# Interaction network among factors involved in heterocyst-patterning in cyanobacteria

Xiaomei Xu, Raphaël Rachedi, Maryline Foglino, Emmanuel Talla, Amel Latifi

Molecular Genetics and Genomics

# Interaction network among factors involved in heterocyst-patterning in cyanobacteria

--Manuscript Draft--

| | |
|---|---|
| **Manuscript Number:** | MGAG-D-21-00741R2 |
| **Full Title:** | Interaction network among factors involved in heterocyst-patterning in cyanobacteria |
| **Article Type:** | Original Article |
| **Corresponding Author:** | Amel Latifi, Ph.D<br>Aix-Marseille Universite<br>Marseille, FRANCE |
| **Corresponding Author Secondary Information:** | |
| **Corresponding Author's Institution:** | Aix-Marseille Universite |
| **Corresponding Author's Secondary Institution:** | |
| **First Author:** | Xiaomei Xu |
| **First Author Secondary Information:** | |
| **Order of Authors:** | Xiaomei Xu |
| | Raphaël Rachedi |
| | Maryline Foglino |
| | Emmanuel Talla |
| | Amel Latifi, Ph.D |
| **Order of Authors Secondary Information:** | |
| **Funding Information:** | Agence Nationale de la Recherche (ANR-21-CE20-0025-01) — Pr. Amel Latifi |
| **Abstract:** | The genetically regulated pattern of heterocyst formation in multicellular cyanobacteria represents the simplest model to address how patterns emerge and are established, the signals that control them, and the regulatory pathways that act downstream. Although numerous factors involved in this process have been identified, the mechanisms of action of many of them remain largely unknown. The aim of this study was to identify specific relationships between 14 factors required for cell differentiation and pattern formation by exploring their putative physical interactions in the cyanobacterium model Nostoc sp. PCC 7120 and by probing their evolutionary conservation and distribution across the cyanobacterial phylum. A bacterial two-hybrid assay indicated that 10 of the 14 factors studied here are engaged in more than one protein-protein interaction. The transcriptional regulator PatB was central in this network as it showed the highest number of binary interactions. A phylum-wide genomic survey of the distribution of these factors in cyanobacteria showed that they are all highly conserved in the genomes of heterocyst-forming strains, with the PatN protein being almost restricted to this clade. Interestingly, eight of the factors that were shown to be capable of protein interactions were identified as key elements in the evolutionary genomics analysis. These data suggest that a network of 12 proteins may play a crucial role in heterocyst development and patterning. Unravelling the physical and functional interactions between these factors during heterocyst development will certainly shed light on the mechanisms underlying pattern establishment in cyanobacteria. |
| **Response to Reviewers:** | Dear editor,<br>in this revised version of our manuscript we introduced all the corrections asked by the editor and Reviewer 1. The modifications in the text are marked in yellow,<br>We thank the reviewers and the communicating editor for their valuable reading of our publication. We hope that this revised manuscript will be suited for publication. |

# Interaction network among factors involved in heterocyst-patterning in cyanobacteria

Xiaomei Xu[1*], Raphaël Rachedi[1*], Maryline Foglino[1], Emmanuel Talla[1**], Amel Latifi[1**]

[1]: Aix Marseille Univ, CNRS, IMM, LCB, Laboratoire de Chimie Bactérienne, Marseille, France

*: these authors contributed equally to the work

**: correspondence:

Pr Amel Latifi

latifi@imm.cnrs.fr

Dr Emmanuel Talla

talla@imm.cnrs.fr

**Key words:** Cell differentiation, Cyanobacteria, Genomics, Pattern establishment, Protein-Protein interaction.

**Running title:** Protein-protein interactions and heterocyst pattern

## Abstract

The genetically regulated pattern of heterocyst formation in multicellular cyanobacteria represents the simplest model to address how patterns emerge and are established, the signals that control them, and the regulatory pathways that act downstream. Although numerous factors involved in this process have been identified, the mechanisms of action of many of them remain largely unknown. The aim of this study was to identify specific relationships between 14 factors required for cell differentiation and pattern formation by exploring their putative physical interactions in the cyanobacterium model *Nostoc* sp. PCC 7120 and by probing their evolutionary conservation and distribution across the cyanobacterial phylum. A bacterial two-hybrid assay indicated that 10 of the 14 factors studied here are engaged in more than one protein-protein interaction. The transcriptional regulator PatB was central in this network as it showed the highest number of binary interactions. A phylum-wide genomic survey of the distribution of these factors in cyanobacteria showed that they are all highly conserved in the genomes of heterocyst-forming strains, with the PatN protein being almost restricted to this clade. Interestingly, eight of the factors that were shown to be capable of protein interactions were identified as key elements in the evolutionary genomics analysis. These data suggest that a network of 12 proteins may play a crucial role in heterocyst development and patterning. Unravelling the physical and functional interactions between these factors during heterocyst development will certainly shed light on the mechanisms underlying pattern establishment in cyanobacteria.

# Introduction

Spatial patterns of cellular differentiation occur in many developmental organisms across the tree of life, and in most cases, key aspects of the molecular interactions that are involved in patterning are poorly understood. In this work, we questioned the possible involvement of protein-protein interactions in patterning. Our study model is the development of heterocysts in multicellular and diazotrophic cyanobacteria. Indeed, in prokaryotes, heterocyst formation in multicellular cyanobacteria represents the simplest model to address one-dimensional patterns based on repeated peaks of diffusible morphogens. These peaks are generated by the activity of a transcriptional regulator that activates the expression of the gene encoding its own inhibitor (the diffusible morphogen); thus, this loop has been called the reaction-diffusion model (Fig. 1a) (Turing 1952).

Heterocysts are nearly anoxic cells that form semi-regularly along the filaments and function to host the oxygen-sensitive nitrogenase. Therefore, the developmental program in these heterocyst-forming bacteria allows the occurrence of two antagonistic metabolic activities (i.e., oxygenic photosynthesis and nitrogen fixation) in the same organism (Meeks and Elhai 2002; Kumar et al. 2010; Herrero et al. 2016). Much of the knowledge about cell differentiation in cyanobacteria stems from studies of the model bacterium *Anabaena/Nostoc* sp. PCC 7120 (termed *Nostoc* hereafter). Combined nitrogen starvation triggers the activation of the expression of the *hetR* gene, which encodes the master and specific transcriptional activator of cell differentiation (Buikema and Haselkorn 1991). HetR is essential for cell differentiation, as its absence abolishes differentiation and its overexpression induces the differentiation of multiple contiguous heterocysts (MCH) under combined nitrogen starvation and allows differentiation even under non-permissive conditions (Buikema and Haselkorn 1991). The regulation of *hetR* transcription is complex and involves the global regulator NtcA and a positive autoregulation loop mediated by HetR itself (Black et al. 1993; Frias et al. 1994; Rajagopalan and Callahan 2010). The inhibition of HetR activity in the neighboring cells of the heterocysts initiates and maintains the pattern of differentiation throughout the growth of the filaments. Three genes (*patS*, *patX*, and *hetN*), the expression of which is induced in developing cells, negatively control this process (Black and Wolk 1994; Yoon and Golden 1998; Yoon and Golden 2001; Higa et al. 2012; Corrales-Guerrero

et al. 2014b; Elhai and Khudyakov 2018; Khudyakov et al. 2020). Deletion of the *patS* gene also leads to the formation of multiple contiguous heterocysts (Fig. 1b). A common feature of the PatS, PatX, and HetN proteins is that they contain a conserved hexapeptide (ERGSGR for PatS and HetN, and HRGTGR for PatX). The RGSGR peptide inactivates the DNA-binding activity of HetR (Huang et al. 2004; Feldmann et al. 2011). These peptides are considered as diffusible morphogens that ensure patterning by lateral inhibition (Corrales-Guerrero et al. 2013). Recently, HetR was shown to interact with a pentapeptide-repeat-containing protein called HetL at the same interface used by the ERGSGR peptide (Xu et al. 2020). Moreover, the transcription of *hetL*, which is a heterocyst-specific gene, is activated by HetR (Xu et al. 2020). *hetL* overexpression suppresses the inhibitory effect of PatS (Liu and Golden 2002) and the interaction of HetL with HetR provides protection against PatS (Xu et al. 2020), thus explaining how the developing cell becomes immune to PatS/HetN/PatX inhibition in the heterocyst (Fig. 1a).

Correct patterning requires the function of several genes that have been termed *pat* genes (Fig. 1b). The phenotype of mutants of these genes provides an important indication of their role and establishes a starting point for the investigation of how the pattern is established. These genes include *patA* and *patL*, the mutation of which results in the formation of rare heterocysts only at the ends of the filaments (Liang et al. 1992; Liu and Wolk 2011). PatA is a two-domain protein with a C-terminal region similar to the CheY-receiver domain and a conserved N-terminal domain called PATAN (Makarova et al. 2006). PatL is a pentapeptide-repeat-containing protein with two transmembrane domains (Liu and Wolk 2011). In addition to sharing similar mutant phenotypes, the PatA and PatL proteins interact, as shown by yeast-two-hybrid assays (Liu and Wolk 2011). A cluster of three genes (*hetZ*, *patU3*, and *patU5*) that are expressed in the heterocysts and encode proteins of unknown function modulates heterocyst formation and pattern establishment (Zhang et al. 2007). Inactivation of *patU3* leads to an MCH phenotype and restores heterocyst formation in a *patA* mutant background (Zhang et al. 2007). A *hetZ* mutant of *Nostoc* is unable to differentiate heterocysts (Zhang et al. 2007; Videau et al. 2018). HetZ action has been proposed to be involved in its interaction with PatU3 and, to a lesser extent, with HetR (Videau et al. 2018; Zhang et al. 2018; Du et al. 2020). In *N. punctiforme*, the deletion of the *patN* gene, which encodes

a putative membrane protein, increases heterocyst frequency and decreases the number of vegetative cells between two heterocysts, thus leading to a multiple singular heterocyst (MSH) phenotype (Risser et al. 2012) (Fig. 1b). A similar phenotype has been observed in a *patN* mutant of *Nostoc* (Abed et al. 2009). PatN was proposed to modulate the competency of a vegetative cell to differentiate into a heterocyst (Risser et al. 2012); therefore, it also affects patterning. Interestingly, the MSH phenotype has been observed in a mutant of the *patB* (*cnfR*) gene, the product of which is predicted to be a transcriptional regulator with a DNA-binding helix-turn-helix at its C-terminal extremity and two putative ferredoxin domains at its N terminus (Jones et al. 2003). In *Anabaena variabilis*, the homolog of PatB (CnfR1) activates the transcription of the heterocyst specific nitrogenase-encoding genes (Pratte and Thiel 2016). Furthermore, in the non-heterocyst-forming cyanobacterium *Leptolyngbya boryana*, CnfR is essential for the transcription of the *nif* genes and for the diazotrophic growth under micro-oxic conditions (Tsujimoto et al. 2014). In *Nostoc*, the transcription of the *nif* genes has been shown to be abolished in a *patB* (*cnfR*) mutant (Kurio et al. 2020) and the DNA-binding sites of this regulator has been predicted in the promoter of the *nif* operon (Brenes-Alvarez et al. 2019). The function of CnfR/PatB as activator of the *nif* genes seems therefore conserved in cyanobacteria, but the manner in which PatB exerts this negative effect on heterocyst spacing in *Nostoc* is unknown. Two other *pat* genes, *patC* and *patD*, both encoding proteins of unknown functions, seem to negatively regulate patterning, as their deletion increases heterocyst frequency (Corrales-Guerrero et al. 2014a; Wang et al. 2019).

In addition to the Pat proteins, other factors involved in the early steps of the differentiation process can affect pattern formation. The *hetC* gene, which encodes a putative bacterial ABC protein transporter, is involved in the regulation of the early steps of heterocyst differentiation (Khudyakov and Wolk 1997; Corrales-Guerrero et al. 2014a; Videau et al. 2015). *hetC* inactivation mutants exhibit a pattern of spaced cells that are still able to divide (Khudyakov and Wolk 1997; Xu and Wolk 2001). These cells are likely to represent an early stage of heterocyst differentiation; thus, HetC has been suggested to be required for the transition of heterocysts to their terminal non-dividing state (Xu and Wolk 2001). *hetF*, which encodes a putative protease, is another gene that is essential for heterocyst formation, as a *hetF* mutant

does not form heterocysts and has altered morphology, whereas *hetF* overexpression results in an MCH phenotype (Wong and Meeks 2001; Risser and Callahan 2008).

Shortly after the initiation step and pattern establishment, the decision to differentiate is irreversibly fixed, which is known as commitment. The HetP protein (encoded by the *alr2818* gene) and its homologs (encoded by the *asl1930*, *alr3234*, and *alr2902* genes) have been proposed as important players in the regulation of this step (Videau et al. 2016). The inactivation of the *hetP* gene affects the differentiation process downstream of initiation and patterning, as the mutant develops heterocysts, albeit at a lower frequency and at a later time than the wild type strain (Fernandez-Pinas et al. 1994; Higa and Callahan 2010). Unlike *hetP*, inactivation of the genes encoding the 3 HetP homologs does not impair cell differentiation, indicating that they are not required for heterocyst formation or patterning (40). Overexpression of *hetP* bypasses the absence of HetZ and vice versa, which indicates a functional overlap between these factors (Zhang et al. 2018). This observation highlights the connection between patterning and commitment and places HetZ at the intersection between the two steps.

