylation is provided. Here we hypothesize that this would be the role of the genetic pathway controlled by *hetF* which has already been presented as a protease [26] and therefore can be expected to have a role in post-transcriptional modifications. Thus, we assign to *hetF* the role of activator of HetR with the mediation of PatA through a post-transcriptional interaction that would change HetR into its enzymatic form. This hypothesis is inferred by observing changes in the phenotypes of several mutant backgrounds and has been considered in Fig 1. However, the nature of this interaction cannot be confirmed, since the observed phenotype could also be explained through indirect interactions mediated by HetF.

To complete the mechanistic model, one must take into account that different genes are expressed in heterocysts or vegetative cells as depicted in Fig 2. Despite the fact that it has also been observed in a tetramer configuration [67], there is still no experimental evidence that HetR binds DNA as a tetramer. Therefore, it will be assumed that it binds to the DNA only as a dimer [70, 71]. Nevertheless, we have checked that this alteration does not change the qualitative system behavior and the same dynamics can be recovered assuming tetrameric binding with an alternative set of parameter values.

As shown in Fig 2, while *hetF* is produced only constitutively at a low-level [72], *patA* only has a regulated expression that depends on the active form of the dimeric transcription factor AHetR, which also activates its own expression. At the protein level: the HetR dimer needs to be activated by HetF (whose enzymatic activity can be enhanced by PatA) to become AHetR. PatS becomes an inhibitor of the transcription factor by protein transformation during cell to cell transport. The inhibitor thus produced is a small molecule that can move along the filament. *hetN* is expressed basally in heterocysts and becomes an inhibitor of the transcription factor, similar to the PatS product, by protein transformation during cell to cell transport. The

fixed nitrogen products produced by the heterocyst can also move to act as an inhibitor of AHetR.

Using the regulatory logic in Fig 2 we have built a mathematical model of gene regulation on a growing filament of cells, as detailed in Materials and Methods.

Study of the wild type and the Δ*patS* and Δ*hetN* mutants

The statistical distribution of vegetative cell intervals between heterocysts may differ from one experiment to another, as one can notice comparing the results from different authors [12–14, 18, 24, 28, 50, 73, 74]. For consistency, to compare our results with the experimental data for the wild type and both the Δ*patS* and the Δ*hetN* mutants, we consider the relative frequency of vegetative intervals of a given length presented in [50], which is the most recent dataset available and one of the most comprehensive. Despite this we also show in S3 Fig a systematic comparison of the Kolmogorov-Smirnov distance between the distribution of the vegetative interval distribution that our model produces and the ones observed in [14, 50, 74] (which are the experimental datasets considered in [42] and in this same paper).

As presented in the Materials and methods section, in order to simplify the description we have modeled both genes *patX*, *patS* using a single variable. Thus, a complete loss of function of this variable represents the experimental Δ*patX*Δ*patS* mutant. This double mutant induces considerably more heterocysts than the single Δ*patS* mutant, S4 Fig, as observed experimentally [20] (see S1 Movie). We have also simulated the Δ*patX* mutant (see S2 Movie), where the production rate is only 10% of the combined PatS+PatX variable. The phenotype observed (S4 and S5 Figs) is still compatible with the wild type data, as reported in [20], with both slightly shorter vegetative intervals and a higher percentage of heterocysts. In Fig 3 we observe the agreement between the simulations and the experimental data from [50] for wild type and *patS* and *hetN* mutants.

For the wild type, it is well known that roughly one of every ten cells differentiates with a slight increase of this interval length with time (see S3 Movie). In the Δ*hetN* mutant, the initial pattern is similar to the wild type, except for the contiguous heterocysts appearing at a sequential pace (see S4 Movie). The Δ*patS* mutant shows a cluster of heterocysts, which appear simultaneously at short times (see S5 Movie). For longer times, the pattern of heterocysts is more

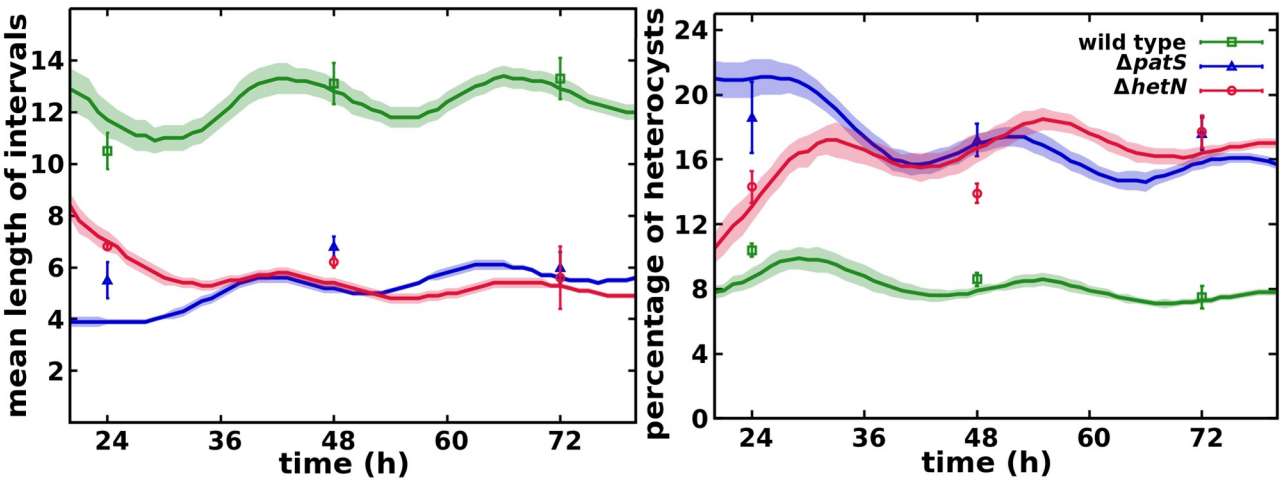

Fig 3. Mean number of vegetative cells between heterocysts and total percentage of heterocysts in the filament. Symbols represent experimental values from [50] and lines are simulation results with their standard deviation as a shadowed area.

