Report Tile: Systems level approaches to understanding and manipulating heterocyst differentiation in *Nostoc punctiforme*: sites of hydrogenase and nitrogenase synthesis and activity

Final Scientific/Technical Report

Reporting period start date: 09/15/2008

Reporting period end date: 09/14/2012

Principle Author: John C. Meeks

Date Report issued: 04/02/2015

DOE Award Number: DE-FG02-08ER64693

Submitting Organization: University of California, OR/Sponsored Programs, 1850 Research Park Drive, Ste. 300, Davis, CA 95618-0000

Abstract

Heterocysts are specialized cells that establish a physiologically low oxygen concentration; they function as the sites of oxygen-sensitive nitrogen fixation and hydrogen metabolism in certain filamentous cyanobacteria. They are present at a frequency of less than 10% of the cells and singly in a nonrandom spacing pattern in the filaments. The extent of differential gene expression during heterocyst differentiation was defined by DNA microarray analysis in wild type and mutant cultures of *Nostoc punctiforme*. The results in wild-type cultures identified two groups of genes; approximately 440 that are unique to heterocyst formation and function, and 500 that respond positively and negatively to the transient stress of nitrogen starvation. Nitrogen fixation is initiated within 24 h after induction, but the cultures require another 24 h before growth is reinitiated. Microarray analyses were conducted on strains with altered expression of three genes that regulate the presence and spacing of heterocysts in the filaments; loss of function or over expression of these genes increases the heterocyst frequency 2 to 3 fold compared to the wild-type. Mutations in the genes *hetR* and *hetF* result in the inability to differentiate heterocysts, whereas over expression of each gene individually yields multiple contiguous heterocysts at sites in the filaments; they are positive regulatory elements. Mutation of the gene *patN* results in an increase in heterocysts frequency, but, in this case, the heterocysts are singly spaced in the filaments with a decrease in the number of vegetative cells in the interval between heterocysts; this is a negative regulatory element. However, over expression of *patN* resulted in the wild-type heterocyst frequency and spacing pattern. Microarray results indicated HetR and HetF influence the transcription of a common set of about 395 genes, as well as about 350 genes unique to each protein. HetR is known to be a transcriptional regulator and HetF is predicted to be a protease, perhaps operating thorough stability of HetR; thus, the influence of HetF on transcription of a unique set of genes was unanticipated. These two proteins are also found in non-heterocyst-forming filamentous cyanobacteria and the results have implications on their other physiological role(s). The PatN protein is unique to heterocyst-forming cyanobacteria. Cytological analysis indicated PatN is present in only one of the two daughter cells following division, but is present in both cell less than 8 h after division. Microarray analysis indicated only five genes were differentially transcribed in the *patN* mutant compared to the wild type; three up-regulated genes that are known to influence heterocyst differentiation and two down-regulated genes that have an unassigned function. Mutational analyses indicted the two down-regulated genes do not have a distinct role in heterocyst differentiation. Thus, PatN only indirectly impacts transcription. These databases provide lists of differentially transcribed genes involved in nitrogen starvation and cellular differentiation that can be mined for detailed genetic analysis of the regulation of heterocyst formation and function for subsequent photo-biohydrogen production.

Table of Contents

Executive Summary

Filamentous heterocyst-forming cyanobacteria have been suggested as the most applicable platform for photo-biohydrogen production (1). Both of nitrogenase and hydrogenase are sensitive to inactivation by oxygen. All cyanobacteria are characterized by the production of oxygen in the photosynthetic generation of ATP and reductant from water. Heterocysts are specialized cells, unable to produce oxygen, that develop from vegetative cells. Pilot plant studies have demonstrated that the typical 5 to 10% heterocyst frequencies of *Nostoc* and *Anabaena* spp. and genetic manipulations of nitrogenase or hydrogenase are insufficient to produce hydrogen on a commercial scale (1). When in symbiosis with a photosynthetic plant partner, the heterocyst frequency of *Nostoc* spp. is 30%, or more, of the cells, nitrogenase activity is elevated about fivefold and growth is slowed (2). We suggest duplication of the symbiotic morphological and physiology, apart from the plant partner, could make commercial hydrogen production by heterocyst-forming cyanobacteria cost effective.