In summary, pattern establishment has clearly been shown to be genetically controlled, and the involvement of several factors in this process has been demonstrated (Fig. 1b). However, their function and their potential molecular interactions remain poorly known. Because protein–protein interaction approaches have been shown to be useful for identifying protein functions and understanding several cellular pathways, we used a bacterial two-hybrid assay (BACTH) to address the possible binary interactions among the 14 factors involved in the early, commitment, and patterning stages in *Nostoc*. In addition, we performed a phylum-wide analysis of the conservation of these factors in 212 cyanobacterial genomes. A comparative genomics analysis of the genes encoding these factors across cyanobacteria supported the contention that they likely were present in the common ancestor of cyanobacteria and that they evolved through gene loss in clades of cyanobacteria that do not form heterocysts. Taken together, our data suggests that a set of 12 proteins (DevH, HetC, HetF', HetL, HetP, HetR, HetZ, PatA, PatB, PatL, PatN, and PatU3), which encoding genes are conserved in heterocyst-forming cyanobacterial genomes, are associated through multiple cross-talk interactions. Our findings provide a framework for a better understanding of heterocyst pattern establishment.

# Materials and methods

## Bacterial strains and growth media

The *E. coli* K-12 strain DH5 alpha (Invitrogen) was used in all the cloning steps. BACTH complementation assays were carried out with the *E. coli cya* strain BTH101 (F− , *cya-99*, *araD139*, *galE15*, *galK16*, *rpsL1* (Strr ), *hsdR2*, *mcrA1*, *mcrB1*) (Karimova et al. 1998). Bacteria were grown at 30°C in Luria-Bertani (LB) broth supplemented with ampicillin at 100 µg/ml and kanamycin at 50 µg/ml when needed. Screening for the ability to complement the *cya* mutation was performed on LB agar plates supplemented with 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-Gal; 40 µg/ml), isopropyl-β-D-galactopyranoside (IPTG; 0.5 mM), and the appropriate antibiotics. The *E. coli* K-12 Lemo21(DE3) strain (Biolabs) was used for protein production and purification. Spectinomycin (50 µg/ml) was used for the selection of pCDF Duet plasmid (Novagen) and derivatives. Kanamycin (100 µg/ml) for the selection of pLIC07 plasmid (Michel-Souzy et al. 2018) and derivatives.

## Plasmid constructions

For BACTH analysis: the genes involved in heterocyst differentiation and patterning were amplified from genomic DNA of *Nostoc* using the primers listed in Table S1. The recombinant plasmids were constructed by inserting gene sequences of interest in pKT25 (T25 at N-terminus), pKNT25 (T25 at the C-terminus), pUT18 (T18 at the C-terminus) and PUT18C (T18 at the N-terminus) (Karimova et al. 1998). All the open reading frames (ORFs) of the analyzed genes, except *patN*, were inserted between Pst1 and EcoRI restriction sites of the BACTH plasmids. *patN* ORF was inserted between PstI and Kpn1 restriction sites.

For pull-down assays: the *hetR*, *hetL* or *hetF'* sequences fused at their 3' extremities to the sequence coding the 89-106 amino acids of the human influenza hemagglutinin protein (HA-tag) were synthesized by Genewiz (https://www.genewiz.com/en-GB/). The obtained genes were inserted between NdeI and XhoI restriction sites of the pCDF Duet plasmid. The *patB* coding sequence was fused to a 6xHis tag at its 5'extremity. For this, the *patB* ORF was amplified by PCR using pLIC07-patB-Fwd and pLIC07-

patB-Rev primers (Table S1) and inserted between the BsaI restriction sites of the pLIC07 plasmid using the SLIC method (Li and Elledge 2012).

All the constructs were analyzed by sequencing. The plasmids used in this study are listed in Table S2.

**SDS-PAGE and immunodetection**

Proteins were separated on 4-20% polyacrylamide gradient gels (NuSep) and transferred onto nitrocellulose membranes using a semidry blotting apparatus. Membranes were blocked with 3% bovine serum albumin (BSA) in TBST (10 mM Tris, 150 mM NaCl, 0.05% Tween 20). Membranes were incubated with monoclonal anti-HA (Enzo) or anti-His (Qiagen) antibodies, diluted at 1:5000 and 1:2000, respectively in TBST, 3% BSA. After three-step washes in TBST, the membranes were incubated with peroxidase-coupled anti-mouse antibodies (1:5000, Promega). The detection was performed using the enhanced chemiluminescence reagents (Cytiva). The membranes were scanned using ImageQuant TL analysis software (GE Healthcare).

**Coproduction of proteins in *E. coli* and Nickel-affinity chromatography**

Lemo21 (DE3) competent cells were co-transformed with pCDF Duet, pLIC07, and their derivatives. The recombinant clones were grown until reaching an optical density at 600 nm of 0.5 at 37 °C in LB medium with appropriate antibiotics (50 µg/ml spectinomycin, 100 µg/ml kanamycin). Induction was performed with 0.5 mM IPTG for 3 hours at 30 °C. Bacteria were collected by centrifugation and broken using a French press (10,000 p.s.i.) in cold buffer (100 mM Tris-HCl, pH 7.3, 400 mM NaCl, 0.5 mg/ml lysozyme, 20 µg/ml DNase, 1X cOmplete™ EDTA-free protease inhibitor cocktail (Roche). The lysate was cleared by ultracentrifugation. The cleared lysate was loaded onto a 1-ml HiFliQ Ni-NTA FPLC column (Generon) using an ÄKTA Prime apparatus. After a washing step (10 ml of buffer: 100 mM Tris-HCl, pH 7.3, 400 mM NaCl), the immobilized proteins were eluted in buffer: 100 mM Tris-HCl, pH 7.3, 400 mM NaCl, 500 mM imidazol. The loaded flow-through, and elution fractions were analyzed by SDS-PAGE and immunodetection.

## Bacterial two hybrid assays

Bacterial two-hybrid assays were performed following the procedure described by Karimova et al (1998) (Karimova et al. 1998). Briefly, after co-transforming the BTH101 strain with the two plasmids expressing the T18- and T25- fusions, LB plates containing ampicillin and kanamycin were incubated at 30° C for 2 days. For each assay, 10 independent colonies were inoculated in 3 ml of LB medium supplemented with ampicillin, kanamycin, and 0.5 mM IPTG and incubated at 30 °C overnight. ß-galactosidase activity was then determined as previously described (Zubay et al. 1972).

## Bioinformatics analysis

### Datasets

The genome data of 212 cyanobacterial strains available in March 2019 were downloaded from NCBI ftp site (ftp:/ftp.ncbi.nih.gov/genomes/) and constituted the primary data source (Table S3, Sheet 1). This data source includes all complete genomes in the highest levels of assembly, their refseq category (reference or representative genome, or none) as well as their annotation features and taxonomy lineages. The ability of each strain to form heterocysts was provided by the literature (Rippka R 1979; Henson B.J 2004; Shih et al. 2013; Komárek J 2014; Hammerschmidt et al. 2021). Hidden Markov Models (HMMs) of protein family profiles (Pfam (version 33.0, February 2020) and Pgap (version 4, November 2020) were downloaded from http://ftp.ebi.ac.uk/pub/databases/Pfam/releases/ and ftp://ftp.ncbi.nlm.nih.gov/hmm/4.0/ ftp sites, respectively. Proteins involved in cell differentiation in *Nostoc* used as reference seed proteins (retrieved from Uniprot database, www.uniprot.org), are indicated in Table S4.

### Identification of homologs

HMMER package (Mistry et al. 2013) and HMM domain profiles were first used to establish the list of seed functional domains associated to reference seed proteins (Table S4). Alignments with a score higher than the Pfam trusted thresholds were considered as significant seed domains. Then, HMMER3 (Mistry et al. 2013) and self-written Perl scripts were used to search for protein homologs (with reference seed domains) in the complete genomes. For each reference protein, the presence of one or more seed domains was a requisite to identify putative homologs (Table S4). As described above, alignments with

a score higher than the Pfam thrusted thresholds were also considered significant. Note that HetP protein and its 3 homologs in *Nostoc* were considered as a unique seed protein HetP since they share an identical functional domain (Table S4). In contrary, the two seed proteins HetF and HetF' were analyzed independently since their domain organizations are different (Table S4). For each reference protein, putative homologs were subsequently analyzed with the same software in order to determine the potential presence of additional functional domains. In-house Perl scripts were used to define the domain organization of each homolog. Putative homologs were then selected for the strictly presence of seed domains (i.e., in the right order and without additional domain), and therefore led to the set of Hmm-homologs. To select specific homologs, additional criteria were added for the selection of HetC, HetL and PatL. For this, we considered that to be a HetC homolog, the gene must be located in the same genomic context as *hetP* (based on the co-localization of these two genes in the genome of *Nostoc*, see Table S4) with a maximal distance of 5 genes upstream or downstream from *hetP*. HetL and PatL homologs were limited to protein homologs sharing similar features with their seed sequences [a protein length between 0.8 and 1.25 times when compared to the reference seed proteins (Table S4) and a number of repeats fixed to 3-5, and 4-6, respectively]. Since PatN and PatU3 do not exhibit a known functional domain, a Blast (Altschul et al. 1990) analysis search was performed using all reference seed proteins as queries and an *E*-value of $10^{-5}$. Blast alignments with a MinLrap [Alignment size / Minimun size (Query, Subject)] >= 0.8 and MaxLrap [Alignment size / Minimun size (Query, Subject)] >= 0.8 were considered as significant and constituted the set of Blast-homologs. For PatU3 and PatN, for which no seed functional domain was found, MinLrap and MaxLrap were set to 0.7. Hmm-homolog and Blast-homolog sets were finally combined in order to yield the final set of homologs for each reference protein. The location of transmembrane domains in the proteins was performed using the TMHMM software (Krogh et al. 2001).

**Statistical analysis.**

β-galactosidase activities: the values presented are means of 3 independent assays containing each 10 independent colonies. The results are expressed as the means ± standard variation.

Genomics analysis: one-sample *t*-test as well as boxplot, heatmap and distribution graphs were done using the *R* (R Core Team, 2018). For boxplot analysis, taxonomic orders with low number of genomes (i.e., Chroococcidiopsidales, Gloeobacterales, Gloeoemargaritales, Pleurocapsales, Spirulinales, and unclassified Cyanobacteria) were fused into a single group (named Others) before computation (Table S3, Sheet 4).

## Results

To characterize the physical interactions between proteins that are involved in heterocyst differentiation and pattern establishment in *Nostoc*, 13 full-length proteins (i.e., DevH, HetC, HetF, HetL, HetP, HetR, HetZ, PatA, PatB, PatD, PatL, PatN, and PatU3) were tested systematically for pairwise interactions using BACTH. We also included the All1730 ORF, which potentially encodes a HetF homolog (named HetF') and the expression of which (i.e., all1730) has been shown to be induced in the same manner as HetF after combined nitrogen starvation, as demonstrated by RNA-deep-sequencing experiments (Flaherty et al. 2011).

As for the BACTH assay the T18 and T25 domains must be in the cytosol, we first determined the functional and structural domain composition of the studied proteins, as described in the Materials and Methods. This allowed the definition of a subset of membrane proteins (HetC, HetF, PatL, and PatN) and cytosolic factors (DevH, HetF', HetL, HetP, HetR, HetZ, PatA PatB, PatD, and PatU3) (Fig. 2, Table S4). Each protein was fused to a part of the catalytic domain of the chimeric adenylate cyclase (T25 or T18) of *Bordetella pertussis*, at the C and N termini, when possible. According to the predicted topology of the membrane proteins, insertions were made at one terminus to have the T18 and T25 domains in the cytosol. Therefore, membrane proteins (HetC, HetF, PatL, and PatN) were fused to T18 and T25 at the extremity of their domain, which was predicted to be in the cytoplasm (Fig. 2). Direct interactions were screened by transforming all possible pairwise combinations into the *E. coli* BTH101 strain (500 in total). Protein interactions were probed as explained in the Materials and Methods section. In addition to the negative T25–T18 negative-control combination, all constructs were tested using a relevant control vector (as an example: T25-HetR+ NT18 and NT25 + T18-HetR served as negative

controls for HetR) and only results above the control values were considered as being positive in the remainder of the study (Fig. 3 and 4). This control is important for the elimination of self-activation values. In addition to the BACTH T18-Zip and T25-Zip positive control, the well-established interactions between HetR subunits and HetL subunits, and between HetR and HetL, served as positive internal controls. The pairwise combinations that yielded positive values are displayed in histograms (Fig. 3 and 4), whereas negative results are grouped in Table 1.

