Genome-wide Expression Analysis of the Responses to Nitrogen Deprivation in the Heterocyst-forming Cyanobacterium Anabaena sp. Strain PCC 7120

Shigeki Ehira,1,2 Masayuki Ohmori,2 and Naoki Sato1,∗

Department of Molecular Biology, Faculty of Science, Saitama University, 255 Shimo-Ohkubo, Sakura-ku, Saitama City, Saitama 338-8570, Japan1 and Department of Life Sciences, Graduate School of Arts and Sciences, The University of Tokyo, Komaba 3-8-1, Meguro, Tokyo 153-8902, Japan2

(Received 17 April 2003; revised 20 May 2003)

Abstract

A heterocyst is a terminally differentiated cell of cyanobacteria which is specialized in dinitrogen fixation. Heterocyst differentiation in Anabaena sp. strain PCC 7120 is triggered by deprivation of combined nitrogen in the medium. Although various genes that are upregulated during heterocyst differentiation have been reported, most studies to date were limited to individual or a small number of genes. We prepared microarrays in collaboration with other members of the Anabaena Genome Project. Here we report on the genome-wide expression analysis of the responses to nitrogen deprivation in Anabaena. Many unidentified genes, as well as previously known genes, were found to be upregulated by nitrogen deprivation at various time points. Three main profiles of gene expression were found: genes expressed transiently at an early stage (1–3 hr) of nitrogen deprivation, genes expressed transiently at a later stage (8 hr), and genes expressed when heterocysts are formed (24 hr). We also noted that many of the upregulated genes were physically clustered to form ‘expressed islands’ on the chromosome. Namely, large, continuous genomic regions containing many upregulated genes were upregulated in a coordinated manner. This suggests a mechanism of global regulation of gene expression that involves chromosomal structure, which is reminiscent of eukaryotic chromatin remodelling. The possible implications of this global regulation are discussed.

Key words: Anabaena sp. PCC 7120; DNA microarray; nitrogen deprivation; heterocyst differentiation

1. Introduction

Differentiation of heterocysts in cyanobacteria provides a unique system of studying prokaryotic cell differentiation as well as a vast resource of biologically utilisable nitrogen to the global ecosystem. The function of the heterocyst is the fixation of dinitrogen, and cyanobacteria are one of the major nitrogen fixers in the biosphere. Heterocystous cyanobacteria such as Anabaena sp. strain PCC 7120 (hereafter referred to as Anabaena), also called Nostoc sp. strain PCC 7120, grow as linear multicellular filaments. Upon limitation of combined nitrogen in the medium, heterocysts are formed with a regular spacing of 10–15 cells.1–3 Although nitrogen fixation is also performed by various bacteria as well as non-heterocyst-forming cyanobacteria, the heterocysts provide a unique system of studying induction of nitrogen fixation by nutritional limitation. This induction is not just induction of enzymes, but involves a complex series of coordinated cellular processes leading to the differentiation of heterocysts.

A cascade-like activation of genes related to cell differentiation has been well studied in the sporulation of Bacillus subtilis, in which sequential activation of different sigma subunits of RNA polymerase temporally and spatially regulates gene expression.4 In Anabaena, the genes encoding one principal sigma factor (sigA) and seven group 2 sigma factors (sigB, C, D, E, F, G, and H) have been identified.5–7 Five of the seven group 2 sigma factors have been inactivated, but there is no evidence that they contribute to heterocyst differentiation. The mechanism of regulation of gene expression during heterocyst differentiation appears to differ from that during sporulation in B. subtilis.

Detailed analysis in the process of heterocyst differentiation dates back half a century.8 The program of protein synthesis during heterocyst differentiation and transcripts expressed in heterocysts have been reported,9,10 but individual genes involved in this process remain to