The experimental organism, *Nostoc punctiforme*, is symbiotically competent with a range of terrestrial plants. Its genome sequence and annotation was completed in 2007 by the DOE Joint Genome Institute. We applied systems level analyses to gene transcription and protein accumulation during heterocyst differentiation by the wild-type and specific mutants to define the regulatory pathway of differentiation and identify steps that may be co-opted by the symbiotic pant partner.

A 24 h DNA microarray with 6 time points of wild-type cultures following combined N starvation revealed that 996 genes were statistically significantly differentially transcribed; 559 and 427 were up- and down-regulated, respectively. This is in contrast to 467 genes differentially transcribed during steady state N_2-dependent growth, compared to NH_3-dependent culture; 344 and 123 were up- and down-regulated. The difference between 24 h after induction of heterocyst differentiation and steady state growth with N_2 is largely due to changes in 4 clusters of genes: 328 genes up-regulated at up 24 h that are transcribed similar to NH_3 grown by N_2 steady state; 389 genes that are down-regulated at 24 h and return to NH_3 levels at N_2 steady state; 113 genes that are uniquely up-regulated at N_2 steady state, while being similar to NH_3-grown up to 24 h after induction; and 75 genes uniquely down-regulated at N_2 steady state, while also similar to NH_3 up to 24 h after induction. This comparison allowed identification of genes unique to heterocyst differentiation and N_2 dependent growth, compared to core metabolic genes required for steady state growth with NH_3 or N_2 as N sources; the latter of which are differentially transcribed during the shock and recovery of N starvation. The 24 h time course of differential gene expression during wild-type heterocyst differentiation yielded 6 temporal patterns with different degrees of change. There were 2 patterns of down-regulation, both of which occurred immediately after induction: 384 genes down about 2-fold and 36 genes down more than 4-fold. There were four pattern of up-regulation, with different times and degrees in change: 282 genes that were immediately, but marginally up-regulated (many gene products of unassigned function); 53 genes up immediately and strongly (>4-fold) (gene encoding enzymes for alternative N sources and regulatory function); 220 genes up between 2- and 4-fold after an approximately 3 h delay (enzymes for structural components of the heterocyst); and 61 genes up between 4- and 16-fold after a 6-12 h delay (nitrogenase and hydrogenase and unique heterocyst metabolism). The up-regulated transcription patterns are consistent with those known from forward genetic analyses (3)

Mutation of the gene *patN* resulted in a frequency (30%) and pattern of heterocyst spacing (ca. 3 vegetative cells in the intervals between heterocysts) similar to that of the symbiotic growth state. However, the process of heterocyst maturation is markedly delayed in the mutant, its steady state rate of N_2 fixation is no greater than the wild-type and the filaments tend to fragment. The patterns of differential gene expression in the *patN* mutant differed from the wild-type due primarily to the delay in heterocyst maturation. However, as cultures approached steady state N_2 growth (5 d) the pattern of gene expression was similar in wild-type and the *patN* mutant. The exceptions included increased expression of *hetZ*, *patU* and *patA*; *hetZ* is a positive regulator that may interact with the co-transcribed negative regulator *patU*, and *patA* is a positive regulator in which heterocysts are located only at the ends of the mutant filaments. These genes may interact in a late heterocyst differentiation regulatory circuit. The negative acting PatN protein is subject to biased inheritance whereby only one daughter cell contains PatN immediately after cell division. This biased inheritance frees the lacking cell to initiate heterocyst differentiation should it receive a signal of N limitation. PatN is the only regulatory protein identified to date that is unique to heterocyst-forming cyanobacteria (4).