## Network analysis of early factors

The major regulator of heterocyst differentiation, HetR, displays a specific fold that has been proposed to allow its interaction with other factors (Kim et al. 2011; Hu et al. 2015). In fact, HetR has been shown to interact with homologs of HetP (Videau et al. 2016), with the serine/threonine kinase Pkn22 (Roumezi et al. 2019), and with HetL (Xu et al. 2020). The potential interaction between HetR and other factors acting early after the initiation step and during pattern establishment was analyzed. Among the factors that are induced early, HetR, HetL, and HetF' were found to interact with the PatB transcription factor, as the β-galactosidase activities obtained were 10- to 40-fold higher than those of the negative controls (Fig. 3a, b). In addition, interactions between HetL (Fig. 3b) and HetF' and multimerization of HetF' (Fig. 3a) were observed. A weaker interaction (80 β-galactosidase units) between HetF' and DevH was also identified (Fig. 3a). Because HetF' has not been studied to date, these results suggest the possible implication of this protein in the differentiation process, which is worth exploring. Among this group of factors, the assays involving HetC and HetF all yielded negative results (Table 1)

## Network analysis of commitment and patterning factors

Among the three factors (HetP, HetZ, and PatU3) that formed the "commitment group", a moderate but significant interaction was observed between HetP and PatB (Fig. 4a). The highest β-galactosidase values (400 units) were obtained for the HetP-T25/PatB-T18 combination. In addition, BACTH data confirmed the previously published interaction between HetZ and PatU3 (Du et al. 2020), as two combinations (T25-HetZ/PatU3-T18 and PatU3-T25/T18-HetZ) showed average β-galactosidase activities of 300 units (Fig. 4a).

The patterning factors consist of PatA, PatD, PatB, PatL, and PatN. Our analysis also included DevH in this group, because it is required for *patB* transcriptional expression (Kurio et al. 2020). In addition to the interactions with HetR, HetL, HetF', and HetP mentioned above, PatB formed multimers and interacted with PatL in the PatL-T25/PatB-T18 combination (Fig. 4b). Moreover, PatA interacted with PatL, which is in agreement with published data obtained using the yeast-two-hybrid approach (Liu and Wolk 2011). PatD and PatN did not show any interaction in our assays (**Table 1**).

## The patterning protein network of *Nostoc* as defined by BACTH

Among the 14 factors analyzed here, 10 were engaged in protein-protein interactions. In total, 16 interactions were detected, including self-associations (i.e. multimerization) (Fig. 5). Five proteins (HetR, HetL, HetF', PatB, and PatL) were able to associate with multiple partners (at least two). The interactions identified form a network that connects the elements involved in the initiation stage of differentiation, commitment, and patterning. As expected, early and commitment factors mostly interact among themselves, whereas the patterning factors, in particular PatB (with five interactions), appear to be the central network and, therefore, serve as the master link between the three developmental stages (Fig. 5). In addition, HetL and HetF' displayed three interactions each, suggesting that these proteins are also important key players in the interaction network.

### Validation of the interaction between PatB and the early factors

To validate our BACTH procedure, we sought to confirm some of the protein-protein interactions described above by another approach. Since PatB appears to be central in the network, we decided to analyze its ability to interact with the early factors by pull-down assays. PatB was produced as a $His_6$-tagged fusion and its capacity to bind nickel column was verified by affinity chromatography purification (Fig. S1a). HetR, HetL and HetF' fused to the HA-tag were produced as shown by the immunoblots of Fig. S1b. To proceed with the pull-down assay, $His_6$-PatB and each of the HA-tagged proteins were coproduced in *E. coli* and protein purifications were performed using Nickel-affinity chromatography. The analysis of the elution fractions by SDS-PAGE followed by immunoblotting with anti-$His_6$ and anti HA antibodies showed a specific co-elution of HetR, HetL and HetF' with PatB (Fig.

6, left panels). To control for non-specific binding of the HA-tagged proteins, HetR-HA, HetL-HA and HetF'-HA produced in the absence of $His_6$-PatB were purified following the same procedure. The data presented in Fig. 6, right panels show that none of them was detected in the elution fractions, indicating that they were unable to bind to the column. Altogether these results reveal direct interactions between PatB and the three early factors and thus confirm the data obtained by BACTH.

## Evolutionary conservation of genes underlying heterocyst formation and patterning

To address the general relevance in cyanobacteria of the interaction network described above, a large-scale (212 complete genomes in total) *in silico* identification and analysis of the set of the 14 protein factors studied was performed (see Materials and methods and Table S4). Note that PatS, PatX and PatC proteins were not included in our *in silico* study. The reasons are that these proteins are too short, do not contain any known functional domain and in particular, *patC* was only present in *Nostoc* PCC7120 and *Trichormus variabilis* NIES-23 genomes. These proteins are thus not compatible with significant sequence comparisons based on functional domains search.

Interestingly, this genome survey detected 63 (for PatN) to 779 (for HetL) homologs within the genomes (Table S3, Sheets 2 and 3). This large variation in the total number of homologs for each reference protein reflects the wide variation in the number of copies of each gene per genome. In fact, the distribution of the number of copies within a genome (Table S3, Sheet 3; Fig. S2) showed that: (*i*) PatN, HetP, PatB, HetF', HetR, PatU3, HetZ, and HetF did not have genes encoding the corresponding homologs in a high number of genomes (≥80); (*ii*) when present, they were on average present in ~1 copy (for PatN, PatB, HetF', HetR, and PatU3) or ~2 copies (for HetZ and HetF); and (*iii*) HetL group of homologs displayed up to 4 copies per genome on average, with 13 copies encoding orthologs for HetL in *Microcoleus* sp. PCC 7113. Moreover, HetL, which is encoded by a gene that is present in multiple copies within most of the cyanobacterial genomes, harbored a low copy number of homologs in Prochloraceae, with no copies detected in some *Prochlorococcus marinus* strains (e.g., *P. marinus* str. NATL2A, *P. marinus* str. MIT9515, and *P. marinus* str. MIT 9211). Interestingly, genes encoding HetR homologs were present in 2 or 3 copies in some genomes (e.g., *Nodosilinea nodulosa* PCC 7104 or *Prochlorothrix hollandica* PCC 9006, respectively). More surprisingly, most of the Nostococaceae

and Aphanizomenonaceae strains displayed homolog(s) for each group of reference proteins, except for few organisms (e.g., *Cylindrospermopsis raciborskii* CS-505 and *Dolichospermum circinale* AWQC310F) for which no homolog of PatU3, PatB, and HetF was found.

Next, we asked whether a relationship exists between the number of homologs in an organism and its taxonomic lineage and, therefore, its ability to form heterocysts. For this purpose, a heatmap graphical representation of the number of homologs within organisms, according to the taxonomic orders, was performed (Fig. 7). Our analysis showed that except for HetC, nearly all the Nostocales contain at least one gene encoding a homolog of each reference group, with a variable number of homologs per organism, as described above. Interestingly, PatN seemed to be specifically present in Nostocales exclusively. To a lesser extent, HetP and PatB also appeared to be poorly represented outside the Nostocales group (Fig. 7). The distributions of the genes encoding HetC, HetF', HetP, HetZ, PatU3, PatB and PatL homologs over the cyanobacterial lineage were similar to the one of *hetR*, suggesting that they might share a similar evolutionary history from the ancestor of cyanobacteria. Considered this observation and the central and essential role of HetR within the heterocyst differentiation process, it can be suggested that HetC, HetF'; HetP, HetZ, PatU3; PatB, PatL, and PatN might also act as key factors for the early, commitment, and patterning stages, respectively, in Nostocales. Fig. 7 also shows that HetL, DevH, PatA, and PatD are conserved over the entire cyanobacterial taxonomic lineages, which suggests that they may be involved in biological processes that are common to all these prokaryotes. Unveiling the mechanisms beyond their speciation for heterocyst formation in the Nostocales will be a significant progress in this field.

Another important question that needed to be addressed was whether the distribution of the studied genes in Nostocales was significantly distinct compared with other taxonomic orders. For this, we computed for each reference group of proteins the average number of genes encoding homologs per genome for each cyanobacterial order (Materials and methods). These data are shown in the boxplots of Fig. 8a (see also Table S3, Sheet 4). With the exception of *patN*, the other analyzed genes exhibited a normal distribution among cyanobacterial orders, according to the Shapiro–Wilk test. As expected, HetL and DevH exhibited a high number of homologs on average, regardless of the cyanobacterial taxonomic

orders. Interestingly, the presence of HetC, HetP, PatA, PatB, PatL and of PatN, was significantly higher (*P*-value < 0.01) in Nostocales compared with other taxonomic orders (Fig. 8a). The high prevalence of these six studied genes in Nostocales indicates their functional importance in heterocyst-forming cyanobacteria, and are also considered as key factors genes. In addition, the data presented in this Fig. highlight again the specific occurrence of PatN in genomes of Nostocales strains.

Altogether, when combined the two comparative genomics analyses (Fig. 7 and 8a), the overall key factors consist of a network encompassing 10 proteins (HetC, HetR, HetF', HetP, HetZ, PatU3, PatA, PatB, PatL, and PatN) that subsequently refers to the patterning factor network of heterocyst-forming cyanobacteria from an evolutionary point of view.

## Discussion

In this study, we combined protein–protein interaction and genomics analyses to determine the physical and evolutionary relationships that exist between the factors involved in heterocyst development and patterning. BACTH has proven to be a powerful approach for probing interactions among the proteins involved in the same process. For instance, it allowed the deciphering of the interactions that occur between the membrane proteins involved in cell division in *E. coli* (Karimova et al. 2005), in cell wall synthesis in *Corynebacterium glutamicum* (Jankute et al. 2014), and in spore coat formation in *Bacillus subtilis* (Krajcikova et al. 2017). In eukaryotes, a yeast-two-hybrid system was used to determine the physical interactions between proteins involved in vulva development in *Caenorhabditis elegans* (Walhout et al. 2000), and this study was the starting point for the elucidation of the global interactome of this organism (Li et al. 2004). In a similar way, our study may help reveal the global interactome in *Nostoc*. The fact that the interactions observed by BACTH for two of the analyzed proteins (HetL and HetR) were confirmed by a BioLayer interferometry (Xu et al. 2020) and that pull-down assay confirmed the interaction of PatB with 3 factors (Fig. 6) validates the BACTH approach. One of the interesting perspectives of this work will be the study of all these possible interactions in *Nostoc* throughout the differentiation process. The development of techniques such as bimolecular fluorescence

complementation (BifC) (Kerppola 2006) and mass spectrometry (Gingras et al. 2007) should help achieve this objective. This study represents an important step-forward future works aiming to reveal the interactions occurring in the heterocyst. We propose to refer to the protein network involved in pattern formation as the "patternome".

The possibility that proteins with transmembrane segments may not be correctly inserted in the membrane of *E. coli* constitutes a limitation of the BACTH assays. We did not observe any interactions for the HetC, HetF, and PatN proteins, which are predicted to be membrane proteins. In the case of HetC, in addition to this prediction, its localization at the inner membrane of the heterocyst has been proposed, as assessed using fusion to the green fluorescent protein (Corrales-Guerrero et al. 2014a). Similarly, membrane localization has been indicated for PatN in *Nostoc punctiforme* (Risser et al. 2012). The possibility that the lack of interactions of these proteins could be related to an incorrect topology of the transmembrane domains in *E. coli* cannot be ruled out. HetF has recently been shown to be part of the divisome of the vegetative cells under specific light conditions (Xing et al. 2021). The fact that no interaction was found with the proteins involved in heterocyst formation and patterning might therefore reflect that it does actually not belong to the protein network formed by these factors.

Fourteen of the 16 protein interactions observed in our BACTH analysis were detected for the first time. These new data provide additional clues about the molecular mechanisms underlying patterning, as they show that the three steps of the process are likely to be connected through physical interactions between key factors (Fig. 5). Strikingly, our results indicate that the transcriptional regulator PatB interacts with multiple partners, which renders it a central factor in the network (Fig. 3, 4, and 6). Among the cytosolic proteins, PatD alone was shown to not interact with any other factor (Fig. 5). This result agrees with the hypothesis that the effect of PatD on patterning might occur indirectly through general regulators, such as ppGpp (Wang et al. 2019). The BACTH analysis also revealed that PatA interacts with two proteins that are required for cell division, ZipN and SepF (Valladares et al. 2020). As PatA also interacts with PatB and PatL, taken together, these data show that the physical protein interactions described here may impact the evolution of the heterocyst toward its terminal differentiation state (i.e., the arrest of cell division). Contrary to what has been published (Du et al. 2020), we did not observe any interaction

between HetR and HetZ (Table 1). The facts that we systematically tested combinations of the fused constructs with the BACTH control vectors as negative controls (Fig. 4a) and that we considered as being positive only values that were higher than those obtained for these controls may explain the discrepancy between our data and the published ones. The genomics analysis indicated that the 14 genes analyzed here are highly conserved in the genomes of heterocyst-forming cyanobacteria **(**Fig. 7); therefore, an evolutionary history of these genes can be drawn. In fact, the location of almost all genes over the taxonomic lineages suggests that they were present within the genome of the cyanobacteria common ancestor. Throughout the evolution process, some of these genes were lost and/or functional specialization might have occurred for those present in multiple copies (e.g., HetL). The process of gene loss appears to be faster in some lineages than others, with Chroococcaceae, Gloeomargaritaceae, Gloeobacteraceae, and Synechococcaceae being the clades in which the gene losses are the most pronounced. Thus, gene loss occurs in clades encompassing strains that do not exhibit cell differentiation, whereas Nostococcales genomes have conserved this set of genes **(**Fig. 7**)**. Regarding functional specialization, the most emblematic example is that of HetF and HetF', as the present copy of HetF would probably be closer to the ones of the common ancestor. If future investigation unveils a role for HetF' in the differentiation process, then one might suggest that its functional specialization during evolution would have consisted in a loss of the transmembrane region, followed by the acquisition of the Forkhead-associated domain (FHA). As the FHA domain has been maintained during evolution, its importance in the differentiation process and in the interaction of HetF' with PatB, DevH, and HetL may be questioned.