https://doi.org/10.1371/journal.pcbi.1010359.g003

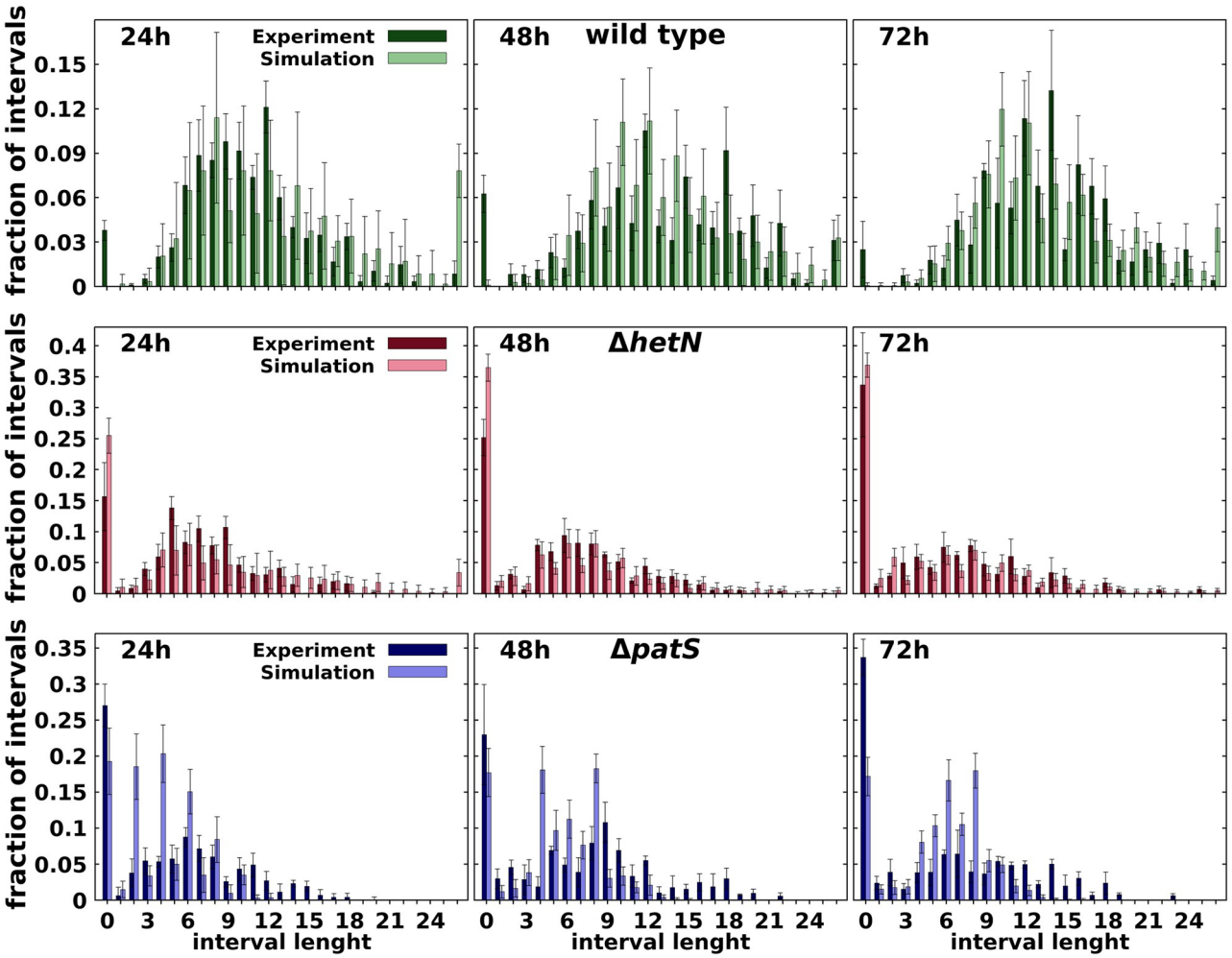

Fig 4. Comparison at different times after nitrogen deprivation (as indicated) between experimental [50] and simulated histograms of the number of vegetative cells between heterocysts for wild type (top row), Δ*hetN* (middle row), Δ*patS* (bottom row). Bars are means of interval lengths, errors are standard deviations. For details of calculations see the Statistics section.

https://doi.org/10.1371/journal.pcbi.1010359.g004

similar to the wild type, but with a higher incidence of contiguous heterocysts, in agreement with the experimental results reported in [14] with a larger statistical sample. This typical phenotype observations are also shown in S6 Fig, where one can observe that both Δ*patS* and Δ*hetN* present multiple clusters of heterocysts. While the distribution for Δ*patS* remains the same through 24 to 72h, the amount of clusters longer than 3 heterocysts increases with time for the Δ*hetN* mutant.

The comparison of the distributions of the number of vegetative cells between heterocysts in the experiment of Ref. [50] and the simulated filaments are shown in Fig 4. The agreement between model and experimental data is very good as one can also observe in S3 Fig, where it is shown that the differences between simulation data and experimental data are of the same order that temporal differences of the same mutant and considerably smaller than the differences between the wild type and both Δ*patS* and Δ*hetN*. Additionally, a closer inspection even reveals that the model produces histograms more similar to the ones presented in [14] than Ref. [50]. A small deviation appears for the early phenotype of the Δ*patS* mutant, especially at 24h. The simulations Δ*patS* of present less contiguous heterocysts and shorter intervals than

both the experimental data (Fig 4) and the Δ*hetN* mutant simulation (S6 Fig) at this time. This difference could be due to the effect of not considering a protoheterocyst phase. Without this phase, the differentiation of adjacent cells is strongly reduced because, once a cell differentiates, it immediately starts producing both HetN and nitrogen products, which inhibit differentiation. Thus, an artificial surplus of one and two cell intervals is observed in the first round of differentiation. After the first round of division, this causes the peaks of two and four cell intervals observed at 24h, Fig 4 and S5 Movie.

The low amount of contiguous heterocysts observed in wild type simulations in comparison with experimental data [50] can be explained using the same argument. A protoheterocyst phase would ease the stochastic formation of multiple simultaneous heterocysts in all genetic backgrounds. On top of that, it is worth noting that these contiguous heterocysts are seldom described in the literature. Their appearance on the wild type can be explained by stochastic fluctuations of the genetic expression that get fixed through the irreversible process that is the differentiation into a heterocyst. In any case, the model also allows for the formation of these contiguous heterocysts, albeit in a much smaller proportion.

These results also improve the phenotypes obtained in the minimal model from [42]. In that model, the Δ*hetN* mutant showed both smaller clusters of heterocysts and shorter intervals between them. Additionally, this previous model could not produce any contiguous heterocysts in the wild type.

Loss of heterogeneity in the HetR concentration profile in a Δ*patA* mutant background

The simulations for the Δ*patA* single mutant show a similar behavior to experimental results [12, 26, 29, 30, 75], with heterocysts forming mostly on the filament ends despite higher global HetR concentration in the filament (approximately 24% higher in our simulations, Fig 5) in Δ*patA* than in the wild type.

The absence of internal heterocysts is caused by the loss of a distinct HetR concentration profile, Fig 6 and S6 Movie. The absence of PatA in the system slows the conversion to active HetR to a minimum level and therefore also the regulatory effect of HetR over both itself and *patS*. This produces a homogenization of the production of PatS and also of the inhibitory hexapeptide along the filament. Then, in the absence of a pronounced inhibitory gradient, the levels of HetR increase uniformly to levels close to the threshold for differentiation (Fig 6). In this conditions, the selection of the few internal cells that will differentiate is exclusively due to stochastic fluctuations on the protein production of both HetF and PatS. On the other hand, the model assumes a certain passive diffusion of both the inhibitory hexapeptide and fixed nitrogen through the filament ends, which causes a local reduction of the inhibitory signals, allowing the differentiation of the boundary cells. If one does not allow the diffusion through the border cells heterocysts do not form in the filament ends as HetR does not accumulate enough on those cells, Fig 6. A simple analogy to a continuous reaction-diffusion system (see Materials and methods) explains how the properties of molecular trafficking at the filament ends, boundary conditions in the mathematical language, can lead to different profiles of heterocyst differentiation.

The results of the simulation for the Δ*patA*Δ*hetN* double mutant also show a phenotypic agreement with observations [12]. The filaments present multiple terminal heterocysts with only the occasional internal heterocyst (S7 Movie). Additionally, the model predicts a higher HetR concentration (approximately 27% higher in our simulations, Fig 5) in Δ*patA*Δ*hetN* than in the wild type, equivalent to the one observed in Δ*patA*. Here the border effect of the inhibitor diffusion through the ends of the filament gets propagated to multiple cells because the