Communicated by Satoshi Tabata
∗ To whom correspondence should be addressed. Tel. +81-48-858-3623, Fax. +81-48-858-3384, E-mail: naokisat@molbiol.saitama-u.ac.jp
be identified. Several techniques including transposon tagging and chemical mutagenesis have been used to identify the genes involved in or regulating the differentiation process.\textsuperscript{11,12} One such gene, \textit{hetR}, is essential for heterocyst differentiation. Extra copies of \textit{hetR} result in ectopic differentiation even when the supply of combined nitrogen is not limited, and multiple contiguous heterocysts are formed on nitrogen-deficient media.\textsuperscript{13} The level of \textit{hetR} transcripts increases within 30 min after nitrogen deprivation.\textsuperscript{14} Induction of \textit{hetR} is dependent on the global nitrogen regulator NtcA.\textsuperscript{15} NtcA is known to regulate the expression of genes related to the function and structure of heterocysts such as \textit{hetC}, \textit{petH}, and \textit{devBCA}.\textsuperscript{16–18} and is essential for heterocyst differentiation.\textsuperscript{19,20} Thus, a number of genes are expressed in an ordered sequence such that expression at one stage depends on the genes expressed at a previous stage during heterocyst differentiation.\textsuperscript{20}

As the complete genomic sequences of many organisms have been determined, DNA microarrays have been used to monitor the expression of thousands of genes simultaneously.\textsuperscript{21} In the cyanobacterium \textit{Synechocystis} sp. PCC 6803, DNA microarray analysis has been carried out to unravel responses to high light,\textsuperscript{22} low temperature,\textsuperscript{23} light-to-dark transition,\textsuperscript{24} salt and osmotic stress,\textsuperscript{25} and to survey the target genes for SYCRP1, a cyanobacterial cAMP receptor protein.\textsuperscript{26} Response to change in nitrogen nutrition has also been analyzed in plants.\textsuperscript{27,28} Gene expression patterns during cell differentiation, such as shoot development in \textit{Arabidopsis thaliana}\textsuperscript{29} and sporulation in yeast,\textsuperscript{30,31} have been reported. The complete genomic sequence of \textit{Anabaena} has been determined\textsuperscript{32} and a research group including the present authors has prepared a custom-made microarray (N. Sato et al. under submission).

To unravel the temporal and spatial changes in gene expression during heterocyst differentiation, we carried out DNA microarray analysis. These analyses allowed us to identify previously uncharacterized genes that are upregulated by nitrogen deprivation and upregulated in heterocysts. In addition, we found that large, contiguous genomic regions were regulated simultaneously during heterocyst differentiation. This suggests a mechanism of global regulation of gene expression that involves chromosomal structure. The possible implications of this global regulation are discussed.

2. Materials and Methods

2.1. Strain and growth conditions

\textit{Anabaena} sp. strain PCC 7120 was grown at 32°C in the light (50 µE m\textsuperscript{−2} s\textsuperscript{−1}) in the BG-11 medium,\textsuperscript{33} buffered with 5 mM N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid (HEPES)-NaOH (pH 7.5). Liquid culture was bubbled with air containing 1.0% (v/v) CO\textsubscript{2}.

For time course studies, a 0.6-liter culture in the mid-log phase grown in the BG-11 medium (containing 17.6 mM NaNO\textsubscript{3} as a nitrogen source) was supplemented with 1 mM ammonium chloride then grown for 12 hr. Then the cells in a 100-ml culture were collected by centrifugation (zero time). In each cell harvest, cells in the culture medium were mixed with an equal volume of ice-cold ethanol containing 5% phenol to kill them rapidly and to eliminate additional effects during centrifugation.\textsuperscript{34} Then, the cells were collected by centrifugation at 4°C, rapidly frozen in liquid nitrogen, and stored at −80°C until isolation of RNA. The remaining culture was divided into two portions, and the cells were collected by centrifugation at room temperature. The cells in each 250 ml portion were washed with the BG-11 medium (BG-11 medium lacking nitrate), and resuspended in the original volume of BG-11 or BG-11\textsubscript{0} medium. For each time point, cells in a 100-ml culture were collected by centrifugation as described above.

2.2. Preparation of heterocyst-enriched fraction

The heterocyst-enriched fraction was obtained from heterocyst-induced cell filaments essentially as described previously.\textsuperscript{35} Filaments in the late-log phase were induced to heterocyst differentiation in BG-11\textsubscript{0} medium for 48 hr. One-fifth of the induced filaments were set aside for the isolation of RNA from the whole filaments. The remaining cells were used to prepare a heterocyst-enriched fraction. Vegetative cells were broken through a French pressure cell. Microscopic examination indicated that about 80% of the cells in the heterocyst-enriched fraction were heterocysts.