The heatmap generated here revealed that various levels can be proposed for the evolutive and functional importance of the analyzed proteins in each step of the differentiation process. Among the early factors, HetF', HetR, and HetC are the major evolutionarily conserved proteins; *hetF'* and *hetR* for their single copy in the genome, and *hetC* for a very significant presence bias in favor of the Nostocales compared with the other orders (Fig. 7). The essential role of HetR in *Nostoc* differentiation is fully understood. As mentioned above, both the protein interaction analysis and the genomics survey suggest that the HetF' protein is an interesting factor that deserves investigation. Genetic analyses revealed the

involvement of HetC and its N-terminal domain in heterocyst differentiation (Khudyakov and Wolk 1997; Xu and Wolk 2001; Corrales-Guerrero et al. 2014a). However, whether HetC actually acts as an ABC transporter remains to be demonstrated. A comparative and evolutionary analysis of HetP, HetZ, and PatU3 showed that all of them are key players in the commitment step, even if there is a significant presence bias in favor of HetP in the Nostococales. It has been proposed that cell-fate commitment in *Nostoc* is modulated by a hierarchy of four homologous HetP proteins, with HetP promoting cell differentiation, whereas the HetP-like proteins seem to exert an opposite control (Videau et al. 2016). These observations combined with the presence of multiple copies of HetP-encoding genes in Nostocales clearly demonstrate that, during the evolutionary time, some HetP paralogs underwent a neofunctionalization process, leading to new functions. The interactions of HetP with PatB and PatU3 with HetZ indicate that the protein complexes formed might explain the role of these proteins in connecting the initiation step with subsequent steps. From an evolutionary point of view, PatA, PatB, PatL and PatN appear as the key players among the patterning factors. The presence of PatN exclusively in Nostocales seems to indicate that its action is important for the formation of heterocysts in all organisms of this clade. Therefore, the elucidation of the molecular mechanism of PatN is crucial for the understanding of this process. The significant enrichment of PatA in the Nostocales also makes it a key factor that deserves further attention. In the protein network, PatB occupies a central position as it interacts with factors belonging to the three defined groups and is mostly present in heterocyst-forming cyanobacteria. Little is known about the mechanism of this regulator and about the genes that are controlled by it in *Nostoc*, although it appears from our study that it deserves additional attention. Among all the key players defined here, 8 out of 10 exhibited interactions with each other and with other partners. This shows that protein–protein interactions may play pivotal roles in heterocyst differentiation. Therefore, unveiling their dynamics through this process is a determining direction for the understanding of development in these prokaryotes.

Given that three factors in the interaction network are transcriptional regulators (HetR, DevH, and PatB), interactions occurring inside the network likely impact the transcriptional program of the developing cell. To obtain insights into this interconnection, we first referred to the literature to build the

transcriptional network (Table S5). Indeed, HetR affects the transcription of genes belonging to the three groups (early events, commitment, and patterning) by acting either directly inside the network (effect on *hetZ*, *hetP*, *patA*, and likely *devH*) or indirectly through the regulation of the expression of the global regulator NtcA (effect on *hetC* and *patD*) (see references cited in Table S5). DevH directly controls the transcription of *patB* (Kurio et al. 2020), which was suggested to act on genes involved in the maturation step (Brenes-Alvarez et al. 2019). Secondly, we combined the network of both protein-protein interactions and comparative genomics, with the transcriptional one, leading to an integrative model of the three distinct approaches on heterocyst differentiation genes (Fig. 8b). From this, several facts can be pinpointed: (i) the activities of the regulators connect the 14 genes through transcriptional relationships, thus rendering the protein and transcriptional networks interdependent; (ii) HetR, PatB and PatA remain the factors that exhibit the highest numbers of biological links; (iii) based on the number of biological links, HetZ and PatB can be considered as the main actors among the commitment and patterning factors, respectively; (iv) cross-talks link early, commitment and patterning factors with 3 key players (as defined by *in silico* analysis) in each differentiation stage. It is important to remind here that in addition to these factors, acting inside the heterocyst, the patterning requires HetN, PatS and PatX which are supposed to generate the diffusible morphogens essential for lateral inhibition along the filaments (Fig. 1a).

**In conclusion**, our study questioned the possible physical and evolutionary interactions between 14 proteins involved in heterocyst formation and patterning in *Nostoc*. BACTH approach was used to analyze the protein-protein interactions and pull-down assays were used to confirm three of the observed interactions. Combining the data obtained to a genomic survey of these 14 genes in cyanobacterial genomes reveals that a core network of 12 proteins (all the analyzed proteins factors except HetF and PatD) may play a crucial role in heterocyst development and patterning not only in *Nostoc* but within all the clade of heterocyst-forming cyanobacteria (Fig. 8b). The results of this study open up interesting perspectives for the elucidation of the mechanism of action of all these proteins and the importance of their interactions in establishing the pattern.

## Acknowledgments:

The authors thank Dr Badreddine Douzi and Dr Romé Voulhoux for helpful discussions regarding pull-down assays. They also thank Anais Scholivet for technical assistance.

## Funding:

X. Xu was funded by a fellowship from the Chinese government. R. Rachedi has a fellowship from the Région SUD and CNRS. The project was supported by the "Agence Nationale pour la Recherche Scientifique" (ANR-21-CE20-0025-01) and by funding from AMU.

## Ethics declarations:

## Conflict of interest

The authors declare that they have no conflict of interest.

## Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

## Informed consent

All the authors have consent for publication.

# REFERENCES


Abed RM, Dobretsov S, Sudesh K (2009) Applications of cyanobacteria in biotechnology. J Appl Microbiol 106:1-12

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403-410

Black TA, Cai Y, Wolk CP (1993) Spatial expression and autoregulation of hetR, a gene involved in the control of heterocyst development in Anabaena. Mol Microbiol 9:77-84

Black TA, Wolk CP (1994) Analysis of a Het- mutation in Anabaena sp. strain PCC 7120 implicates a secondary metabolite in the regulation of heterocyst spacing. J Bacteriol 176:2282-2292

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

Buikema WJ, Haselkorn R (1991) Characterization of a gene controlling heterocyst differentiation in the cyanobacterium Anabaena 7120. Genes Dev 5:321-330

Corrales-Guerrero L, Flores E, Herrero A (2014a) Relationships between the ABC-exporter HetC and peptides that regulate the spatiotemporal pattern of heterocyst distribution in Anabaena. PLoS One 9:e104571

Corrales-Guerrero L, Mariscal V, Flores E, Herrero A (2013) Functional dissection and evidence for intercellular transfer of the heterocyst-differentiation PatS morphogen. Mol Microbiol 88:1093-1105

Corrales-Guerrero L, Mariscal V, Nurnberg DJ, Elhai J, Mullineaux CW, Flores E, Herrero A (2014b) Subcellular localization and clues for the function of the HetN factor influencing heterocyst distribution in Anabaena sp. strain PCC 7120. J Bacteriol 196:3452-3460

Du Y, Zhang H, Wang H, Wang S, Lei Q, Li C, Kong R, Xu X (2020) Expression from DIF1-motif promoters of hetR and patS is dependent on HetZ and modulated by PatU3 during heterocyst differentiation. PLoS One 15:e0232383

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

Feldmann EA, Ni S, Sahu ID, Mishler CH, Risser DD, Murakami JL, Tom SK, McCarrick RM, Lorigan GA, Tolbert BS, Callahan SM, Kennedy MA (2011) Evidence for direct binding between HetR from Anabaena sp. PCC 7120 and PatS-5. Biochemistry 50:9212-9224

Fernandez-Pinas F, Leganes F, Wolk CP (1994) A third genetic locus required for the formation of heterocysts in Anabaena sp. strain PCC 7120. J Bacteriol 176:5277-5283

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

Frias JE, Flores E, Herrero A (1994) Requirement of the regulatory protein NtcA for the expression of nitrogen assimilation and heterocyst development genes in the cyanobacterium Anabaena sp. PCC 7120. Mol Microbiol 14:823-832

Gingras AC, Gstaiger M, Raught B, Aebersold R (2007) Analysis of protein complexes using mass spectrometry. Nat Rev Mol Cell Biol 8:645-654

Hammerschmidt K, Landan G, Domingues Kummel Tria F, Alcorta J, Dagan T (2021) The Order of Trait Emergence in the Evolution of Cyanobacterial Multicellularity. Genome Biol Evol 13

Henson B.J HSM, Watson L.E, Barnum S.R (2004) Molecular phylogeny of the heterocystous cyanobacteria (subsections IV and V) based on nifD. Journal of Systematic and Evolutionary Microbiology 54:493–497

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

Higa KC, Callahan SM (2010) Ectopic expression of hetP can partially bypass the need for hetR in heterocyst differentiation by Anabaena sp. strain PCC 7120. Mol Microbiol 77:562-574

Higa KC, Rajagopalan R, Risser DD, Rivers OS, Tom SK, Videau P, Callahan SM (2012) 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 83:682-693

Hu HX, Jiang YL, Zhao MX, Cai K, Liu S, Wen B, Lv P, Zhang Y, Peng J, Zhong H, Yu HM, Ren YM, Zhang Z, Tian C, Wu Q, Oliveberg M, Zhang CC, Chen Y, Zhou CZ (2015) Structural insights into HetR-PatS interaction involved in cyanobacterial pattern formation. Sci Rep 5:16470

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

Jankute M, Byng CV, Alderwick LJ, Besra GS (2014) Elucidation of a protein-protein interaction network involved in Corynebacterium glutamicum cell wall biosynthesis as determined by bacterial two-hybrid analysis. Glycoconj J 31:475-483

Jones KM, Buikema WJ, Haselkorn R (2003) Heterocyst-specific expression of patB, a gene required for nitrogen fixation in Anabaena sp. strain PCC 7120. J Bacteriol 185:2306-2314

Karimova G, Dautin N, Ladant D (2005) Interaction network among Escherichia coli membrane proteins involved in cell division as revealed by bacterial two-hybrid analysis. J Bacteriol 187:2233-2243

Karimova G, Pidoux J, Ullmann A, Ladant D (1998) A bacterial two-hybrid system based on a reconstituted signal transduction pathway. Proc Natl Acad Sci U S A 95:5752-5756

Kerppola TK (2006) Design and implementation of bimolecular fluorescence complementation (BiFC) assays for the visualization of protein interactions in living cells. Nat Protoc 1:1278-1286

Khudyakov I, Gladkov G, Elhai J (2020) Inactivation of Three RG(S/T)GR Pentapeptide-Containing Negative Regulators of HetR Results in Lethal Differentiation of Anabaena PCC 7120. Life (Basel) 10

Khudyakov I, Wolk CP (1997) hetC, a gene coding for a protein similar to bacterial ABC protein exporters, is involved in early regulation of heterocyst differentiation in Anabaena sp. strain PCC 7120. J Bacteriol 179:6971-6978

Kim Y, Joachimiak G, Ye Z, Binkowski TA, Zhang R, Gornicki P, Callahan SM, Hess WR, Haselkorn R, Joachimiak A (2011) Structure of transcription factor HetR required for heterocyst differentiation in cyanobacteria. Proc Natl Acad Sci U S A 108:10109-10114

Komárek J KJ, Mareš J, Johansen J.R (2014) Taxonomic classification of cyanoprokaryotes (cyanobacterial genera) 2014, using a polyphasic approach. Preslia 86

Krajcikova D, Forgac V, Szabo A, Barak I (2017) Exploring the interaction network of the Bacillus subtilis outer coat and crust proteins. Microbiol Res 204:72-80

Krogh A, Larsson B, von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305:567-580

Kumar K, Mella-Herrera RA, Golden JW (2010) Cyanobacterial heterocysts. Cold Spring Harb Perspect Biol 2:a000315

Kurio Y, Koike Y, Kanesaki Y, Watanabe S, Ehira S (2020) The CRP-family transcriptional regulator DevH regulates expression of heterocyst-specific genes at the later stage of differentiation in the cyanobacterium Anabaena sp. strain PCC 7120. Mol Microbiol 114:553-562

Li MZ, Elledge SJ (2012) SLIC: a method for sequence- and ligation-independent cloning. Methods Mol Biol 852:51-59

Li S, Armstrong CM, Bertin N, Ge H, Milstein S, Boxem M, Vidalain PO, Han JD, Chesneau A, Hao T, Goldberg DS, Li N, Martinez M, Rual JF, Lamesch P, Xu L, Tewari M, Wong SL, Zhang LV, Berriz GF, Jacotot L, Vaglio P, Reboul J, Hirozane-Kishikawa T, Li Q, Gabel HW, Elewa A, Baumgartner B, Rose DJ, Yu H, Bosak S, Sequerra R, Fraser A, Mango SE, Saxton WM, Strome S, Van Den Heuvel S, Piano F, Vandenhaute J, Sardet C, Gerstein M, Doucette-Stamm L, Gunsalus KC, Harper JW, Cusick ME, Roth FP, Hill DE, Vidal M (2004) A map of the interactome network of the metazoan C. elegans. Science 303:540-543

Liang J, Scappino L, Haselkorn R (1992) 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 89:5655-5659