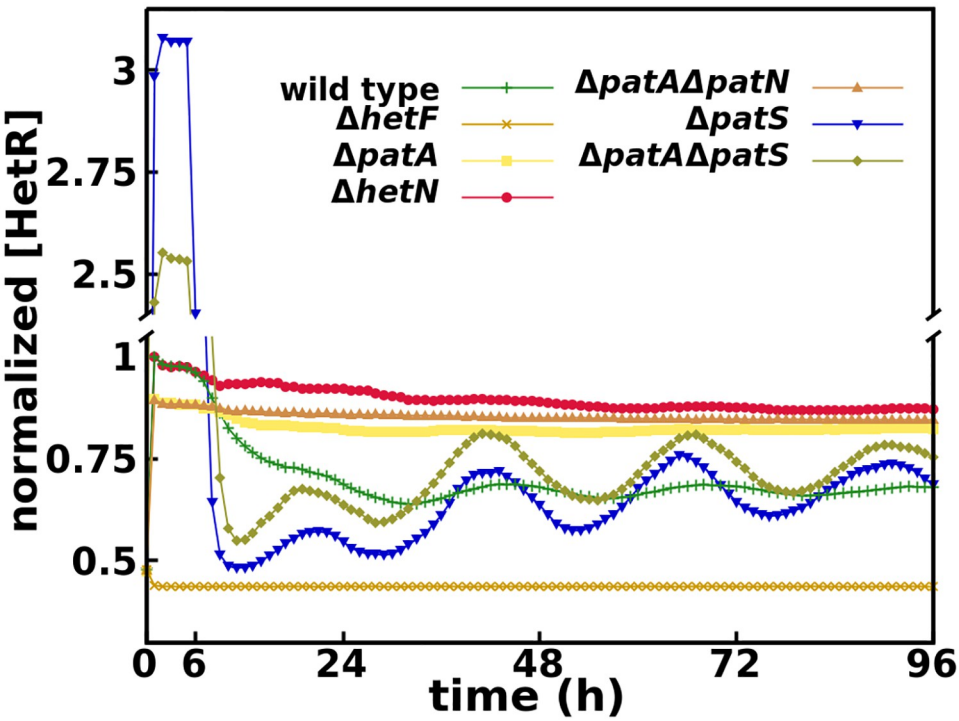

Fig 5. Time evolution of HetR concentration on vegetative cells for all mutants considered, normalized by the maximum wild type concentration.

https://doi.org/10.1371/journal.pcbi.1010359.g005

Δ*patA*Δ*hetN* double mutant, besides the homogenization of the HetR concentration, does not present the inhibitory gradient around existent heterocysts produced by *hetN*.

In Fig 7 we present the temporal evolution of the amount of terminal heterocyst for both Δ*patA* and Δ*patA*Δ*hetN* mutants. These results are in agreement with experimental data [12]

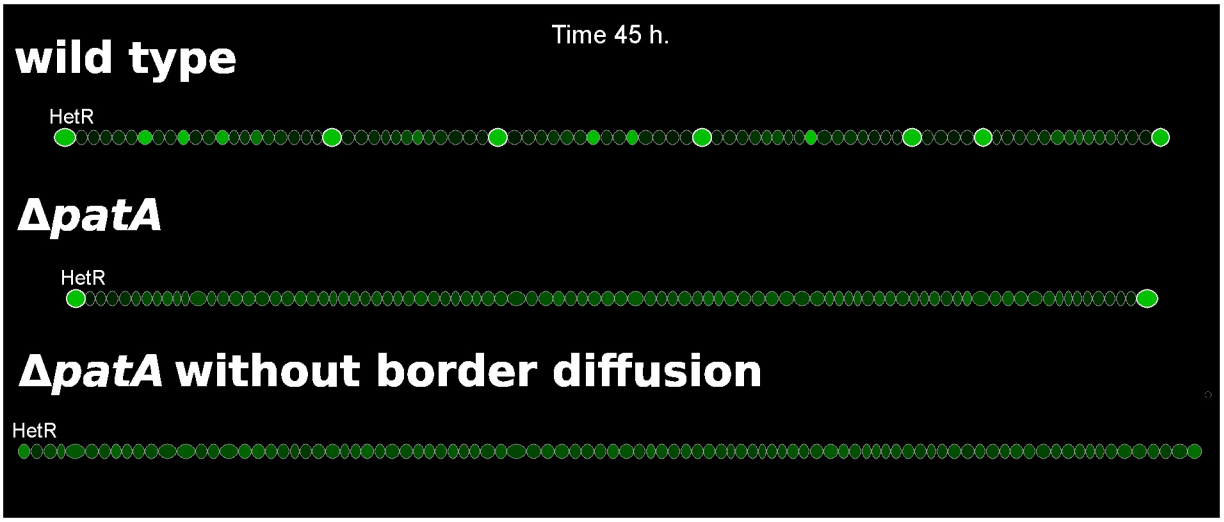

Fig 6. Three simulations showing HetR profiles in filaments of wild type, Δ*patA*, and Δ*patA* with no inhibitor leakage from the terminal cells ($d_{broder} = 0$). HetR concentration is represented by the brightness of green and the heterocysts present an additional white cell wall. For the wild type one can observe cells with high levels of HetR, candidates to differentiate into heterocysts, between existing heterocysts.

https://doi.org/10.1371/journal.pcbi.1010359.g006

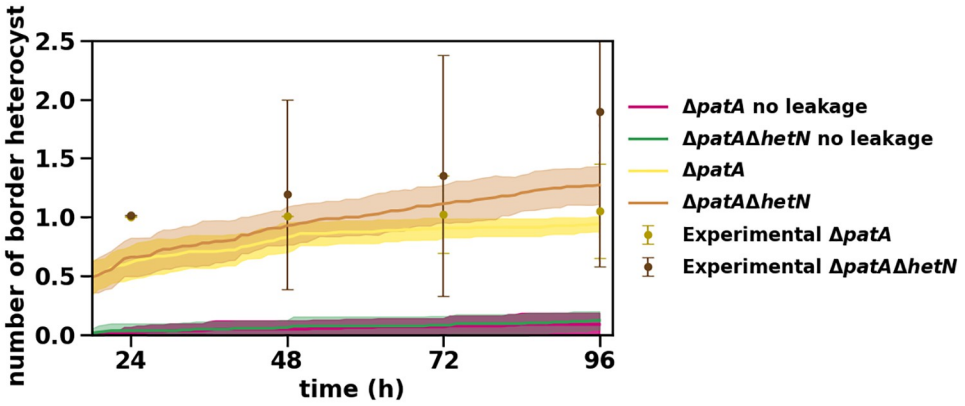

Fig 7. Temporal evolution of the number of heterocysts in the Δ*patA* and Δ*patA*Δ*hetN* mutants. Symbols represent experimental values from [12] and lines are simulation results with their standard deviation as a shadowed area.

https://doi.org/10.1371/journal.pcbi.1010359.g007

as shown in Fig 7 and S7 Fig, where we present the distribution of the number of terminal heterocysts for the Δ*patA*Δ*hetN* mutant. The model seems to have a small delay in the formation of heterocysts (also present in the Δ*patA* mutant), but except for early times the model shows a good agreement with the experimental data. This delay could be related to the mechanism of commitment to the differentiation of a given cell. In our model, this decision is exclusively linked to a sustained high concentration of HetR and not to the expression of a supplementary gene (*hetP* and/or *hetZ*) as recent publications [76–81] seem to indicate. Hence, incorporating a gene-controlled differentiation commitment would surely improve these results, because then self-regulation of that gene could amplify the differentiation signal.

Finally, we studied the Δ*patA*Δ*patS* double mutant, of which to our knowledge there are only phenotypical observations [12]. Its phenotype is described as similar to the single Δ*patS* mutant phenotype, but with longer distances between heterocysts. The simulated filament (S8 Movie) fits this behavior, S4 and S8 Figs. One can notice that the Δ*patA*Δ*patS* mutant presents a higher amount of contiguous heterocysts but with a smaller frequency and larger intervals of vegetative cells between heterocysts. Additionally, the simulated filament does not present an increase in HetR concentration typical to other *patA* mutants. This also confirms that the full Δ*patA* phenotype, which shows almost no internal heterocysts and a high concentration of HetR, requires a functional *patS* gene [12].

As we discussed previously, the increase of the HetR concentration is produced due to the homogenization of the PatS production along the filament, which produces a state where all the cells have a homogeneous HetR concentration lower than the decision threshold. Therefore if *patS* is not functional this effect would not be observed and the weak activation of HetR due to the absence of PatA precludes the formation of as many clusters of heterocysts as one observes in the Δ*patS* mutant.