2.3. Preparation of fluorescent probes

RNA was extracted from the harvested cells as described previously\textsuperscript{34} and purified by ultracentrifugation. Single-stranded cDNA was prepared in the presence of Cy3-dUTP or Cy5-dUTP (Amersham Biosciences, NJ, USA). At first, a mixture of RNA and random primer (Invitrogen, CA, USA) was heated to 70°C for 5 min, and then rapidly chilled on ice. The reverse transcription was performed in a 20-µl volume containing 20 µg RNA, 50 µM random primer, 10 mM dithiothreitol (DTT), 500 µM each of dATP, dCTP, and dGTP, 200 µM dTTP, 100 µM Cy3-dUTP or Cy5-dUTP, 25 U of RNAguard (Amersham Biosciences), and 200 U of SuperScript II (Invitrogen) in 1 x first-strand buffer (50 mM Tris-HCl, 75 mM KCl, 3 mM MgCl\textsubscript{2}, pH 8.3). The mixture was incubated at 42°C for 1 hr. Then, 200 U of SuperScript II was added and the mixture was incubated for another hour. The enzyme was inactivated at 70°C for 10 min. The reaction products of two reactions (one with Cy3 labeling and the other with Cy5 labeling) were combined. Unincorporated fluorescent nucleotides were removed by QIAquick PCR purification kit (Qiagen,
2.4. DNA microarray analysis

The fluorescent cDNA solution was applied to a glass microarray slide and the slide was covered with a cover slip. The glass slide was placed in an ArrayIt hybridization cassette (TeleChem International, CA, USA) which was subsequently incubated for 12 to 15 hr in a water bath at 65°C. After hybridization, the glass slide was washed three times in 2× SSC, 0.2% SDS at room temperature for 10 min each, three times in 0.2× SSC, 0.2% SDS at room temperature for 10 min each, and three times in 0.2× SSC, 0.2% SDS at 65°C for 10 min each. To eliminate SDS, the glass slide was washed two times in 0.2× SSC at room temperature for 5 min each, and finally dried by gentle centrifugation. The glass slide was scanned with a scanning laser microscope (ScanArray 4000; GSI Lumonics, Ottawa, Canada) at a resolution of 10 µm per pixel. Data analysis was performed using the ArrayVision software (ver. 6.0; Imaging Research, Ont, Canada).

The relative ratio for each DNA segment was defined as normalized fluorescence intensity for each DNA segment at each time point relative to that at the zero time. The relative ratio in the base-2 logarithm was assessed to be significantly different from zero by applying t-test to a significance level of 1%. In the evaluation of significance level by the t-test, the average standard deviation of relative ratio for all DNA segments at each time point was used in place of the actual standard deviation when the actual one was smaller than the average one. Any DNA segment for which the relative ratio was significantly different from zero was judged to be either upregulated or downregulated.

2.5. RNA gel blot analysis

RNA gel blot analysis was performed essentially as described previously with digoxigenin-labeled DNA probes. Each DNA probe was prepared by PCR with a set of specific primers, which were designed based on the nucleotide sequence in the CyanoBase database (http://www.kazusa.or.jp/cyano/).

3. Results

3.1. Construction of DNA microarray for the genome of Anabaena sp. PCC 7120

To construct a microarray that covers the whole genomic sequence, we selected 2407 clones from the sequencing clones that had been used for the assembly of genomic sequences at the Kazusa DNA Research Institute. Each of the selected clones bore an insert of approximately 3 kbp, containing one to eight open reading frames (ORFs) of various sizes. The selection was made in such a way that each clone overlapped another clone on each end. To ensure reproducibility of the microarray analysis, amplified DNA fragments were arrayed in duplicate on each glass slide. The microarray used in the present study covered about 90% of the whole chromosomal sequence. We used only the sequences of the main chromosome, because of the limitation in the arrayer that we originally planned to use.

A typical result indicating the reliability of the microarray analysis is shown in Figure 1A. Cy3- and Cy5-labeled cDNAs were prepared from a single preparation of total RNA that had been isolated from the cells grown in the standard condition, and then they were mixed and hybridized to a microarray. Each data point represents the average of fluorescence intensities for each DNA segment arrayed in duplicate. The background was estimated from the average intensity over a rectangular outline with a 2-pixel width surrounding each circular spot, and was subtracted from the intensity for each DNA segment. The normalization of fluorescence intensity was made in such a way that the sum of the corrected intensities of Cy3- and Cy5-labeled cDNAs for all DNA segments were equalized, and the average of intensity was defined as unity. The normalized fluorescence intensity for each DNA segment was greater than 0.04, while the average fluorescence intensity of the background was 0.009. The data points aligned on the diagonal with low deviation and confined between the two reference lines, which corresponded to a twofold difference. A few data points were outside the two reference lines. In general, these data resulted from the DNA segments of which the duplicated spots differed significantly in intensity owing to incomplete washing. In further analysis, we excluded data from DNA segments for which the duplicated spots differed in intensity by more than twofold. The normalized fluorescence intensities for 27 DNA segments were smaller than 0.01, a level equivalent to the background. In these cases, an insufficient amount of DNA was loaded on the glass slide, because of failure in the PCR amplification of DNA. Such DNA segments were also excluded from further analysis.