Liu D, Golden JW (2002) hetL overexpression stimulates heterocyst formation in Anabaena sp. strain PCC 7120. J Bacteriol 184:6873-6881

Liu J, Wolk CP (2011) Mutations in genes patA and patL of Anabaena sp. strain PCC 7120 result in similar phenotypes, and the proteins encoded by those genes may interact. J Bacteriol 193:6070-6074

Makarova KS, Koonin EV, Haselkorn R, Galperin MY (2006) Cyanobacterial response regulator PatA contains a conserved N-terminal domain (PATAN) with an alpha-helical insertion. Bioinformatics 22:1297-1301

Meeks JC, Elhai J (2002) Regulation of cellular differentiation in filamentous cyanobacteria in free-living and plant-associated symbiotic growth states. Microbiol Mol Biol Rev 66:94-121; table of contents

Michel-Souzy S, Douzi B, Cadoret F, Raynaud C, Quinton L, Ball G, Voulhoux R (2018) Direct interactions between the secreted effector and the T2SS components GspL and GspM reveal a new effector-sensing step during type 2 secretion. J Biol Chem 293:19441-19450

Mistry J, Finn RD, Eddy SR, Bateman A, Punta M (2013) Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. Nucleic Acids Res 41:e121

Pratte BS, Thiel T (2016) Homologous regulators, CnfR1 and CnfR2, activate expression of two distinct nitrogenase gene clusters in the filamentous cyanobacterium Anabaena variabilis ATCC 29413. Mol Microbiol 100:1096-1109

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

Rippka R D, J, Wterbury JB, Herdman M, Stanier RY (1979) Generic assignments, strain stories and properties of pure cultures of cyanobacteria. . J. Gent. Microbiol. 111:1-61

Risser DD, Callahan SM (2008) HetF and PatA control levels of HetR in Anabaena sp. strain PCC 7120. J Bacteriol 190:7645-7654

Risser DD, Wong FC, Meeks JC (2012) Biased inheritance of the protein PatN frees vegetative cells to initiate patterned heterocyst differentiation. Proc Natl Acad Sci U S A 109:15342-15347

Roumezi B, Xu X, Risoul V, Fan Y, Lebrun R, Latifi A (2019) 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 10:3140

Shih PM, Wu D, Latifi A, Axen SD, Fewer DP, Talla E, Calteau A, Cai F, Tandeau de Marsac N, Rippka R, Herdman M, Sivonen K, Coursin T, Laurent T, Goodwin L, Nolan M, Davenport KW, Han CS, Rubin EM, Eisen JA, Woyke T, Gugger M, Kerfeld CA (2013) Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. Proc Natl Acad Sci U S A 110:1053-1058

Tsujimoto R, Kamiya N, Fujita Y (2014) Transcriptional regulators ChlR and CnfR are essential for diazotrophic growth in nonheterocystous cyanobacteria. Proc Natl Acad Sci U S A 111:6762-6767

Turing A (1952) The chemical basis of morphogenesis. Philosophical Transactions of the Royal Society B 237:37-72

Valladares A, Velazquez-Suarez C, Herrero A (2020) Interactions of PatA with the Divisome during Heterocyst Differentiation in Anabaena. mSphere 5

Videau P, Rivers OS, Higa KC, Callahan SM (2015) ABC Transporter Required for Intercellular Transfer of Developmental Signals in a Heterocystous Cyanobacterium. J Bacteriol 197:2685-2693

Videau P, Rivers OS, Hurd K, Ushijima B, Oshiro RT, Ende RJ, O'Hanlon SM, Cozy LM (2016) The heterocyst regulatory protein HetP and its homologs modulate heterocyst commitment in Anabaena sp. strain PCC 7120. Proc Natl Acad Sci U S A 113:E6984-E6992

Videau P, Rivers OS, Tom SK, Oshiro RT, Ushijima B, Swenson VA, Philmus B, Gaylor MO, Cozy LM (2018) The hetZ gene indirectly regulates heterocyst development at the level of pattern formation in Anabaena sp. strain PCC 7120. Mol Microbiol

Walhout AJ, Sordella R, Lu X, Hartley JL, Temple GF, Brasch MA, Thierry-Mieg N, Vidal M (2000) Protein interaction mapping in C. elegans using proteins involved in vulval development. Science 287:116-122

Wang L, Lin GM, Niu TC, Zhang SR, Zhang JY, Tang GF, Chen W, Zhang CC (2019) patD, a Gene Regulated by NtcA, Is Involved in the Optimization of Heterocyst Frequency in the Cyanobacterium Anabaena sp. Strain PCC 7120. J Bacteriol 201

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

Xing WY, Liu J, Wang ZQ, Zhang JY, Zeng X, Yang Y, Zhang CC (2021) HetF Protein Is a New Divisome Component in a Filamentous and Developmental Cyanobacterium. mBio:e0138221

Xu X, Risoul V, Byrne D, Champ S, Douzi B, Latifi A (2020) HetL, HetR and PatS form a reaction-diffusion system to control pattern formation in the cyanobacterium nostoc PCC 7120. Elife 9

Xu X, Wolk CP (2001) Role for hetC in the transition to a nondividing state during heterocyst differentiation in Anabaena sp. J Bacteriol 183:393-396

Yoon HS, Golden JW (1998) Heterocyst pattern formation controlled by a diffusible peptide. Science 282:935-938

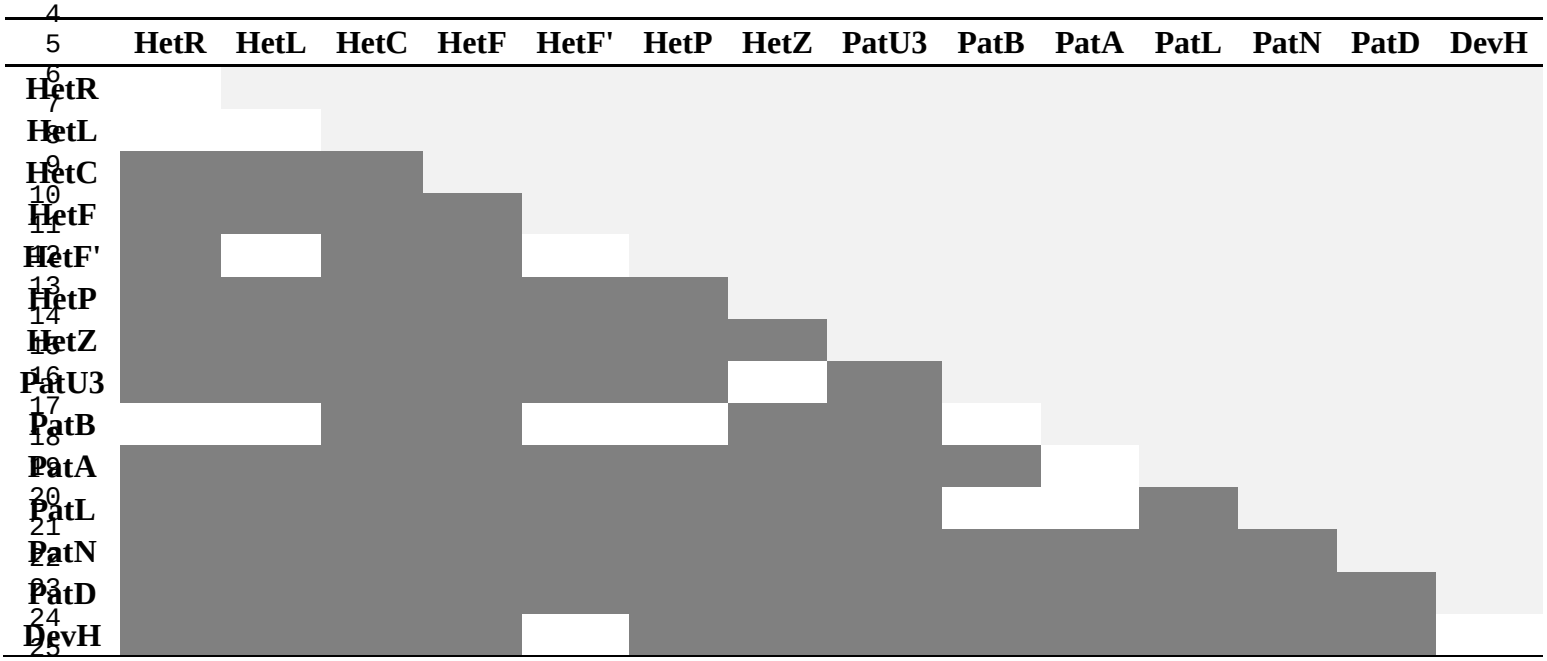
Yoon HS, Golden JW (2001) PatS and products of nitrogen fixation control heterocyst pattern. J Bacteriol 183:2605-2613

Zhang H, Wang S, Wang Y, Xu X (2018) Functional Overlap of hetP and hetZ in Regulation of Heterocyst Differentiation in Anabaena sp. Strain PCC 7120. J Bacteriol 200

Zhang W, Du Y, Khudyakov I, Fan Q, Gao H, Ning D, Wolk CP, Xu X (2007) A gene cluster that regulates both heterocyst differentiation and pattern formation in Anabaena sp. strain PCC 7120. Mol Microbiol 66:1429-1443

Zubay G, Morse DE, Schrenk WJ, Miller JH (1972) Detection and isolation of the repressor protein for the tryptophan operon of Escherichia coli. Proc Natl Acad Sci U S A 69:1100-1103

**Table 1: Matrix of BACTH data highlighting negative interactions between proteins involved in heterocyst differentiation within *Nostoc*.**




Each rectangle of the matrix represents a tested two hybrid interactions (from horizontal and vertical axes). The β-galactosidase activities correspond to the functional complementation of the T18/T25 adenylate cyclase domains following interaction of the two hybrid proteins. β-galactosidase activity lower than 20 Miller units, or lower than the self-activation controls (ie: the fused-construct tested with the control vector) was considered as negative. Those whose interactions were detected as negative by all possible combinations are marked as dark grey. Those whose interactions were detected at least by one combination are marked as blank (see Fig. 3 and 4). The repeat copy of each protein pair in this table is marked as light grey.

## Legends to figures

### Figure 1

**(a)** Local activation/long range inhibition model for heterocyst pattern establishment in *Nostoc*. The red areas indicate the concentration of the inhibitory peptides (derived from PatS, HetN and probably PatX). A: activator (HetR). I: immunity protein (HetL). A/I: complex formed by the interaction between HetR and HetL.

**(b)** Phenotype of mutants of genes involved in heterocyst formation and patterning. o.e. stands for overexpression of the corresponding genes. Δ stands for deletion of the corresponding genes.

### Figure 2

Structural and functional domains located within the reference proteins analyzed in this study.

The domains predicted by Pfam are indicated by blue boxes. For HetR, the domains correspond to those identified by the resolution of the structure (Hu et al. 2015). The topology of the membrane protein was predicted by the TMHMM software (Krogh et al. 2001).
Protein sizes are in parenthesis (aa: amino acids). The frame of each protein is colored corresponding to its functional group in our study (bleu: early factors, yellow: commitment, green: patterning).
TM, Transmembrane domains. Functional domain descriptions are : Crp, Bacterial regulatory proteins, crp family; HTH_Crp, Crp-like helix-turn-helix domain; Peptidase_C39, Peptidase C39 family; ABC_tran, ABC transporter; CHAT, CHAT domain; FHA, FHA domain; Pentapeptide, Pentapeptide repeats (8 or 9 copies); HetP_family, HetP family heterocyst commitment protein; Peptidase_S48, Peptidase family S48; HetR_C, Heterocyst differentiation regulator C-terminal Hood domain; DUF4388, Domain of unknown function (DUF4388); Response_reg, Response regulator receiver domain; heterocyst_HetZ, heterocyst differentiation protein HetZ; HTH_3, Helix-turn-helix; HTH_31, Helix-turn-helix domain; het_cyst_PatD, heterocyst frequency control protein PatD.

### Figure 3

BACTH analysis of protein-protein interactions of factors involved in early stage of differentiation. **(a)** Positive assays of HetR or HetF' interacting with other proteins involved in pattern formation. **(b)** Positive assays for HetL interacting with other proteins involved in pattern formation.
For both, BTH101 strains were transformed with the indicated plasmids. β-galactosidase activities were measured as described in section "Materials and Methods" and were expressed in Miller units. Strains producing T25-Zip and T18-Zip served as positive control (PC). Strains producing the T25 with T18 served as negative control (NC). Strains producing T25-HetR and T18-HetR served as inner positive control. For each combination, the two corresponding self-activation controls were analyzed and shown in gray columns. Blue columns indicate high BACTH signals (more than 300 Miller units); yellow columns indicate intermediate BACTH signals (100-300 Miller units); green columns indicate weak BACTH signals (70-100 Miller units).

### Figure 4

BACTH analysis of protein-protein interactions of factors involved in commitment **(a)** and patterning **(b).**
BTH101 strains were transformed with the indicated plasmids. β-galactosidase activities were measured as described in section "Materials and Methods" and were expressed in Miller units. Strains producing T25-Zip and T18-Zip served as positive control (PC). Strains producing the T25 with T18 served as negative control (NC). Strains producing T25-HetR and T18-HetR served as inner positive control. For each combination, the two corresponding self-activations controls were analyzed and shown in gray columns. Blue columns indicate high BACTH signals (more than 300 Miller units); yellow columns indicate intermediate BACTH signals (100-300 Miller units).