One can observe this in Fig 5, where the HetR concentration for this mutant does not behave like the Δ*patA* single mutant or the Δ*patA*Δ*hetN* double mutant, with high constant values of HetR. It behaves similarly to the simple *patS* mutant, with a slightly higher concentration than that mutant after the first outburst of HetR that marks the first differentiation round [72]. On the other hand, the concentration is higher in the Δ*patS* single mutant during the first round of differentiation. This inversion is due to the faster production of HetR with a functional *patA*. Therefore, the decision to differentiate is reached in a shorter time than in the Δ*patA*Δ*patS* mutant. Hence, on a homogeneous initial condition (the first round of

differentiation), the concentration would be higher in the Δ*patS* mutant because all cells start producing HetR at a much faster rate. But starting from heterogeneous initial conditions (all the following differentiation rounds), the concentration would be higher in the double mutant. Due to the slower commitment, more cells (closer to the heterocysts) start producing HetR before being shut down by the newly formed heterocysts. This effect can also be observed by comparing S5 and S8 Movies.

Complete loss of HetR regulatory function in the Δ*hetF* mutant background

If HetF is necessary to produce the active form of HetR, in its absence HetR should lose its regulatory function. This prediction is in agreement with experimental observations in *Nostoc punctiforme* [72], which presents a similar transcriptional induction pattern of *hetR* than *Anabaena* sp. PCC 7120. There, the induction of *hetR* is dependent on the presence of an intact copy of *ntcA*. Moreover, the induction of *hetR* is still present in the Δ*hetF* background but with an altered induction pattern that eliminates the 0 to 6h burst.

Our simulated Δ*hetF* successfully eliminates the initial burst of HetR production observed in all other mutants, Fig 5. The absence of this burst in Δ*hetF* implies that it is exclusively produced by the positive self-regulation of *hetR* [26, 72]. Due to this, the introduction of an additional *hetR* promoter activated as a response to nitrogen deprivation would improve the agreement between our model and experiments by increasing only the overall HetR concentration in our simulations (especially on the Δ*hetF* and Δ*patA* simple and multiple mutants) without altering much of the dynamics.

Conclusion

The formation and maintenance of heterocyst patterns is a paradigmatic example in which many processes, such as complex gene regulatory networks, interactions at different time scales, molecular trafficking, and cell growth, act together to give rise to a multicellular pattern. All these aspects form an intriguing puzzle for which a complete understanding is still elusive. A practical way to expand our knowledge about this problem is to investigate what are the functions of some of its pieces. Thus, increasing the complexity of a previous minimal model, we have been able to gain insight into the functions of the players involved. We have focused on the *patA* gene, and based on experimental phenotypical evidence, we have formulated a mathematical model that includes the interactions of this gene with the key genes responsible for heterocyst pattern formation. Our model shows that considering PatA as a collaborator of the activation process of HetR directed by *hetF* is capable of explaining all the phenotypes of the genes considered in our genetic network. This agreement suggests that there is some interaction, direct or indirect, between *hetF* and *patA* that has not been reported experimentally. This consideration, together with the existence of an active form of HetR, is also enough to account for the paradoxical changes in HetR concentration in the *patA* mutant, that seemed to question the central role of *hetR* in heterocyst differentiation.

New experiments are required to confirm the validity of the interactions proposed. In particular, experimental information regarding protein translation [25, 82] could be useful in order to have more detailed information regarding the effects of a given gene on the regulatory network. The roles of many other actors are still to be elucidated and could be included in the core processes to obtain a more extensive mathematical description. For example, recent work [83] presents evidence that *hetL*, a gene previously shown to alter PatS mediated inhibition of heterocyst differentiation [84], interacts with HetR without inhibiting its DNA-binding activity in *Nostoc* sp. PCC 7120. This interaction protects HetR from the inhibitory effects of the

Pat-derived hexapeptide and seems to be essential for the proper function of HetR as a genetic regulator. That would be the role that we have assigned to hetF based on the phenotypic evidence on *Anabaena* sp. PCC 7120.

Finally, our work shows that it is possible to reproduce the *patA* mutant phenotype without considering a differentiation mechanism depending on a cell's position on the filament. The model also expands the characterization presented in [12] by directly linking the increase of the HetR concentration in all the cells with the absence of internal heterocysts in both Δ*patA* and Δ*patA*Δ*hetN* mutants. This is obtained by a slowing of the *hetR* transcription which produces a homogenization of HetR concentration through the filament. Then it is easy to see why this phenotype is not present in the Δ*patA*Δ*patS* mutant where this reduction on the transcription rate is completely compensated by the absence of the early inhibitor PatA. The intriguing differentiation of only terminal heterocysts appears as a consequence of the boundary conditions of the system: leakage of inhibitors out of the filament through the terminal cells. Hence, despite the apparent simplicity of *Anabaena* compared to other developmental systems, it is already clear that genetic and metabolic interactions result in patterns shaped by physical constraints.

Supporting information

S1 Table. Parameter values used for the wild type simulations. The abbreviation DL is used for dimensionless variables.
(PDF)

S1 Fig. Effect of noise on the mean interval of vegetative cells between heterocyst for wild type parameters. Symbols represent experimental values from [50] and lines are simulation results with their standard deviation as a shadowed area.
(TIF)

S2 Fig. Sensitivity of the mean distance between heterocysts after 72h with respect to 10% changes in the indicated parameters. Changes are with respect to the wild type values in S1 Table.
(TIF)

S3 Fig. Kolmogorov–Smirnov distance between histograms of vegetative intervals length obtained with our model and the experimental datasets from [14], [50] and [74]. The X axis presents the common variables between the histograms and the legend the variable that we are comparing. In the first subplot we compare the histograms of a given condition at a given time from two different datasets to evaluate the agreement of our model with the available experimental data. In the second, we obtain the histogram K-S distance between conditions from the same dataset at a given time. And finally in the third subplot we present the temporal change of a given condition for all the datasets considered.
(TIF)

S4 Fig. Mean number of vegetative cells between heterocysts and total percentage of heterocysts in the filament for different conditions, as indicated. Symbols represent experimental values from [50] and lines are simulation results with their standard deviation as a shadowed area.
(TIF)

S5 Fig. Comparison at different times after nitrogen deprivation (as indicated) between simulated histograms of the number of vegetative cells between heterocysts for wild type and Δ*patX*, as indicated. Bars are means of interval lengths, errors are standard deviations.
(TIF)

S6 Fig. Histogram of the number of clusters of a given size observed at 24h, 48h and 72h. Bars are the mean number of clusters of each size at each time, error bars are standard deviation.
(TIF)

S7 Fig. Comparison at different times after nitrogen deprivation (as indicated) between simulated histograms of the number of vegetative cells between heterocysts for wild type, Δ*patS* and Δ*patA*Δ*patS*, as indicated. Bars are means of interval lengths, errors are standard deviations.
(TIF)

S8 Fig. Comparison at different times after nitrogen deprivation (as indicated) between experimental [12] and simulated histograms of the number of heterocysts in the filaments ends for the Δ*patA*Δ*hetN* double mutant. Bars are means of the number of birder heterocysts, errors are standard deviations.
(TIF)

S1 Movie. Temporal evolution of gene expression in a filament for theΔ*patS*Δ*patX* mutant. Time counts hours after nitrogen deprivation. Heterocyst cells have a thicker white membrane. The intensity of the green, purple and cyan colors show the level of HetR, ERGSGR inhibitor, and fixed-nitrogen products concentrations, respectively. When a filament is too long to fit in the width of the movie, it is continued in a row below. The last cell on the right of a row is a neighbor of the first cell on the left of the row immediately below.
(GIF)