The transcriptional activator HetR serves as the master positive regulator of heterocyst differentiation and the protease HetF modeled to modulate HetR concentration or activity. In such a model, one predicts the pattern of differential gene expression following N starvation to be similar in the two mutants. However, even under permissive growth conditions, gene expression was different in the two mutants. In steady state NH_3 cultures and compared to the wild-type strain, 210 and 58 genes were differentially transcribed in the *hetF* and *hetR* mutants, respectively. Genes encoding the heterocyst regulatory proteins PatA, HetZ and PatU were up-regulated in the NH_3-grown *hetF*, but not the *hetR* mutant. Interestingly, 80% of the time upon N-starvation, both mutants defaulted to differentiate into motile hormogonium filaments. Differential transcription was analyzed only in cultures that did not form hormogonium-like filaments; even in those filaments a few genes characteristic of hormogonia were up-regulated in both mutants. There were roughly equal numbers of differentially transcribed genes over the 24 h time course in common (365) or unique to the wild-type (440), *hetF* (341) and *hetR* (391) mutants. Up-regulated genes following N-starvation in common to both mutants and the wild-type were those encoding for the alternative N sources nitrate, urea transport, heterocyst regulatory protein NrrA, phycobilisome degrading NblA, cell envelope, transport proteins and unassigned proteins. Common down-regulated genes included those for core-metabolism, secondary metabolites and unassigned proteins. Up-regulated genes unique to *hetF*, encode PatU, HetZ, molybdenum processing MoaA and those to *hetR* included HetN, DevC and 2-component signal transduction proteins. In all cases, the majority of genes encode unassigned proteins. These results will lead to an understanding of the roles HetR and HetF may have in vegetative filaments and non-heterocyst-forming cyanobacteria (5).

A shotgun proteomics profile from duplicate samples was defined for steady state free-living N_2-grown cultures that yielded 1574 proteins. This profile was similar to a prior analysis of soluble proteins from an NH_3-grown culture. The majority of known proteins identified were associated with core metabolism, but heterocyst, nitrogenase and hydrogenase proteins were under-represented. Forty separate preparations of *N. punctiforme* colonies were isolated from tissue of the hornwort *Anthoceros punctatus*, yielding about 1.8 g fresh weight; the colonies are stored at -80 C, but were not processed for analysis. The proteomics analyses have not been published.

Report Details

A. Experimental Methods. Global transcription analysis was performed using DNA microarrays. The initial arrays were PCR products of internal gene fragments generated in-house by support of the NSF. The arrays were also printed in-house and contained representatives of 6,893 genes; absent were 351 genes encoding transposases and 200 hypothetical genes less than 20 bp in length. The PCR-based arrays were subjected to two-color hybridization; the three biological replicate samples were also dye swapped, yielding 6 data points for each array spot and since the arrays were printed in duplicated on each slide, 12 data points were used for analysis. The PCR-based array was replaced by a NimbleGen platform; the array spots were composed of 6 to 8, 20 bp oligonucleotides per ORF, synthesized on the substratum and representing 7323 ORFs, including pseudogenes and those encoding transposases. These slides were subjected to one color hybridization and were conducted with triplicate biological/technical replicates. The NimbleGen software package processed and compiled the replicate oligonucleoitides for each ORF. The raw data were normalized in the R-statistical platform, via the LIMMA GUI (6) for PCR-based arrays, or the R-oligo (7) and R-oneChanelGUI (8) for NimbleGen arrays. For time course analyses, the PCR-based array normalized M values ($\log_2$ experimental – $\log_2$ reference) were computed in R and Genesis (9) used to for *K*-means clustering. Time course NimbleGen-based array data were first analyzed with BATS (Bayesian user-friendly software for analyzing time series microarray experiments; 10) and then *K*-means clustered with Cluster 3.0 (11).