3.2. Global changes in gene expression after nitrogen deprivation

We used DNA microarrays to unravel global changes in gene expression after nitrogen deprivation. RNA was isolated from the whole filaments of cells that had been subjected to 0, 1, 3, 8, and 24 hr of nitrogen deprivation and were then used to generate fluorescently labeled cDNA. The relative ratio was calculated for the zero-time sample compared with each time-point sample. The experiment was carried out three times with independently
We first examined the expression of the *hetR* and *ntcA* genes, which encode heterocyst differentiation regulatory protein HetR and global nitrogen regulator NtcA, respectively, to confirm that the cells used in this series of experiments correctly represent cells deprived of combined nitrogen. Expression of *hetR* slightly increased at the third hour from the start of nitrogen deprivation and its increase was maximal at the eighth hour (Figure 2A). Transcripts of *ntcA* were barely detectable by the third hour but very prominent at the eighth hour. These results are consistent with the results of a previous report and indicate that the cells used in this experiment underwent nitrogen starvation.

Deprivation of combined nitrogen for 24 hr had a dramatic effect on the expression of many genes (Figure 1B). Especially, the fluorescence intensity for DNA segment containing the nitrogen fixation genes *nifUHD* was high after nitrogen deprivation (Figure 3A, outer major circle in red), while it was very low when provided with combined nitrogen (Figure 3A, middle major circle in blue). Figure 3A is a ‘GenoMap’ representation of the fluorescence intensity for each DNA segment as a function of its position within the genome. The fluorescence intensities for DNA segments containing the *atp* genes (encoding ATP synthase subunits), the *rbcLS* genes (encoding ribulose 1,5-bisphosphate carboxylase/oxygenase large and small subunits, respectively), and genes related to transcription and translation such as *rpl*, *rps*, *rpoA*, *infAB*, and *nusA* were high in a manner independent of the nitrogen status. Of the 2380 DNA segments (after excluding the 27 segments for which fluorescence was always faint), expression of 126 DNA segments increased and that of 44 DNA segments decreased at the 24th hour after the start of nitrogen deprivation, based on the result of the t-test (Figure 3B).