S2 Movie. Temporal evolution of gene expression in a filament for the Δ*patX* mutant. Time counts hours after nitrogen deprivation. Heterocyst cells have a thicker white membrane. The intensity of the green, purple and cyan colors show the level of HetR, ERGSGR inhibitor, and fixed-nitrogen products concentrations, respectively. When a filament is too long to fit in the width of the movie, it is continued in a row below. The last cell on the right of a row is a neighbor of the first cell on the left of the row immediately below.
(GIF)

S3 Movie. Temporal evolution of gene expression in a filament for the wild type. Time counts hours after nitrogen deprivation. Heterocyst cells have a thicker white membrane. The intensity of the green, purple and cyan colors show the level of HetR, ERGSGR inhibitor, and fixed-nitrogen products concentrations, respectively. When a filament is too long to fit in the width of the movie, it is continued in a row below. The last cell on the right of a row is a neighbor of the first cell on the left of the row immediately below.
(GIF)

S4 Movie. Temporal evolution of gene expression in a filament for the Δ*hetN* mutant. Time counts hours after nitrogen deprivation. Heterocyst cells have a thicker white membrane. The intensity of the green, purple and cyan colors show the level of HetR, ERGSGR inhibitor, and fixed-nitrogen products concentrations, respectively. When a filament is too long to fit in the width of the movie, it is continued in a row below. The last cell on the right of a row is a neighbor of the first cell on the left of the row immediately below.
(GIF)

S5 Movie. Temporal evolution of gene expression in a filament for the Δ*patS* mutant. Time counts hours after nitrogen deprivation. Heterocyst cells have a thicker white membrane. The intensity of the green, purple and cyan colors show the level of HetR, ERGSGR inhibitor, and

fixed-nitrogen products concentrations, respectively. When a filament is too long to fit in the width of the movie, it is continued in a row below. The last cell on the right of a row is a neighbor of the first cell on the left of the row immediately below.
(GIF)

S6 Movie. Temporal evolution of gene expression in a filament for the Δ*patA* mutant. Time counts hours after nitrogen deprivation. Heterocyst cells have a thicker white membrane. The intensity of the green, purple and cyan colors show the level of HetR, ERGSGR inhibitor, and fixed-nitrogen products concentrations, respectively. When a filament is too long to fit in the width of the movie, it is continued in a row below. The last cell on the right of a row is a neighbor of the first cell on the left of the row immediately below.
(GIF)

S7 Movie. Temporal evolution of gene expression in a filament for the Δ*patA*Δ*hetN* mutant. Time counts hours after nitrogen deprivation. Heterocyst cells have a thicker white membrane. The intensity of the green, purple and cyan colors show the level of HetR, ERGSGR inhibitor, and fixed-nitrogen products concentrations, respectively. When a filament is too long to fit in the width of the movie, it is continued in a row below. The last cell on the right of a row is a neighbor of the first cell on the left of the row immediately below.
(GIF)

S8 Movie. Temporal evolution of gene expression in a filament for the Δ*patA*Δ*patS* mutant. Time counts hours after nitrogen deprivation. Heterocyst cells have a thicker white membrane. The intensity of the green, purple and cyan colors show the level of HetR, ERGSGR inhibitor, and fixed-nitrogen products concentrations, respectively. When a filament is too long to fit in the width of the movie, it is continued in a row below. The last cell on the right of a row is a neighbor of the first cell on the left of the row immediately below.
(GIF)

Acknowledgments

We thank Joel Stavans and Rinat Goren for fruitful discussions, Alicia M. Muro-Pastor for feedback on the preprint version of the work, and Pablo Catalán for help with the simulated annealing algorithm.

Author Contributions

Conceptualization: Saúl Ares, Javier Muñoz-García.

Funding acquisition: Saúl Ares, Javier Muñoz-García.

Investigation: Pau Casanova-Ferrer, Javier Muñoz-García.

Supervision: Saúl Ares, Javier Muñoz-García.

Writing – original draft: Pau Casanova-Ferrer.

Writing – review & editing: Pau Casanova-Ferrer, Saúl Ares, Javier Muñoz-García.

References

1. Muro-Pastor AM, Hess WR. Heterocyst Differentiation: From Single Mutants to Global Approaches. Trends Microbiol. 2012; 20(11):548–557. https://doi.org/10.1016/j.tim.2012.07.005 PMID: 22898147
2. Zhang XSW. Synthetic Biology of Cyanobacteria. vol. 1080. Springer, Singapore; 2018.

3. Herrero A, Flores E. Genetic Responses to Carbon and Nitrogen Availability in *Anabaena*. Environ Microbiol. 2019; 21(1):1–17. https://doi.org/10.1111/1462-2920.14370 PMID: 30066380
4. Flores E, Herrero A. Compartmentalized Function through Cell Differentiation in Filamentous Cyanobacteria. Nat Rev Microbiol. 2010; 8(1):39–50. https://doi.org/10.1038/nrmicro2242 PMID: 19966815
5. Herrero A, Picossi S, Flores E. Gene Expression during Heterocyst Differentiation. In: Chauvat F, Cassier-Chauvat C, editors. Genomics of Cyanobacteria. vol. 65 of Advances in Botanical Research. Academic Press; 2013. p. 281–329.
6. Herrero A, Muro-Pastor AM, Valladares A, Flores E. Cellular Differentiation and the NtcA Transcription Factor in Filamentous Cyanobacteria. FEMS Microbiol Rev. 2004; 28(4):469–487. https://doi.org/10.1016/j.femsre.2004.04.003 PMID: 15374662
7. Zhang CC, Laurent S, Sakr S, Peng L, Bédu S. Heterocyst Differentiation and Pattern Formation in Cyanobacteria: A Chorus of Signals. Mol Microbiol. 2006; 59(2):367–375. https://doi.org/10.1111/j.1365-2958.2005.04979.x PMID: 16390435
8. Valladares A, Flores E, Herrero A. Transcription Activation by NtcA and 2-Oxoglutarate of Three Genes Involved in Heterocyst Differentiation in the Cyanobacterium *Anabaena* Sp. Strain PCC 7120. J Bacteriol. 2008; 190(18):6126–6133. https://doi.org/10.1128/JB.00787-08 PMID: 18658268
9. Buikema WJ, Haselkorn R. Characterization of a Gene Controlling Heterocyst Differentiation in the Cyanobacterium *Anabaena 7120*. Genes Dev. 1991; 5(2):321–330. https://doi.org/10.1101/gad.5.2.321 PMID: 1840555
10. Black TA, Cai Y, Wolk CP. Spatial Expression and Autoregulation of *hetR*, a Gene Involved in the Control of Heterocyst Development in *Anabaena*. Mol Microbiol. 1993; 9(1):77–84. https://doi.org/10.1111/j.1365-2958.1993.tb01670.x PMID: 8412673
11. Rajagopalan R, Callahan SM. Temporal and Spatial Regulation of the Four Transcription Start Sites of *hetR* from *Anabaena* Sp. Strain PCC 7120. J Bacteriol. 2010; 192(4):1088–1096. https://doi.org/10.1128/JB.01297-09 PMID: 20008074
12. Orozco CC, Risser DD, Callahan SM. Epistasis Analysis of Four Genes from *Anabaena* Sp. Strain PCC 7120 Suggests a Connection between *patA* and *patS* in Heterocyst Pattern Formation. J Bacteriol. 2006; 188(5):1808–1816. https://doi.org/10.1128/JB.188.5.1808-1816.2006 PMID: 16484191
13. Yoon HS, Golden JW. Heterocyst Pattern Formation Controlled by a Diffusible Peptide. Science. 1998; 282(5390):935–938. https://doi.org/10.1126/science.282.5390.935 PMID: 9794762
14. Yoon HS, Golden JW. *patS* and Products of Nitrogen Fixation Control Heterocyst Pattern. J Bacteriol. 2001; 183(8):2605–2613. https://doi.org/10.1128/JB.183.8.2605-2613.2001 PMID: 11274121
15. Feldmann EA, Ni S, Sahu ID, Mishler CH, Levengood JD, Kushnir Y, et al. Differential Binding between *patS* C-terminal Peptide Fragments and *hetR* from *Anabaena* Sp. PCC 7120. Biochemistry. 2012; 51(12):2436–2442. https://doi.org/10.1021/bi300228n PMID: 22397695
16. Huang X, Dong Y, Zhao J. *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. 2004; 101(14):4848–4853. https://doi.org/10.1073/pnas.0400429101 PMID: 15051891
17. Risser DD, Callahan SM. Mutagenesis of *hetR* Reveals Amino Acids Necessary for *hetR* Function in the Heterocystous Cyanobacterium *Anabaena* Sp. Strain PCC 7120. J Bacteriol. 2007; 189(6):2460–2467. https://doi.org/10.1128/JB.01241-06 PMID: 17220221
18. Corrales-Guerrero L, Mariscal V, Flores E, Herrero A. Functional Dissection and Evidence for Intercellular Transfer of the Heterocyst-Differentiation *patS* Morphogen. Mol Microbiol. 2013; 88(6):1093–1105. https://doi.org/10.1111/mmi.12244 PMID: 23663167
19. Elhai J, Khudyakov I. Ancient Association of Cyanobacterial Multicellularity with the Regulator *hetR* and an RGSGR Pentapeptide-Containing Protein (PatX). Mol Microbiol. 2018; 110(6):931–954. https://doi.org/10.1111/mmi.14003 PMID: 29885033
20. Khudyakov I, Gladkov G, Elhai J. Inactivation of Three Rg(S/t)Gr Pentapeptide-Containing Negative Regulators of *hetR* Results in Lethal Differentiation of *Anabaena* PCC 7120. Life. 2020; 10(12):1–16. https://doi.org/10.3390/life10120326
21. Higa KC, Rajagopalan R, Risser DD, Rivers OS, Tom SK, Videau P, et al. The RGSGR Amino Acid Motif of the Intercellular Signalling Protein, *hetN*, Is Required for Patterning of Heterocysts in *Anabaena* Sp. Strain PCC 7120. Mol Microbiol. 2012; 83(4):682–693. https://doi.org/10.1111/j.1365-2958.2011.07949.x PMID: 22220907
22. Rivers OS, Beurmann S, Dow A, Cozy LM, Videau P. Phenotypic Assessment Suggests Multiple Start Codons for *hetN*, an Inhibitor of Heterocyst Differentiation, in *Anabaena* Sp. Strain PCC 7120. J Bacteriol. 2018; 200(16):1–14. https://doi.org/10.1128/JB.00220-18