Shotgun proteomics were performed with cell lysates in high osmotic buffer to stabilize protein complexes, separated by standard definition into soluble and particulate (membrane and wall) fractions at 30,000 x g; the supernatent fraction was further separated at 150,000 x g into protein complexes (pellet; physobilisomes and carboxysomes) and soluble proteins (supernatant). Subsamples of about 200 µg from each fraction were separated on SDS PAGE, gel slices were eluted, digested with trypsin and the resulting peptides analyzed online by MudPit with a Michrom Paradigm multidimensional HPLC connected to a Thermo-Fisher linear iontrap ms/ms via a vacuum electrospray ionization source. The raw data were processed via Normalized Spectral Abundance Factor (12).

B. Results and Discussion. At the time of publication, the free-living N_2-grown transcriptome of *N. punctiforme* was the most robust database available (3). Subsequently, RNAseq analyses of N_2-grown *Anabaena* sp. strain PCC 7120 have been published by two different groups. The *N. punctiforme* microarray N_2-transcriptome provided two significant observations. First, the 6 point time course of gene expression, from induction by NH_3-deprivation up to the presence of structurally mature heterocysts at 24 h after induction, remains the most intensive database, with robust statistical analysis; RNAseq experiments include 2-3 time points and lack rigorous statistical analysis. Clustering of the time course date allowed for organization of genes with respect to when their transcription was enhanced. The time of induction for a number of heterocyst genes has been determined during forward genetic analysis and the previously reported induction times served as internal controls in the clustering of the DNA microarray data. Four clusters of up-regulated genes were identified; they differ by the time of enhanced expression after N-starvation and the degree (intensity) of differential expression. Genes encoding proteins for assimilation of alternative N-sources, such as nitrate assimilation and urea transport, *nblA* for degradation of phycobilisomes, *hep* for heterocyst polysaccharides, the signal transduction protein NrrA and the heterocyst master regulator HetR were amongst the first to be up-regulated. Also included in this cluster were 17 unassigned proteins and 5 export proteins

which are targets for more detailed analysis. The early expressed alternative N source genes imply *N. punctiforme* would default to combined N, if available, before committing to heterocyst differentiation. That is to say, nitrogen fixation is a last resort, which must be recognized in attempts to engineer higher H_2 production strains. A second cluster of up-regulated genes induced about 3 h after N-starvation included enhanced expression of the regulatory proteins HetF, PatU and DevH, transhydrogenase, 3 export proteins, and 77 proteins of unknown function. The third cluster of up-regulated genes were initiated between 6 and 12 h after starvation and included those encoding enzymes of heterocyst glycolipid synthesis, Hup and Nif proteins, oxidative pentose phosphate enzymes for reductant supply and 3 terminal cytochrome oxidases as part of the oxygen protection mechanism. These are the functional enzymatic components of heterocysts. There were 13 unassigned proteins in this cluster. This cluster also includes 4 co-located genes encoding a secondary modified peptide (putative toxin) and a multidrug exporter, which may invoke a potential signaling function for these metabolites. The forth cluster includes genes that were immediately statistically significantly up-regulated, but only about 1.5-fold, and did not markedly vary through the time course. This cluster includes genes of relevance to nitrogen metabolism such as TCA cycle proteins in generation of 2-oxoglutarate and glutamine synthetase for ammonium assimilation, as well as cell envelope, some signal transduction proteins, 13 transporters, including 5 exporters and 143 proteins of unknown function. The known proteins are also present in vegetative cells which may account for their apparent marginal enhancement in the population of cells of which heterocysts consist of about 10%. This database is available on the PI's web page. The second observation is that, although at 24 h after induction the heterocysts are functional in nitrogen fixation, cultures require an additional 24 h before they commence growth. This temporal requirement was verified by measurements of biomass increase over time, a simple experiment that had never previously been conducted in conjunction with the differentiation process. The lag in growth indicates that N-starvation leading to heterocyst differentiation is a significant stress to the cultures and has implications of the selective pressures during differentiation and subsequent engineering. Comparison of the transcriptome at 24 h after induction to the steady state N_2-grown transcriptome confirmed that differential gene expression was required after the 24 h time point before growth could commence. A substantial fraction of the genes encoding core metabolic proteins, required to make two cells from one, were down-regulated following N-starvation; transcription of these genes was not enhanced until after the 24 h time point. We speculate that global transcription factors may be involved in suppressing transcription of core metabolic genes until heterocyst structural and functional genes had been expressed, translated and proteins complexes assembled.