Our microarray contains multiple ORFs in each DNA segment. The changes in expression of individual ORFs were then analyzed by RNA gel blot analysis to identify new genes upregulated by nitrogen deprivation. DNA segment 3-1-5-5, in which expression increased by 12.0-fold (marked ‘a’ in Figure 3B), contained three ORFs, *all0473* (encoding a hypothetical protein), *all0474* (encoding a hypothetical protein), and *all0475* (encoding a putative oxidoreductase) (Figure 4A). The results indicated that *all0473* and *all0475* were upregulated by nitrogen deprivation. The *all0473*-specific probe detected a smear of RNA below 1.3 kb. A band at 0.6 kb was considered to be an artifact caused by the dense band of the endonucleolytic fragment of 23S rRNA. *alr1174*, encoding a rubrerythrin, which is a non-haem iron protein implicated in oxidative stress protection in anaerobic bacteria, was included in two DNA segments, 2-1-5-10 and 7-2-5-9, both of which showed increased expression by 2.2-fold and 2.7-fold, respectively (marked ‘b’ in Figure 3B, and Figure 4B). *alr1174* was, indeed, upregulated by nitrogen deprivation. Interestingly, the size of
Figure 2. RNA gel blot analysis of expression of genes upregulated by nitrogen deprivation. Total RNA was isolated from ammonium-grown cells (lane 0) and from cells subjected to nitrogen deprivation for 3, 8, and 24 hr (lanes marked with 3, 8, and 24, respectively). An aliquot of 5 µg of glyoxylated total RNA was loaded in each lane. (A) The membranes were hybridized to a hetR probe (upper panel) or an ntcA probe (middle panel). The sizes of transcripts are shown on the right of the panels. The band of 16S rRNA stained with methylene blue is shown in the lower panel as a loading control. (B) The membranes were hybridized to an all2004 probe (upper panel), an all2005 probe (middle panel) or an all2006 probe (lower panel).
Figure 3. ‘GenoMap’ representation of expression data. The data for each DNA segment is represented as a function of its position within the genome. The bars in the outermost and the second narrow circular band show the positions of ORFs in the clockwise and counter-clockwise directions, respectively. Inner major circles show expression data. (A) The labeled cDNAs prepared from RNA isolated from ammonium-grown cells and from cells deprived of combined nitrogen for 24 hr were mixed and hybridized to a microarray. The fluorescence intensities in the cells deprived of combined nitrogen for 24 hr are shown in the outer circle in red (-N) and those in ammonium-grown cells are in the middle circle in blue (NH₃). The innermost circle in purple shows the average GC content of every 2-kb segment of the genome. ORFs included in DNA segments for which fluorescence intensities are larger than 10.0 are labeled. (B) The increase ratio in the base-10 logarithm for each DNA segment at the 24th hour of nitrogen deprivation is shown in the outer circle and the decrease ratio is in the inner circle. DNA segments judged as being upregulated are showed in red and those judged as being downregulated are in blue, while the data without significant changes (p > 0.01 by t-test) are shown in gray. fdxH–nifB, all2511–all2528, alr2823–alr2843, and all5343–alr5360 indicate ORFs contained in each expressed island. DNA segments that have been subjected to RNA gel blot analysis in Figure 4 are indicated by Symbols a, b, and c. (C) The increase ratio in the base-10 logarithm for each DNA segment during the 24 hr growth without combined nitrogen. From the innermost circle towards the outermost circle, the increase ratios at the first, the third, the eighth and the 24th hour of nitrogen deprivation are shown, respectively. The scale of each circle is fixed to 2.2 in the base-10 logarithm. DNA segments judged as being upregulated at each time point are shown in green, green yellow, orange and red, respectively, while the data without significant changes are shown in gray. DNA segments upregulated by more than 5.0-fold are numbered and listed in Table 1. DNA segments which have indicated no significant changes in expression by RNA gel blot analysis are marked by filled circles. (D) DNA segments upregulated or downregulated in the heterocysts. The labeled cDNAs prepared from RNA isolated from whole filaments, consisting of vegetative cells and heterocysts, and from heterocyst-enriched fraction were mixed and hybridized to a microarray. The increase ratio in heterocyst-enriched fraction is shown in the outer circle and the decrease ratio is in the inner circle. DNA segments judged as being upregulated are showed in red and those judged as being downregulated are in blue, while the data without significant changes are shown in gray. DNA segments upregulated in heterocysts by more than 3.0-fold and downregulated by more than 2.0-fold are numbered and listed in Table 3.
Figure 4. RNA gel blot analysis of expression of each ORF contained in DNA segments upregulated by nitrogen deprivation for 24 hr. The uppermost horizontal lines indicate DNA segments and arrows indicate ORFs contained within each DNA segment. The increase ratio for each DNA segment is shown on the right of each line. Total RNA was isolated from ammonium-grown cells (lane NH$_3$) and from cells subjected to nitrogen deprivation for 24 hr (lane -N). An aliquot of 5 µg of glyoxylated total RNA was loaded in each lane. (A) Expression in DNA segment 3-1-5-5, containing three ORFs, all0473, alr0474, and all0475. The membranes were hybridized to an all0473 probe (left panel), an alr0474 probe (central panel) and an all0475 probe (right panel). (B) Expression in DNA segments 2-1-5-10 and 7-2-5-9, which overlapped each other on one end. alr1171, all1172, all1173, alr1174, and all1175 are contained in these two DNA segments. Probes used for each hybridization are indicated under each panel. (C) Expression in DNA segment 4-1-1-6, containing all1088 and all1089. The membranes were hybridized to an all1088 probe (left panel) and an all1089 probe (right panel).