23. Callahan SM, Buikema WJ. The Role of *hetN* in Maintenance of the Heterocyst Pattern in *Anabaena* Sp. PCC 7120. Mol Microbiol. 2001; 40(4):941–950. https://doi.org/10.1046/j.1365-2958.2001.02437.x PMID: 11401701

24. Borthakur PB, Orozco CC, Young-Robbins SS, Haselkorn R, Callahan SM. Inactivation of *patS* and *hetN* Causes Lethal Levels of Heterocyst Differentiation in the Filamentous Cyanobacterium *Anabaena* Sp. PCC 7120. Mol Microbiol. 2005; 57(1):111–123. https://doi.org/10.1111/j.1365-2958.2005.04678.x PMID: 15948953

25. Di Patti F, Lavacchi L, Arbel-Goren R, Schein-Lubomirsky L, Fanelli D, Stavans J. Robust Stochastic Turing Patterns in the Development of a One-Dimensional Cyanobacterial Organism. PLoS Biol. 2018; 16(5):e2004877. https://doi.org/10.1371/journal.pbio.2004877 PMID: 29727442

26. Risser DD, Callahan SM. *hetF* and *patA* Control Levels of *hetR* in *Anabaena* Sp. Strain PCC 7120. J Bacteriol. 2008; 190(23):7645–7654. https://doi.org/10.1128/JB.01110-08 PMID: 18835986

27. Brenes-Álvarez M, Minguet M, Vioque A, Muro-Pastor AM. NsiR1, a Small RNA with Multiple Copies, Modulates Heterocyst Differentiation in the Cyanobacterium *Nostoc* Sp. PCC 7120. Environ Microbiol. 2020; 22(8):3325–3338. https://doi.org/10.1111/1462-2920.15103 PMID: 32468657

28. Risser DD, Callahan SM. Genetic and Cytological Evidence That Heterocyst Patterning Is Regulated by Inhibitor Gradients That Promote Activator Decay. Proc Natl Acad Sci U S A. 2009; 106(47):19884–19888. https://doi.org/10.1073/pnas.0909152106 PMID: 19897721

29. Liang J, Scappino L, Haselkorn R. The *patA* Gene Product, Which Contains a Region Similar to CheY of *Escherichia coli*, Controls Heterocyst Pattern Formation in the Cyanobacterium *Anabaena 7120*. Proc Natl Acad Sci U S A. 1992; 89(12):5655–5659. https://doi.org/10.1073/pnas.89.12.5655 PMID: 1608976

30. Young-Robbins SS, Risser DD, Moran JR, Haselkorn R, Callahan SM. Transcriptional Regulation of the Heterocyst Patterning Gene *patA* from *Anabaena* Sp. Strain PCC 7120. J Bacteriol. 2010; 192(18):4732–4740. https://doi.org/10.1128/JB.00577-10 PMID: 20622060

31. Wilcox M, Mitchison GJ, Smith RJ. Pattern Formation in the Blue-Green Alga, *Anabaena*. I. Basic Mechanisms. J Cell Sci. 1973; 12(3):707–723. https://doi.org/10.1242/jcs.12.3.707 PMID: 4198321

32. Wolk CP. Formation of One-Dimensional Patterns by Stochastic Processes and by Filamentous Blue-Green Algae. Dev Biol. 1975; 46(2):370–382. https://doi.org/10.1016/0012-1606(75)90113-X PMID: 810378

33. Meinhardt H. Models of Biological Pattern Formation: From Elementary Steps to the Organization of Embryonic Axes. Curr Top Dev Biol. 2008; 81:1–63. https://doi.org/10.1016/S0070-2153(07)81001-5 PMID: 18023723

34. Pinzon NM, Ju LK. Modeling Culture Profiles of the Heterocystous N2-fixing Cyanobacterium *Anabaena flos-aquae*. Biotechnol Prog. 2006; 22(6):1532–1540. https://doi.org/10.1021/bp060163c PMID: 17137298

35. Allard JF, Hill AL, Rutenberg AD. Heterocyst Patterns without Patterning Proteins in Cyanobacterial Filaments. Dev Biol. 2007; 312(1):427–434. https://doi.org/10.1016/j.ydbio.2007.09.045 PMID: 17976569

36. Gerdtzen ZP, Salgado CJ, Osses A, Asenjo JA, Rapaport I, Andrews BA. Modeling Heterocyst Pattern Formation in Cyanobacteria. BMC Bioinformatics. 2009; 10(SUPPL. 6):S16. https://doi.org/10.1186/1471-2105-10-S6-S16 PMID: 19534741

37. Zhu M, Callahan SM, Allen JS. Maintenance of Heterocyst Patterning in a Filamentous Cyanobacterium. J Biol Dyn. 2010; 4(6):621–633. https://doi.org/10.1080/17513751003777507 PMID: 22881208

38. Brown AI, Rutenberg AD. Reconciling Cyanobacterial Fixed-Nitrogen Distributions and Transport Experiments with Quantitative Modelling. Phys Biol. 2012; 9(1):16007. https://doi.org/10.1088/1478-3975/9/1/016007 PMID: 22313598