The *patN* mutant is the only representative of a genetic alteration leading to an increase in heterocyst frequency, where the heterocysts are singly present at sites in the filament, with a decreased interval of vegetative cells between adjacent heterocysts (called multiple singular heterocysts or Msh) (4). In all other pattern mutants with increased heterocyst frequency, the heterocysts are clustered as pairs, triplets, or, more rarely, quadruplets at sites in the filaments (called multiple contiguous heterocysts or Mch); these mutations include deletion of the negative regulators PatS or PatU, or over expression of the positive regulators HetR or HetF. The DNA microarray results indicated the heterocyst regulatory genes *hetZ*, *patU* and *patA* were up-regulated in the PatN deletion mutant, implying repression in the presence of PatN. The heterocyst frequency and spacing pattern in filaments of the PatN mutant is similar to those observed in the symbiotic growth state of *N. punctiforme* in association with the hornwort *A.*

punctatus. The *patN* phenotype indicates it is a negative regulator of heterocyst differentiation. Since PatN is a membrane protein with no discernible DNA binding motif, we assume the transcriptional repression of *hetZ*, *patU* and *patA* is indirect. Cytological analyses with a fluorescent marker determined the protein is present in vegetative cells and undergoes biased inheritance following cell division. Based on PatN-GFP localization, vegetative cells appear to grow in a polar fashion; i.e. the one growing end of the cell lacked membrane localized PatN, but PatN's influence can be exerted throughout the growing cell prior to cell division. Following division, the "younger" daughter cell now lacked detectable PatN, but PatN appeared at the new septum in the lacking daughter cell within less than 8 h after division. The daughter cell lacking PatN can respond to the signal of N-limitation and initiate differentiation. The results of our initial analysis of PatN function supports a two-stage model of the initiation of heterocyst differentiation following N-starvation that we proposed in 2002 (13). We suggested in the first stage that all cells sense the signal of N limitation, but only some cells are competent to respond to the signal; we speculated that competence was dependent on a stage in the cell cycle. The cell division cycle determines which daughter cell lacks PatN; thus, cells singly or in a cluster (depending on the asynchrony of division in the population) could initiate differentiation. The second stage involves resolution of either a cluster or single, but closely spaced, initiating cells. Resolution is dependent on the interactions between the master regulator HetR and its diffusible antagonist PatS. Lack of PatS results in dominance of HetR activity, as does over expression of HetR, both of which result in Mch. Exactly how PatU and HetF influence the activities of PatS and HetR, respectively, to yield similar Mch phenotypes is unresolved. Rather than an interesting oddity in a group of bacteria, this model now places study of heterocyst differentiation in the mainstream of developmental biology of higher organisms. Differential inheritance of cell fate determinants (e.g. PatN) (14) and autoregulatory activator-inhibitor systems (e.g. HetR-PatS) (15) are two broad strategies in the developmental biology of complex systems ranging from hydra to vertebrates.