The sizes of transcripts are shown on the right of the panels.

differentiation.$^{41,42}$ After excision events, these genes are transcribed as several distinct operons, such as nifB-fdxN-nifS-nifU, nifHDK, and hesAB-fdxH.$^{35,43,44}$ fdxH is also known to be transcribed as a monocistronic transcript.$^{45}$ In addition, nifENX is supposed to be transcribed as an operon, because in the soil bacterium Klebsiella pneumoniae, which has a nif gene organization similar to Anabaena, nifENX is transcribed as an operon.$^{46}$ Therefore, the expressed island around 1.70 Mb is 23 kb long and contains 17 ORFs, which consists of at least four transcription units, after the excision of two excisions.

Another large expressed island was located around position 6.39 Mb (No. 19 in Table 1). This region was 29 kb long and contained 18 ORFs, all5343 to alr5360 including hglE, D, and C genes (each encoding a heterocyst glycolipid synthase) and hetM, N, and I genes (encoding polyketide synthase, ketoacyl reductase and phosphopantetheinyltransferase, respectively).$^{47-49}$ It is indicated that the hetN gene is expressed primarily in heterocysts and is involved in the maintenance of heterocyst spacing.$^{50}$ In this region, all ORFs from all5343 to all5347 and from alr5348 to alr5358 (hetN) are oriented in the counter-clockwise and clockwise directions, respectively. Therefore, this expressed island also consists of several transcription units. Moreover, around position 3.02 Mb, there was a small expressed island of 16 kb (No. 9 in Table 1), which included fdxB (encoding a ferredoxin) and coxBACH (each encoding a cytochrome c oxidase subunit), which are expressed specifically in
Thus, Anabaena responds to the nitrogen deprivation with extensive and dynamic changes in gene expression. Its regulation might involve not only a single gene or transcription unit but also a larger region that we name ‘expressed island,’ encompassing multiple genes and multiple transcription units that are regulated in a coordinated manner.

### 3.4. Sequential changes in gene expression after nitrogen deprivation

Of the 2380 DNA segments, 340 DNA segments were upregulated at one or more time points during the 24 hr growth without combined nitrogen (Figure 3C and Table 1). Within 1 hr of nitrogen deprivation, expression in DNA segments including the *nir* operon (*nirA-nrtABCD-narB*, which encode nitrite reductase, four subunits of the ABC-type transport system for nitrate and nitrite, and nitrate reductase, respectively) increased by more than 10.0-fold (No. 1 in Figure 3C and Table 1). The expression in this operon decreased before the third hour and returned to the basal level at the eighth hour of nitrogen deprivation. At the third hour of nitrogen deprivation, expression in DNA segment 3-3-3-9 containing four ORFs (*all2004* to *all2007*) increased by 5.2-fold (No. 2 in Figure 3C and Table 1). The expression in this DNA segment continued to increase until the eighth hour. This change in expression was confirmed by RNA blot analysis (Figure 2B). The expression level of *all2007* was under the detection limit. Thus, the expression of genes in each DNA segment showed a different time course. The expression pattern for each DNA segment is profiled in Figure 6A. Of the 340 DNA segments that showed increased expression, 331 DNA segments that we were able to get complete data at all time points were used in the profiling of the expression pattern. Expression patterns were classified into three types: cluster 1 in which the expression in each DNA segment increased transiently at an early stage (1–3 hr) and then returned to the basal level (in the presence of ammonium...
Table 1. DNA segments upregulated by nitrogen deprivation. DNA segments in which expression increased by more than 5.0-fold at one or more time points during the 24 hr growth without combined nitrogen are listed.

| No.| DNA segment | ORF Cluster | Ref. |
|----|------------|-------------|------|
| 1.2-1.2 | 2394659 | alr0607(porC), alr0607(ntrA) | 1 |
| 4.2-4.7 | 704078 | alr0607(ntrA), alr0608(ntrA), alr0609(ntrB) | (52) |
| 1.3-1.4 | 708291 | alr0609(ntrB), alr0610(ntrC), alr0611(ntrD) | (76) |
| 5.2-2.9 | 710775 | alr0611(ntrD), alr0612(ntrB) | 1 |
| 1.4-2.2 | 712558 | alr0612(ntrB), alr0613, alr0614 | 1 |

No. corresponds to the number in Figure 3C. Expressed islands are indicated by bold face.

b DNA segments in expression was judged to be upregulated at each time point are shaded in gray. DNA segments for which data have shown high deviation between each experimental set are indicated by n.d. (no data).

c Cluster in which each DNA segment has been classified in Figure 6 is shown.