39. Brown AI, Rutenberg AD. A Storage-Based Model of Heterocyst Commitment and Patterning in Cyanobacteria. Phys Biol. 2014; 11(1):16001. https://doi.org/10.1088/1478-3975/11/1/016001 PMID: 24384886

40. Torres-Sánchez A, Gómez-Gardeñes J, Falo F. An Integrative Approach for Modeling and Simulation of Heterocyst Pattern Formation in Cyanobacteria Filaments. PLoS Comput Biol. 2015; 11(3):e1004129. https://doi.org/10.1371/journal.pcbi.1004129 PMID: 25816286

41. Ishihara Ji, Tachikawa M, Iwasaki H, Mochizuki A. Mathematical Study of Pattern Formation Accompanied by Heterocyst Differentiation in Multicellular Cyanobacterium. J Theor Biol. 2015; 371:9–23. https://doi.org/10.1016/j.jtbi.2015.01.034 PMID: 25665721

42. Muñoz-García J, Ares S. Formation and Maintenance of Nitrogen-Fixing Cell Patterns in Filamentous Cyanobacteria. Proc Natl Acad Sci U S A. 2016; 113(22):6218–6223. https://doi.org/10.1073/pnas.1524383113 PMID: 27162328

43. Malatinszky D, Steuer R, Jones PR. A Comprehensively Curated Genome-Scale Two-Cell Model for the Heterocystous Cyanobacterium *Anabaena* Sp. PCC 7120. Plant Physiol. 2017; 173(1):509–523. https://doi.org/10.1104/pp.16.01487 PMID: 27899536

44. Grover JP, Scott JT, Roelke DL, Brooks BW. Dynamics of Nitrogen-Fixing Cyanobacteria with Heterocysts: A Stoichiometric Model. Mar Freshw Res. 2019; 71(5):644–658. https://doi.org/10.1071/MF18361

45. Gardiner CW. Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences. vol. 13 of Springer Series in Synergetics. 3rd ed. Berlin: Springer-Verlag; 2004.

46. Frigola D, Casanellas L, Sancho JM, Ibañes M. Asymmetric Stochastic Switching Driven by Intrinsic Molecular Noise. PLoS ONE. 2012; 7(2):e31407. https://doi.org/10.1371/journal.pone.0031407 PMID: 22363638

47. Kirkpatrick S, Gelatt CD, Vecchi MP. Optimization by Simulated Annealing. Science. 1983; 220(4598):671–680. https://doi.org/10.1126/science.220.4598.671 PMID: 17813860

48. Asai H, Iwamori S, Kawai K, Ehira S, Ishihara JI, Aihara K, et al. Cyanobacterial Cell Lineage Analysis of the Spatiotemporal *hetR* Expression Profile during Heterocyst Pattern Formation in *Anabaena* Sp. PCC 7120. PLoS ONE. 2009; 4(10):e7371. https://doi.org/10.1371/journal.pone.0007371 PMID: 19823574

49. Massey FJ Jr. The Kolmogorov-Smirnov Test for Goodness of Fit. J Am Stat Assoc. 1951; 46(253):68–78. https://doi.org/10.1080/01621459.1951.10500769

50. Corrales-Guerrero L, Flores E, Herrero A. Relationships between the ABC-exporter *hetC* and Peptides That Regulate the Spatiotemporal Pattern of Heterocyst Distribution in *Anabaena*. PLoS ONE. 2014; 9(8):e104571. https://doi.org/10.1371/journal.pone.0104571 PMID: 25121608

51. Burnat M, Schleiff E, Flores E. Cell Envelope Components Influencing Filament Length in the Heterocyst-Forming Cyanobacterium *Anabaena* Sp. Strain PCC 7120. J Bacteriol. 2014; 196(23):4026–4035. https://doi.org/10.1128/JB.02128-14 PMID: 25201945

52. Olmedo-Verd E, Flores E, Herrero A, Muro-Pastor AM. *hetR*-Dependent and Independent Expression of Heterocyst-Related Genes in an *Anabaena* Strain Overproducing the NtcA Transcription Factor. J Bacteriol. 2005; 187(6):1985–1991. https://doi.org/10.1128/JB.187.6.1985-1991.2005 PMID: 15743946

53. Ehira S, Ohmori M. NrrA Directly Regulates Expression of *hetR* during Heterocyst Differentiation in the Cyanobacterium *Anabaena* Sp. Strain PCC 7120. J Bacteriol. 2006; 188(24):8520–8525. https://doi.org/10.1128/JB.01314-06 PMID: 17041048

54. Black TA, Wolk CP. Analysis of a Het- Mutation in *Anabaena* Sp. Strain PCC 7120 Implicates a Secondary Metabolite in the Regulation of Heterocyst Spacing. J Bacteriol. 1994; 176(8):2282–2292. https://doi.org/10.1128/jb.176.8.2282-2292.1994 PMID: 8157596

55. Bauer CC, Ramaswamy KS, Endley S, Scappino LA, Golden JW, Haselkorn R. Suppression of Heterocyst Differentiation in *Anabaena* PCC 7120 by a Cosmid Carrying Wild-Type Genes Encoding Enzymes for Fatty Acid Synthesis. FEMS Microbiol Lett. 1997; 151(1):23–30. https://doi.org/10.1111/j.1574-6968.1997.tb10390.x PMID: 9198279

56. Li B, Huang X, Zhao J. Expression of *hetN* during Heterocyst Differentiation and Its Inhibition of *hetR* Up-Regulation in the Cyanobacterium *Anabaena* Sp. PCC 7120. FEBS Lett. 2002; 517(1-3):87–91. https://doi.org/10.1016/S0014-5793(02)02582-6 PMID: 12062415

57. Corrales-Guerrero L, Mariscal V, Nürnberg DJ, Elhai J, Mullineaux CW, Flores E, et al. Subcellular Localization and Clues for the Function of the *hetN* Factor Influencing Heterocyst Distribution in *Anabaena* Sp. Strain PCC 7120. J Bacteriol. 2014; 196(19):3452–3460. https://doi.org/10.1128/JB.01922-14 PMID: 25049089

58. El-Shehawy RM, Kleiner D. Effect of Controlled Expression of the *hetR* Gene on Heterocyst Formation in the Filamentous Cyanobacterium *Anabaena* Sp. PCC 7120. Physiol Plant. 2003; 119(1):44–48. https://doi.org/10.1034/j.1399-3054.2003.00108.x

59. Kim Y, Joachimiak G, Ye Z, Binkowski TA, Zhang R, Gornicki P, et al. Structure of Transcription Factor *hetR* Required for Heterocyst Differentiation in Cyanobacteria. Proc Natl Acad Sci U S A. 2011; 108(25):10109–10114. https://doi.org/10.1073/pnas.1106840108 PMID: 21628585

60. Zhou R, Wei X, Jiang N, Li H, Dong Y, Hsi KL, et al. Evidence That *hetR* Protein Is an Unusual Serine-Type Protease. Proc Natl Acad Sci U S A. 1998; 95(9):4959–4963. https://doi.org/10.1073/pnas.95.9.4959 PMID: 9560210

61. Buikema WJ, Haselkorn R. Expression of the *Anabaena hetR* Gene from a Copper-Regulated Promoter Leads to Heterocyst Differentiation under Repressing Conditions. Proc Natl Acad Sci U S A. 2001; 98(5):2729–2734. https://doi.org/10.1073/pnas.051624898 PMID: 11226308

62. Rivers OS, Videau P, Callahan SM. Mutation of *sepJ* Reduces the Intercellular Signal Range of a *hetN*-Dependent Paracrine Signal, but Not of a *patS*-Dependent Signal, in the Filamentous Cyanobacterium *Anabaena* Sp. Strain PCC 7120. Mol Microbiol. 2014; 94(6):1260–1271. https://doi.org/10.1111/mmi.12836 PMID: 25336355