## Figure 5

Protein-protein interaction network of 14 factors involved in heterocyst formation and patterning.

The thickness of the lines is relative to the level of BACTH signals. ▬ : indicate high levels (β-galactosidase values above 300 Miller units). ▬ : indicate intermediate levels (β-galactosidase values between 100-300 Miller units). ▬ : indicates weak levels (β-galactosidase between 70-100 Miller units).

## Figure 6

Interactions between PatB and the early factors analyzed by pull-down assays.
Left panels: Nickel-affinity based copurification and immunoblotting experiments of His-tagged PatB co-produced with HetR-HA (**a**), HetL-HA (**b**) or HetF'-HA (**c**) in *E. coli*.
Right panels: negative controls corresponding to the purification of each HA-tagged protein on a Nickel column, after heterologous production in *E. coli*.
M: molecular size of the protein marker (kDa), L: loading material; FT: flow-through; W: wash fraction; E1–E6: elution fractions.
The antibodies used are indicated at the top of each immunoblot.

## Figure 7

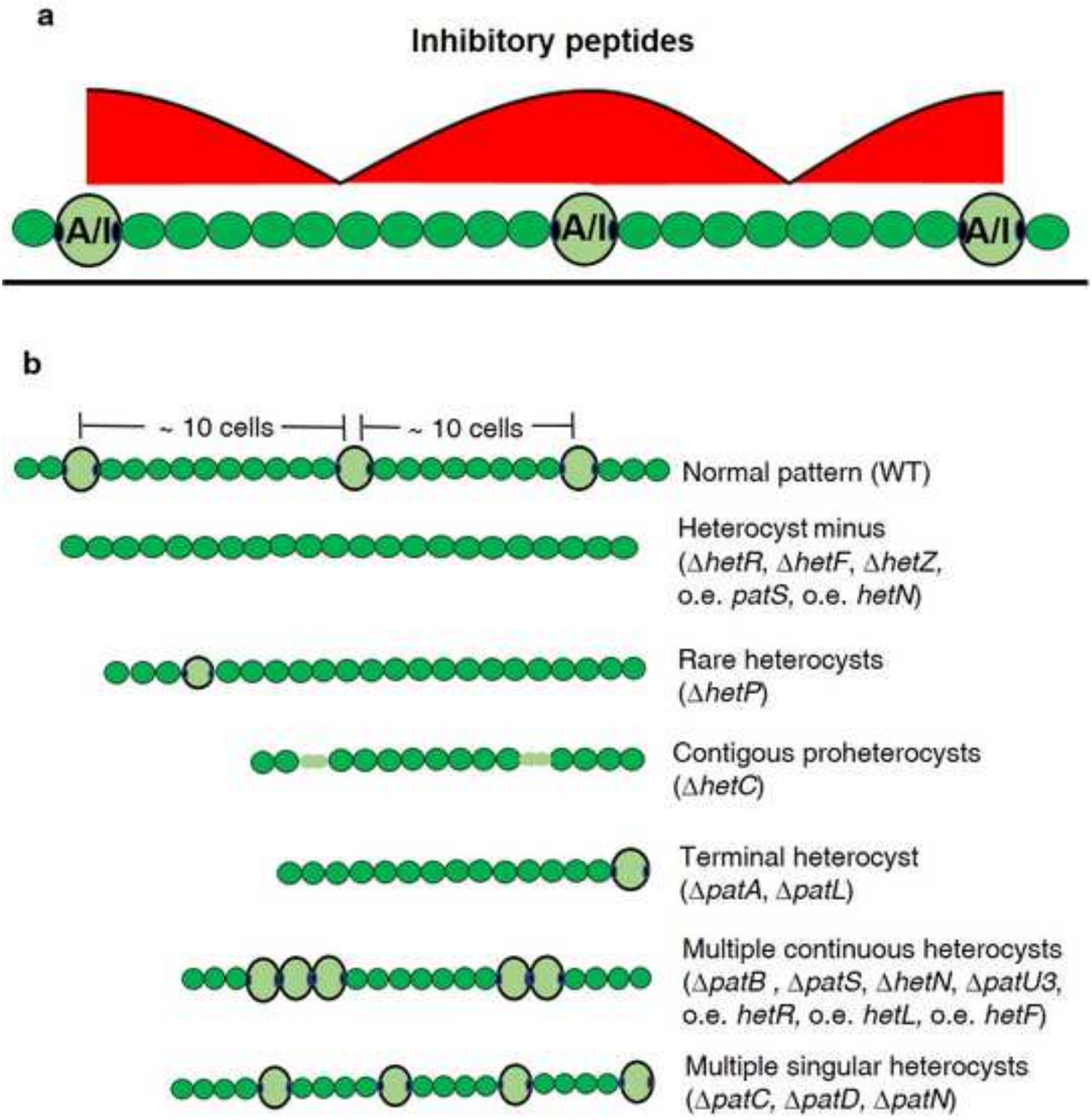
Heatmap graphical representation of the number of homologs within organisms, with their associated taxonomic levels (order and family) over early, commitment and patterning factors. Each line corresponds to one cyanobacterial strain. The color scale (0, 1 or >= 2) represents the number of homologs in each genome. Taxonomic name abbreviations are: c, Chroococcidiopsidales; p, Pleurocapsales; s, Spirulinales; g, Gloeoemargaritales

## Figure 8

**(a)** Boxplot showing the average number of homolog genes per genome within the cyanobacterial taxonomic orders. Blue dots are related to *Nostocales* order, with * or ** on the top when the statistical significance (one-sample *t*-test, *p*-value) was below 0.05 or 0.01, respectively.
**(b)** Integrative map of protein-protein interactions, transcriptional regulations and comparative genomics of heterocyst differentiation genes.
The size of the gray cycles is relative to the total number of biological links, including protein-protein interactions and transcription regulatory actions. Key player proteins (as defined by *in silico* analysis) are underlined. The transcriptional regulations include direct and indirect actions (see Table S4 for details). As described within the main text, HetN, PatS and PatS were not included in our analyses.