HetR and HetF proteins are found in heterocyst-forming and non-heterocyst-forming filamentous cyanobacteria. Their roles in vegetative growth are unknown; loss of function mutation does not impair permissive growth in *N. punctiforme* or *Anabaena* 7120. However, the absence of either HetR or HetF leads to not only an inability to differentiate heterocysts, but also a propensity to hormogonium differentiation following N-starvation in *N. punctiforme*. This latter property contributes, in part, to our developing concept of decision-making in cyanobacterial development that involves common and branched regulatory pathways. HetR is now known to function as a transcriptional regulator; thus, with respect to the similar loss of function and over expression HetR and HetF mutant phenotypes, HetF is thought to function through HetR. We tested this hypothesis by microarray analysis, comparing transcription patterns in steady state NH_3-grown cultures and over a 24 h period following N-starvation. The absence of HetR and/or HetF resulted in significant alteration in expression of 58 and 210 genes, respectively, in NH_3-grown cultures, compared to the wild-type, and 71 and 74% of the genes, respectively, were up-regulated (derepressed). Four genes associated with heterocyst differentiation (*patA*, *patU*, *hetZ* and *hepN*) were derepressed in the HetF mutant, whereas in the HetR mutant *patA* and *patU* were repressed and *hetZ* and *hepN* unchanged. The similarity of transcriptional expression of *patA*, *patU* and *hetZ* in the PatN and HetF mutants may have implications on the regulatory pathway of PatN. Several genes encoding proteins characteristic of hormogonia function (gas vesicles, taxis and pili) were up-regulated in the HetF and HetR mutants. We interpret these results to indicate the NH_3-grown mutant cultures were poised to differentiate into hormogonia

even before the N-stress is imposed, and this is the reason why, in 80% of the N-starvation experiments with these mutants, nearly 100% of the filaments differentiated into hormogonia. The cause and nature of the poising are unknown. In time course N-starvation experiments, only those minority 20% of the samples that did not differentiate into hormogonium-like filaments were processed. A large number of genes were differentially transcribed uniquely and in common with the wild-type, HetR and HetF mutants: There were 365 genes in common in the three strains and 395 in common between the HetF and HeR mutants, as well as 440, 341 and 391 unique differentially transcribed genes in the wild-type, HetF and HetR mutants, in the order stated. The vast majority of the 365 common genes displayed the same patterns of expression (up- or down-regulated, relative to NH_3-grown cultures). These included genes encoding proteins for assimilation of alternative N-sources; this result indicated the N-starvation response was induced in all strains. However, 29 genes showed different patterns and many of these encode either unassigned proteins or proteins for synthesis of secondary metabolites; of particular interest is *patA* which showed no differential transcription in the wild-type, was up-regulated in the HetF mutant and down-regulated in the HetR mutant. Genes unique to the wild-type included all of those encoding heterocyst structural (polysaccharide and glycolipid synthesis) and functional (Nif, Hup, cytochrome oxidase, carbon catabolism) proteins, as well as 35 signal transduction and transcriptional regulation proteins. The 395 genes with similar transcription patterns in HetR and HetF mutants were predominantly (65%) up-regulated and the majority (65%) are unassigned proteins. Up-regulated genes with known function encode proteins for hormogonia function that were enhanced from their previously high transcription level in permissive conditions, and a variety of core metabolism and transport proteins. The distribution of common down-regulated genes, encoding proteins of known function, was also grouped in a variety of metabolic roles. These results imply HetR and HetF have at least an indirect effect on transcription under conditions of N-stress. Most of the 341 unique genes differentially transcribed genes in the HetF mutant were altered between 1.5 to 2.0 fold up or down and they encode proteins with various metabolic functions. However, proteins with greater than 4-fold up-regulation included PatU, cell division factor FtsQ, GTP synthetase , and molybdenum cofactor MoaA; an unassigned protein containing a TRP domain was down-regulated 8-fold. In contrast to the HetF mutant, many of the up-regulated genes in the HetR mutant were enhanced between 2 to 32-fold, relative to the permissive growth state. Examples include approximately 45 signal transduction proteins, 17 cell envelope polysaccharide synthetic factors, as well as heterocyst related HetN and DevC, and hormogonium repressor HrmI, but not the remainder of the Hrm operon. Genes distinctly down-regulated 2 to 4-fold encode porphyrin subunits, transport and carbon storage proteins. We conclude that there is large scale disruption of wild-type transcription patterns in both HetF and HetR mutants, and those changes indicate both mutual and distinct pathways of influence during steady state growth and N-stress that are peripheral to heterocyst differentiation and nitrogen fixation (5).