63. Hou S, Zhou F, Peng S, Gao H, Xu X. The *hetR*-Binding Site That Activates Expression of *patA* in Vegetative Cells Is Required for Normal Heterocyst Patterning in *Anabaena* Sp. PCC 7120. Sci Bull. 2015; 60(2):192–201. https://doi.org/10.1007/s11434-014-0724-5

64. Wang Y, Xu X. Regulation by *hetC* of Genes Required for Heterocyst Differentiation and Cell Division in *Anabaena* Sp. Strain PCC 7120. J Bacteriol. 2005; 187(24):8489–8493. https://doi.org/10.1128/JB.187.24.8489-8493.2005 PMID: 16321953

65. Videau P, Rivers OS, Higa KC, Callahan SM. ABC Transporter Required for Intercellular Transfer of Developmental Signals in a Heterocystous Cyanobacterium. J Bacteriol. 2015; 197(16):2685–2693. https://doi.org/10.1128/JB.00304-15 PMID: 26055115

66. Zhou R, Cao Z, Zhao J. Characterization of *hetR* Protein Turnover in *Anabaena* Sp. PCC 7120. Arch Microbiol. 1998; 169(5):417–426. https://doi.org/10.1007/s002030050592 PMID: 9560423

67. Valladares A, Flores E, Herrero A. The Heterocyst Differentiation Transcriptional Regulator *hetR* of the Filamentous Cyanobacterium *Anabaena* Forms Tetramers and Can Be Regulated by Phosphorylation. Mol Microbiol. 2016; 99(4):808–819. https://doi.org/10.1111/mmi.13268 PMID: 26552991

68. Makarova KS, Koonin EV, Haselkorn R, Galperin MY. Cyanobacterial Response Regulator *patA* Contains a Conserved N-terminal Domain (PATAN) with an Alpha-Helical Insertion. Bioinformatics. 2006; 22(11):1297–1301. https://doi.org/10.1093/bioinformatics/btl096 PMID: 16543275

69. Roumezi B, Xu X, Risoul V, Fan Y, Lebrun R, Latifi A. The Pkn22 Kinase of *Nostoc* PCC 7120 Is Required for Cell Differentiation via the Phosphorylation of *hetR* on a Residue Highly Conserved in Genomes of Heterocyst-Forming Cyanobacteria. Front Microbiol. 2020; 10(January):1–18. https://doi.org/10.3389/fmicb.2019.03140 PMID: 32038573

70. Kim Y, Ye Z, Joachimiak G, Videau P, Young J, Hurd K, et al. Structures of Complexes Comprised of *Fischerella* Transcription Factor *hetR* with *Anabaena* DNA Targets. Proc Natl Acad Sci U S A. 2013; 110(19):E1716–E1723. https://doi.org/10.1073/pnas.1305971110 PMID: 23610410

71. Hu HX, Jiang YL, Zhao MX, Cai K, Liu S, Wen B, et al. Structural Insights into *hetR*'*patS* Interaction Involved in Cyanobacterial Pattern Formation. Sci Rep. 2015; 5:16470. https://doi.org/10.1038/srep16470 PMID: 26576507

72. Wong FCY, Meeks JC. The *hetF* Gene Product Is Essential to Heterocyst Differentiation and Affects *hetR* Function in the Cyanobacterium *Nostoc punctiforme*. J Bacteriol. 2001; 183(8):2654–2661. https://doi.org/10.1128/JB.183.8.2654-2661.2001 PMID: 11274126

73. Wu X, Liu D, Lee MH, Golden JW. *patS* Minigenes Inhibit Heterocyst Development of *Anabaena* Sp. Strain PCC 7120. J Bacteriol. 2004; 186(19):6422–6429. https://doi.org/10.1128/JB.186.19.6422-6429.2004 PMID: 15375122

74. Khudyakov IY, Golden JW. Different Functions of *hetR*, a Master Regulator of Heterocyst Differentiation in *Anabaena* Sp. PCC 7120, Can Be Separated by Mutation. Proc Natl Acad Sci U S A. 2004; 101(45):16040–16045. https://doi.org/10.1073/pnas.0405572101 PMID: 15520378

75. Risser DD, Wong FCY, Meeks JC. Biased Inheritance of the Protein PatN Frees Vegetative Cells to Initiate Patterned Heterocyst Differentiation. Proc Natl Acad Sci U S A. 2012; 109(38):15342–15347. https://doi.org/10.1073/pnas.1207530109 PMID: 22949631

76. Higa KC, Callahan SM. Ectopic Expression of *hetP* Can Partially Bypass the Need for *hetR* in Heterocyst Differentiation by *Anabaena* Sp. Strain PCC 7120. Mol Microbiol. 2010; 77(3):562–574. https://doi.org/10.1111/j.1365-2958.2010.07257.x PMID: 20545862

77. Du Y, Cai Y, Hou S, Xu X. Identification of the *hetR* Recognition Sequence Upstream of *hetZ* in *Anabaena* Sp. Strain PCC 7120. J Bacteriol. 2012; 194(9):2297–2306. https://doi.org/10.1128/JB.00119-12 PMID: 22389489

78. Videau P, Rivers OS, Hurd K, Ushijima B, Oshiro RT, Ende RJ, et al. The Heterocyst Regulatory Protein HetP and Its Homologs Modulate Heterocyst Commitment in *Anabaena* Sp. Strain PCC 7120. Proc Natl Acad Sci U S A. 2016; 113(45):E6984–E6992. https://doi.org/10.1073/pnas.1610533113 PMID: 27791130

79. Videau P, Rivers OS, Tom SK, Oshiro RT, Ushijima B, Swenson VA, et al. The *hetZ* Gene Indirectly Regulates Heterocyst Development at the Level of Pattern Formation in *Anabaena* Sp. Strain PCC 7120. Mol Microbiol. 2018; 109(1):91–104. https://doi.org/10.1111/mmi.13974

80. Zhang H, Wang S, Wang Y, Xu X. Functional Overlap of *hetP* and *hetZ* in Regulation of Heterocyst Differentiation in *Anabaena* Sp. Strain PCC 7120. J Bacteriol. 2018; 200(9):e00707–17. https://doi.org/10.1128/JB.00707-17 PMID: 29440250

81. Du Y, Zhang H, Wang H, Wang S, Lei Q, Li C, et al. Expression from DIF1-motif Promoters of *hetR* and *patS* Is Dependent on HetZ and Modulated by PatU3 during Heterocyst Differentiation. PLoS ONE. 2020; 15(7):e0232383. https://doi.org/10.1371/journal.pone.0232383 PMID: 32701963

82. Corrales-Guerrero L, Tal A, Arbel-Goren R, Mariscal V, Flores E, Herrero A, et al. Spatial Fluctuations in Expression of the Heterocyst Differentiation Regulatory Gene *hetR* in *Anabaena* Filaments. PLoS Genet. 2015; 11(4):e1005031. https://doi.org/10.1371/journal.pgen.1005031 PMID: 25830300

83. Xu X, Risoul V, Byrne D, Champ S, Douzi B, Latifi A. *hetL*, *hetR* and *patS* Form a Reaction-Diffusion System to Control Pattern Formation in the Cyanobacterium Nostoc PCC 7120. eLife. 2020; 9:1–48. https://doi.org/10.7554/eLife.59190

84. Liu D, Golden JW. *hetL* Overexpression Stimulates Heterocyst Formation in *Anabaena* Sp. Strain PCC 7120. J Bacteriol. 2002; 184(24):6873–6881. https://doi.org/10.1128/JB.184.24.6873-6881.2002 PMID: 12446638