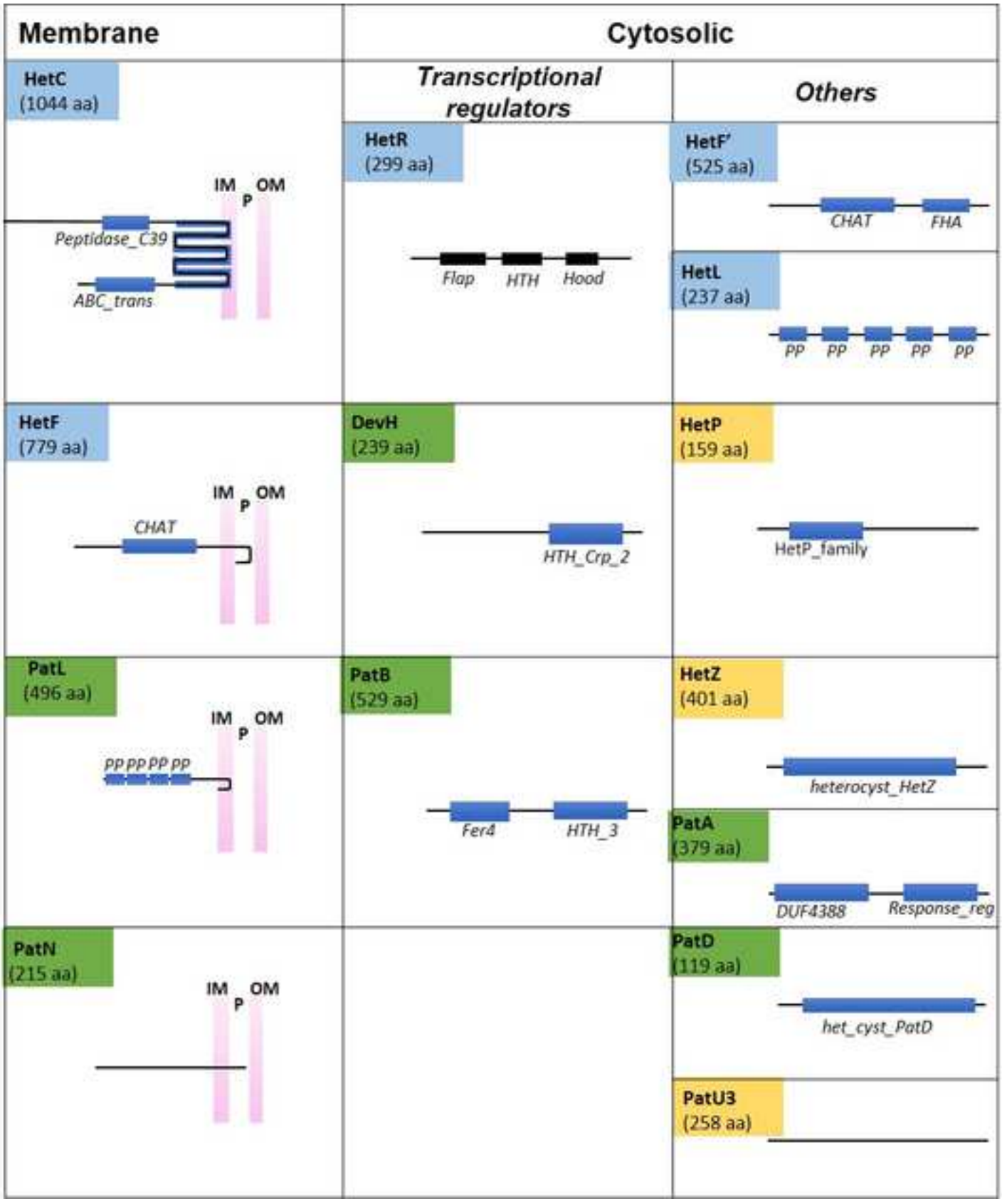
Figure





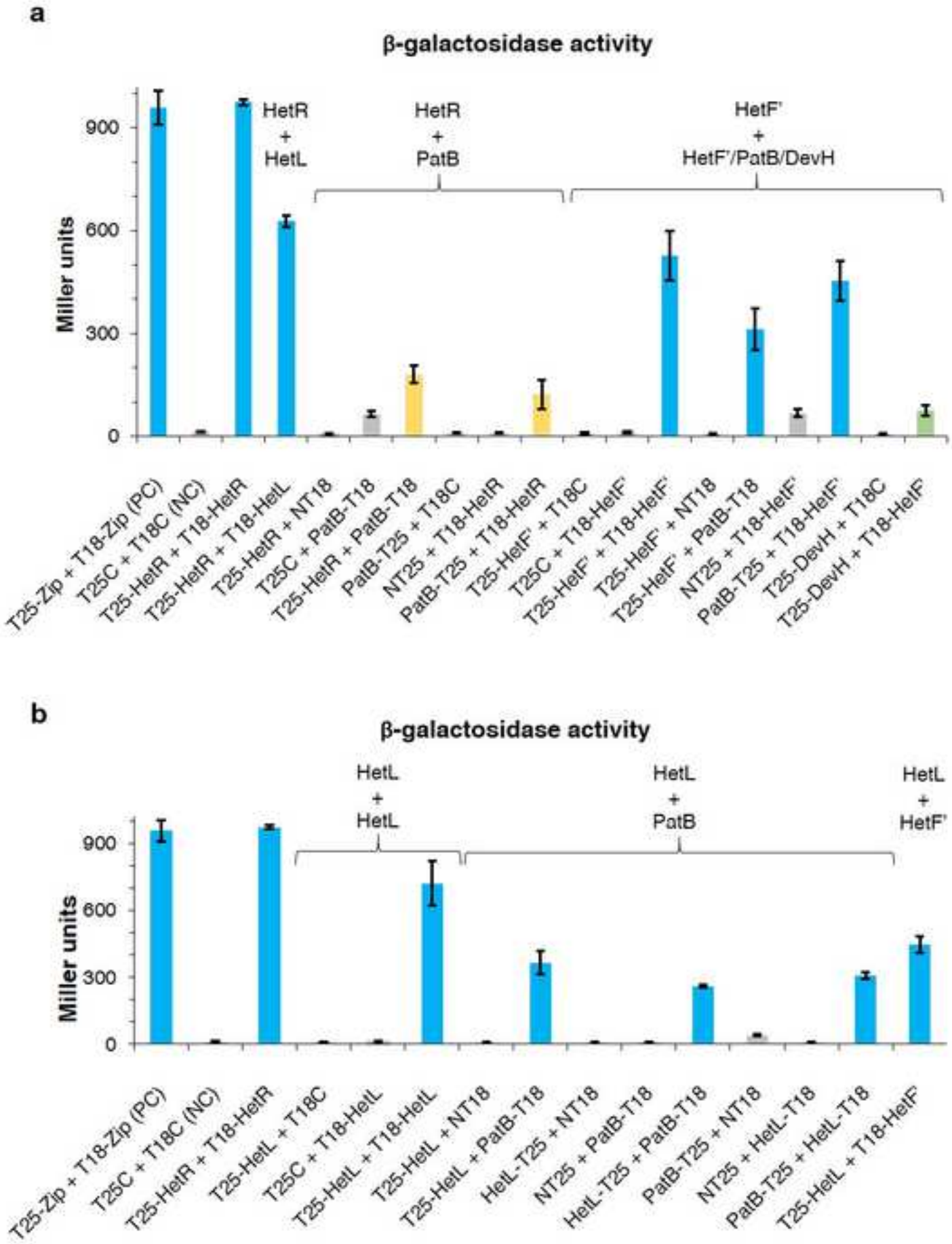
Figure

Click here to access/download;Figure;Figure 2.jpg



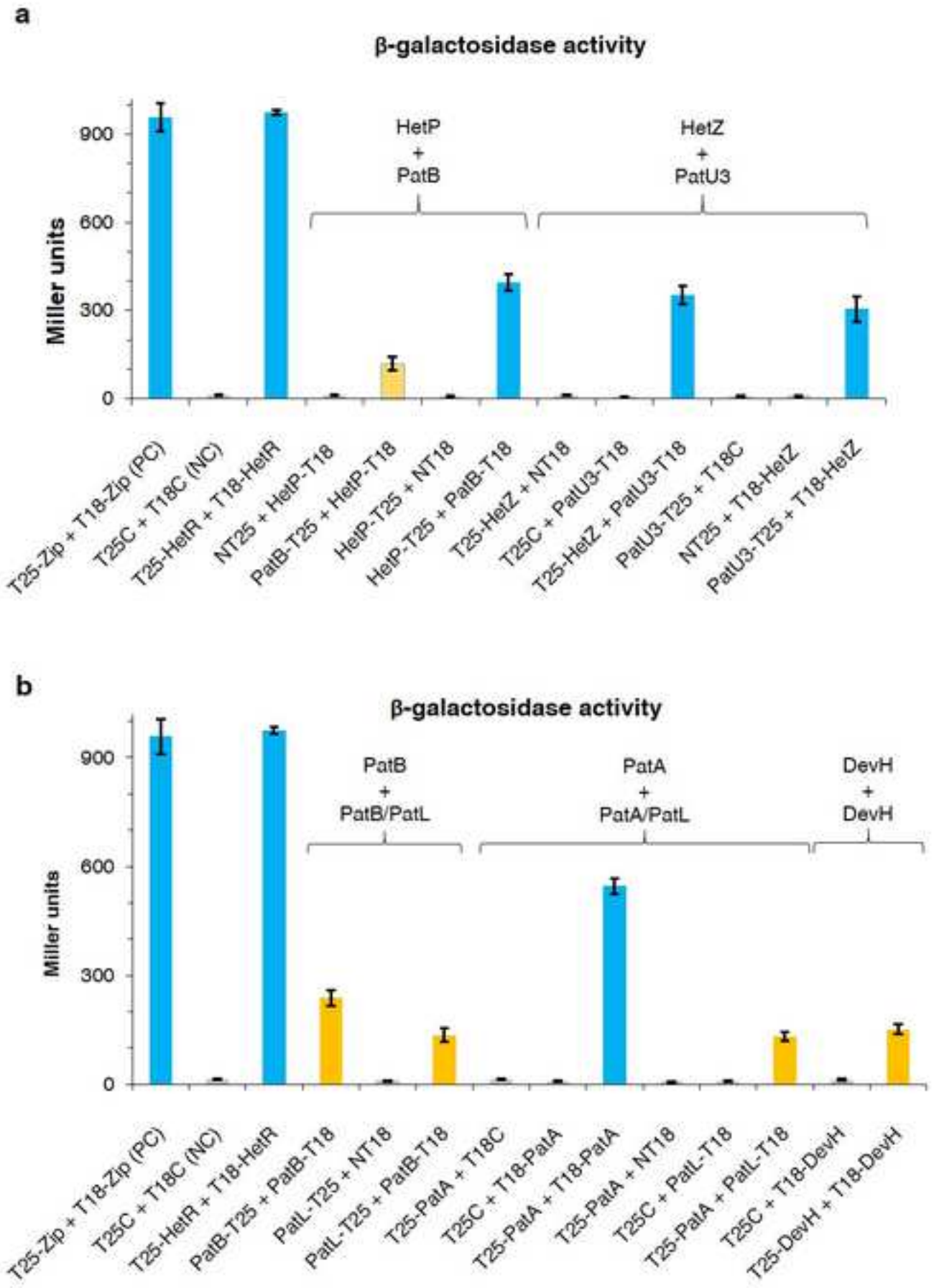
Figure





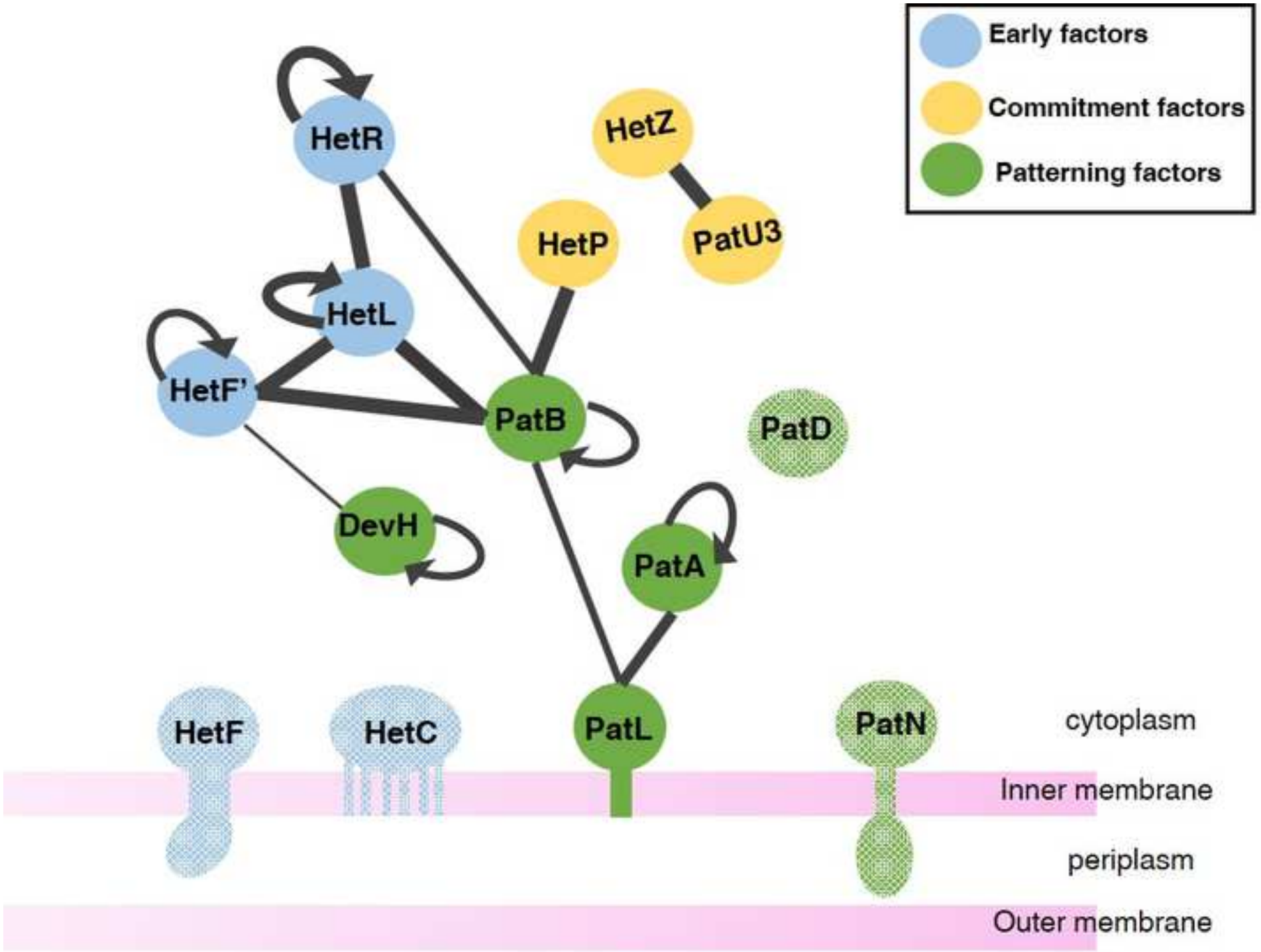
Figure

Click here to access/download;Figure;Figure 4.jpg



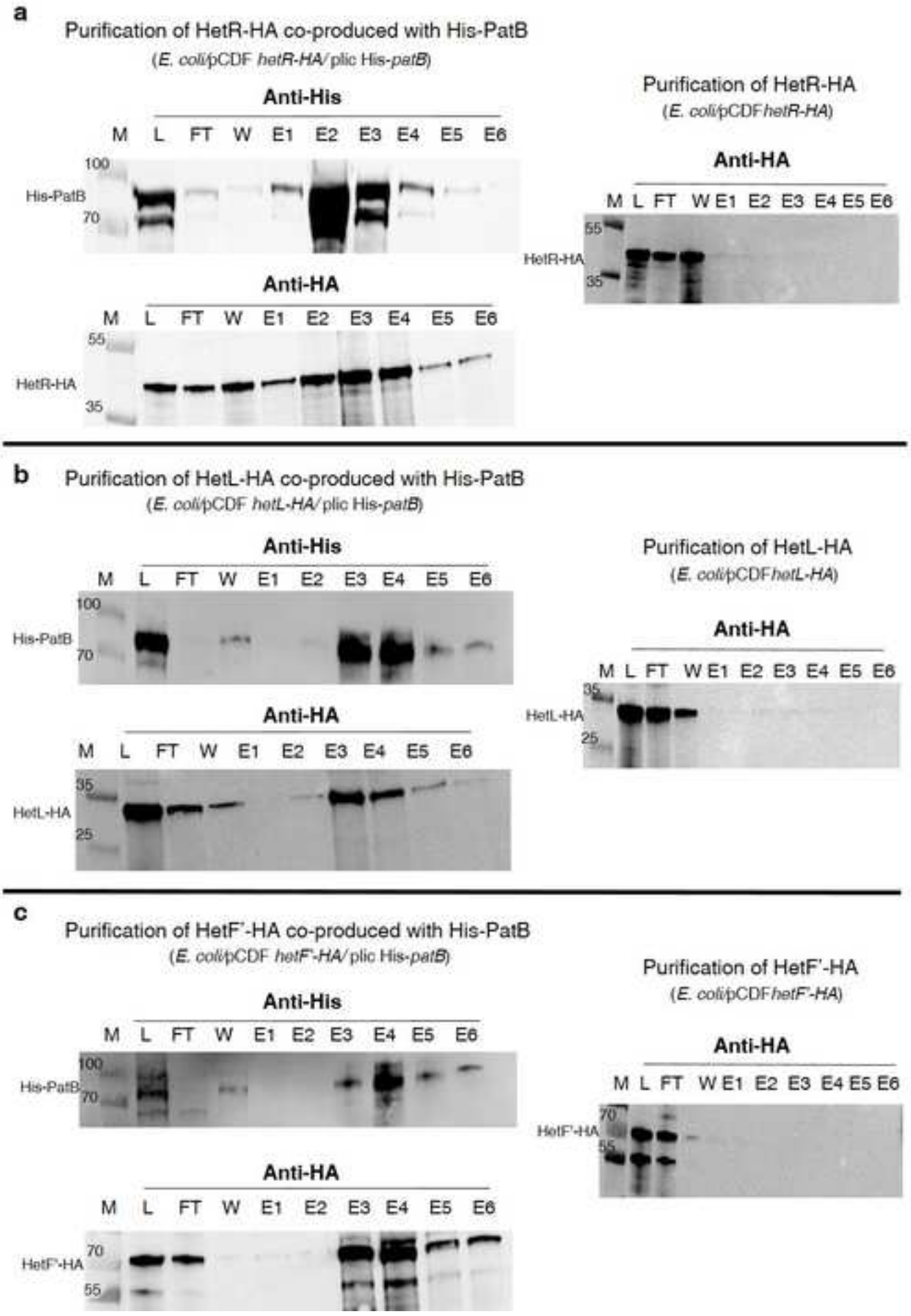
Figure

Click here to access/download;Figure;Figure 5.jpg



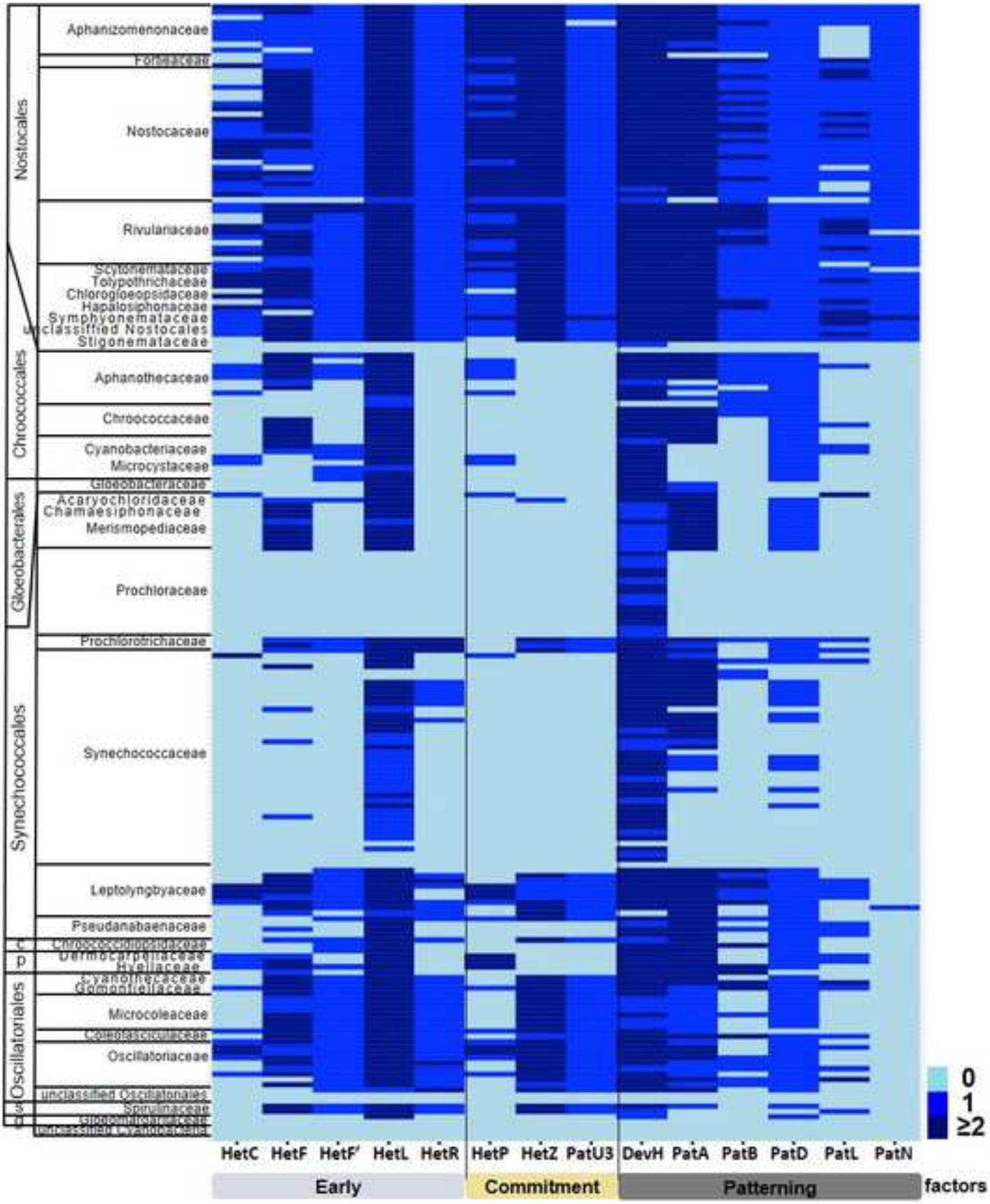
Figure

Click here to access/download;Figure;Figure 6.jpg



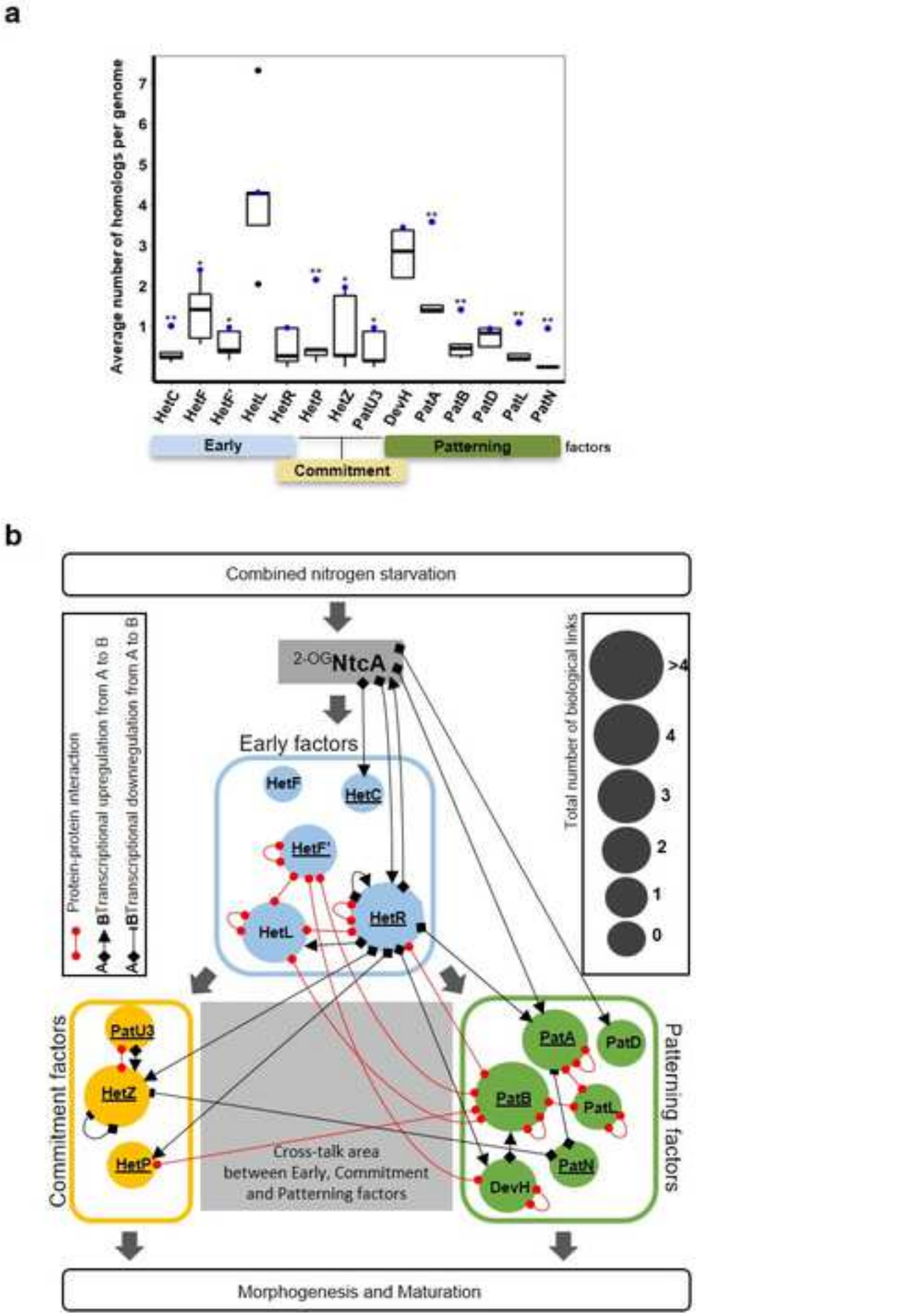
Figure

Click here to access/download;Figure;Figure 7.jpg



Figure

Click here to access/download;Figure;Figure 8.jpg



# Change of authorship request form (pre-acceptance)

## Please read the important information on page 4 before you begin

This form should be used by authors to request any change in authorship including changes in corresponding authors. Please fully complete all sections. Use black ink and block capitals and provide each author's full name with the given name first followed by the family name.

Please note: In author collaborations where there is formal agreement for representing the collaboration, it is sufficient for the representative or legal guarantor (usually the corresponding author) to complete and sign the Authorship Change Form on behalf of all authors.

**Section 1: Please provide the current title of manuscript**
(For journals: Please provide the manuscript ID, title and/or DOI if available.)
(For books: Please provide the title, ISBN and/or DOI if available.)

| | |
|---|---|
| **Manuscript ID no. in case of unpublished manuscript:** | |
| **DOI in case of published manuscript:** | MGAG-D-21-00741R1 |
| **ISBN (for books):** | |

Title: Interaction network among factors involved in heterocyst-patterning in cyanobacteria

**Section 2: Please provide the previous authorship, in the order shown on the manuscript before the changes were introduced. Please indicate the corresponding author by adding (CA) behind the name.**

| | First name(s) | Family name | ORCID or SCOPUS id, if available |
|---|---|---|---|
| 1st author | XIAOMEI | Xu | |
| 2nd author | MARYLINE | FOGLINO | |
| 3rd author | EMMANUEL | TALLA | |
| 4th author | AMEL | LATIFI | |
| 5th author | | | |
| 6th author | | | |
| 7th author | | | |

Please use an additional sheet if there are more than 7 authors.

# Change of authorship request form (pre-acceptance)

**Section 3: Please provide a justification for change. Please use this section to explain your reasons for changing the authorship of your manuscript, e.g. what necessitated the change in authorship? Please refer to the (journal) policy pages for more information about authorship. Please explain why omitted authors were not originally included and/or why authors were removed on the submitted manuscript.**

The change in the author-list consists in including Raphaël Rachedi as co-first author
The reason for this addition is thar R. Rachedi performed all the co-expression and copurification and immunoblotting experiments that were performed to answer the reviewers and editor requests (Figure 6 and Figure S1)

**Section 4: Proposed new authorship. Please provide your new authorship list in the order you would like it to appear on the manuscript. Please indicate the corresponding author by adding (CA) behind the name. If the corresponding author has changed, please indicate the reason under section 3.**

| | First name(s) | Family name (this name will appear in full on the final publication and will be searchable in various abstract and indexing databases) |
|---|---|---|
| 1st author | XIAOMEI | XU |
| ~~2nd~~ co-first author | RAPHAEL | RACHEDI |
| 3rd author | MARYLINE | FOGLINO |
| 4th author | EMMANUEL | TALLA |
| 5th author | AMEL | LATIFI |
| 6th author | | |
| 7th author | | |

Please use an additional sheet if there are more than 7 authors.