We observed a defect in our experiments to define the steady state N_2-grown proteome. In two replicates, the proteome was deficient in proteins involved in heterocyst function, including nitrogenase and hydrogenase enzymes. We were unable to determine the causes; the processed cultures contained heterocysts detectable by microscopy. Otherwise, we collected a reasonably robust catalog of 1,574 proteins, although the PI is unsure of the biological value of the catalog. We did not commit to processing of the symbiotic *N*, *punctiforme* colonies that were collected over a period of one-year and stored at -80 C.

References.

1. Sakurai H, Masukawa H, Kitashima M, Inoue K. 2015. How close we are to achieving commercially viable large-scale photobiological hydrogen production by Cyanobacteria: A review of the biological aspects. Life 5:997-1018.

2. Meeks, JC. 1998. Symbiosis between nitrogen-fixing cyanobacteria and plants. BioScience 48:266-276.

3. Christman HD, Campbell EL, Meeks JC. 2011. Global transcription profiles of the nitrogen stress response resulting in heterocyst or hormogonium development in *Nostoc punctiforme*. J. Bacteriol. 193:6874-6886.

4. Risser DD, Wong FCY, Meeks JC. 2012. Biased inheritance of the protein PatN frees vegetative cells to initiate patterned heterocyst differentiation. Proc. Natl. Acad. Sci. USA 109:15342-15347.

5. Christman HD, Campbell EL, Meeks JC. 2015. The proteins HetR and HetF, required for heterocyst differentiation, influence transcription of both a common set and unique sets of genes in *Nostoc punctiforme*. In preparation.

6. Smyth GK. 2005. Limma: linear models for microarray data. In: R Gentleman, V Carry, S Dudoit, R Irizarry and W Huber (eds), Bioinformatics and computational biology solutions using R and Bioconductor, p. 397-420. Springer, New York.

7. Carvalho B, Bengtsson H, Speed TP, Irizarry RA. 2007. Exploration, normalization, and genotype calls of high-density oligonucleotide SNP array data. Biostatistics 8:485-499.

8. Sanges R, Cordero F, Calogero RA. @007. OneChannelGUI: A graphical interface to Bioconductor tools, designed for life scientists who are not familiar with R language. Bioinformatics 23:3406-3408.

9. Sturn A, Quackenbush J, Trajanoski Z. 2002. Genesis: Cluster analysis of microarray data. Bioinformatics 18:207-208.

10. Angelini C, Cutillo L, De Canditi D, Mutarelli M, Pensky M. 2008. BATS: a Bayesian user-friendly software for analyzing time series microarray experiments. BMC Boinformatics 9:415.

11. de Hoon MJL, Imoto S, Nolan J, Wood P. 2004. Open source clustering software. Bioinformatics 20:1453-1454.

12. Swanson SK, Pavelka N, Florens L, Parmely, T, Conway JW, Conway RC, Washburn MP. 2007. Evaluation of 36 MudPIT replicate datasets using normalized spectral counting and PLGEM. J. Biomol. Tech. 18:64-65.

13. 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.

14. Meinhardt H. 2008. Models of biological pattern formation: From elementary steps to the organization of embryonic axes. Curr. Top. Dev. Biol. 81:1-63.

15.Macara IG, Mili S. 2008. Polarity and differential inheritance – universal attributes of life? Cell 135:801-812.

Bibliography

Christman HD, Campbell EL, Meeks JC. 2011. Global transcription profiles of the nitrogen stress response resulting in heterocyst or hormogonium development in *Nostoc punctiforme*. Journal of Bacteriology 193:6874-6886. doi:10.1128/JB.0599-11.

Risser DD, Wong FCY, Meeks JC. 2012. Biased inheritance of the protein PatN frees vegetative cells to initiate patterned heterocyst differentiation. Proceedings of the National Academy of Sciences USA 109: 15342-15347. doi: 10.1073/pnas.1207530109