# Change of authorship request form (pre-acceptance)

**Section 5: Author contribution, Acknowledgement and Disclosures. Please use this section to provide a new disclosure statement and, if appropriate, acknowledge any contributors who have been removed as authors and ensure you state what contribution any new authors made (if applicable per the journal or book (series) policy). Please ensure these are updated in your manuscript - after approval of the change(s) - as our production department will not transfer the information in this form to your manuscript.**

**New acknowledgements:**

No author has been removed from author-list

**New Disclosures (financial and non-financial interests, funding):**

Not applicable

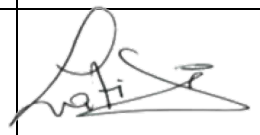
**New Author Contributions statement (if applicable per the journal policy):**

Not applicable

State 'Not applicable' if there are no new authors.

# Change of authorship request form (pre-acceptance)

**Section 6: Declaration of agreement. *All* authors, unchanged, new and removed *must* sign this declaration.**
(NB: Please print the form, sign and return a scanned copy. Please note that signatures that have been inserted as an image file are acceptable as long as it is handwritten. Typed names in the signature box are unacceptable.) *** Please delete as appropriate. Delete all of the bold if you were on the original authorship list and are remaining as an author.**

| | First name | Family name | | Signature | Affiliated institute | Date |
|---|---|---|---|---|---|---|
| 1st author | XIAOMEI | XU | I agree to the proposed new authorship shown in section 4 /~~**and the addition/removal*of my name to the authorship list.**~~ |  | Aix-Marseille University CNRS, IMM, Laboratoire de Chimie Bactérienne | 03-17th-2022 |
| 2nd author | RAPHAEL | RACHEDI | I agree to the proposed new authorship shown in section 4 /**and the addition~~/removal*~~of my name to the authorship list.** |  | Aix-Marseille University CNRS, IMM, Laboratoire de Chimie Bactérienne | 03-17th-2022 |
| 3rd author | MARYLINE | FOGLINO | I agree to the proposed new authorship shown in section 4 /~~**and the addition/removal*of my name to the authorship list.**~~ |  | Aix-Marseille University CNRS, IMM, Laboratoire de Chimie Bactérienne | 03-17th-2022 |
| 4th authors | EMMANUEL | TALLA | I agree to the proposed new authorship shown in section 4 /~~**and the addition/removal*of my name to the authorship list.**~~ |  | Aix-Marseille University CNRS, IMM, Laboratoire de Chimie Bactérienne | 03-17th-2022 |
| 5th author | AMEL | LATIFI | I agree to the proposed new authorship shown in section 4 /~~**and the addition/removal*of my name to the authorship list.**~~ |  | Aix-Marseille University CNRS, IMM, Laboratoire de Chimie Bactérienne | 03-17th-2022 |
| 6th author | | | I agree to the proposed new authorship shown in section 4 /**and the addition/removal*of my name to the authorship list.** | | | |
| 7th author | | | I agree to the proposed new authorship shown in section 4 /**and the addition/removal*of my name to the authorship list.** | | | |

Please use an additional sheet if there are more than 7 authors.

---



---

# Change of authorship request form (pre-acceptance)

## Important information. Please read.

- Please return this form, fully completed, to Springer Nature. We will consider the information you have provided to decide whether to approve the proposed change in authorship. We may choose to contact your institution for more information or undertake a further investigation, if appropriate, before making a final decision.

- By signing this declaration, all authors guarantee that the order of the authors are in accordance with their scientific contribution, if applicable as different conventions apply per discipline, and that only authors have been added who made a meaningful contribution to the work.

- Please note, we cannot investigate or mediate any authorship disputes. If you are unable to obtained agreement from all authors (including those who you wish to be removed) you must refer the matter to your institution(s) for investigation. Please inform us if you need to do this.

- If you are not able to return a fully completed form within **30 days** of the date that it was sent to the author requesting the change, we may have to withdraw your manuscript. We cannot publish manuscripts where authorship has not been agreed by all authors (including those who have been removed).

- Incomplete forms will be rejected.