References indicate evidence for regulated gene expression. TS (this study) indicates DNA segments that have been confirmed to exhibit changed expression by RNA gel blot analysis.
anion) by the eighth hour, cluster 2 in which the expression in each DNA segment increased transiently at the eighth hour and then decreased, and cluster 3 in which expression was upregulated at the eighth hour or later and maximal at the 24th hour among the time points tested. Representative (i.e., highly expressed) DNA segments that fall into each group are summarized in Table 1.

Most of DNA segments belonging to cluster 1 also showed increased expression when cells were transferred to a medium containing nitrate as the sole nitrogen source from ammonium-containing medium (data not shown). DNA segments including the nir operon belonged to cluster 1. Expression of the nir operon is repressed by ammonium ion and induced by the removal of ammonium ion whether or not nitrate ion is present. This indicates that DNA segments belonging to cluster 1

**Figure 6.** (A) Profiling of expression pattern for DNA segments upregulated by nitrogen deprivation. DNA segments upregulated at one or more time points during the 24 hr growth without combined nitrogen were used in this profiling. Each unit on the x-axis represents one DNA segment and each time point was plotted on the y-axis. ‘H’ on the y-axis indicates the comparison of heterocyst-enriched fraction with whole filaments at the 48th hour of nitrogen deprivation. The height of the bar is proportional to the relative ratio in the base-2 logarithm for each DNA segment at each time point. The relative ratio in the base-2 logarithm for DNA segments judged as ‘no change’ were assigned a value of 0. DNA segments were classified into three clusters depending on the expression pattern, and then arranged in the ascending order of relative ratio within each cluster. (B) Schematic representation of various stages of heterocyst differentiation. The results of expression profiling are also shown.
Table 2. The number of ORFs belonging to each organism group.

|          | Ana   | Het   | All    | Others | total |
|----------|-------|-------|--------|--------|-------|
| cluster1 | 13 (19.7%) | 9 (13.6%) | 34 (51.5%) | 10 (15.1%) | 66    |
| cluster2 | 111 (14.7%) | 138 (18.3%) | 366 (48.4%) | 141 (18.7%) | 756   |
| cluster3 | 25 (13.6%) | 40 (21.7%) | 94 (51.1%) | 25 (13.6%) | 184   |
| chromosome | 832 (15.5%) | 900 (16.8%) | 2477 (46.1%) | 1160 (21.6%) | 5369 |

This analysis is based on comparative analysis of the genome of six species of cyanobacteria (Anabaena, Nostoc punctiforme PCC 73102, Synechocystis sp. PCC 6803, Synechococcus sp. WH8102, Prochlorococcus marinus MED4 and Prochlorococcus marinus MIT9313).\(^{58}\) Ana, Anabaena-specific homologue groups; Het, homologue groups shared by the heterocyst-forming cyanobacteria, Anabaena and N. punctiforme; All, homologue groups shared by all six species of cyanobacteria; Others, homologue groups not belonging to three organism groups mentioned above.

are not upregulated by nitrogen deprivation but rather are derepressed from repression by ammonium ion.

Cluster 2 included 252 DNA segments. Extensive changes in gene expression were detected at the eighth hour after the start of nitrogen deprivation. A large expressed island around 3.45 Mb appeared at the eighth hour after the first time (No. 10 in Figure 3C). Cluster 2 included genes required for the heterocyst differentiation, such as ntcA, hetC, and hetP.\(^{15,19,53,54}\) Those required for the formation of the heterocyst envelope, such as hgpB, C, hglK, and devBCA.\(^{40,55,56}\) and those required for the full activation of nitrogenase, such as prpA (encoding a protein phosphatase) and pknE (encoding a protein kinase).\(^{57}\) We still do not know if other unknown genes in this cluster are related to the stress response to nitrogen starvation or to the induction cascade of heterocyst differentiation.

In cluster 3, the genes related to the synthesis of glycolipids of the heterocyst envelope such as hglE, D, and C, and those related to the function of heterocysts such as the nif genes and hupSL (encoding the two subunits of uptake hydrogenase) were included. The result of homology-based grouping of ORFs contained in each cluster is shown in Table 2. This analysis was carried out based on the result of Sato (2002).\(^{58}\) ORFs belonging to groups shared by the only heterocyst-forming cyanobacteria (Het) was 16.8% of total ORFs on the chromosome, while 21.7% of ORFs contained in cluster 3 belonged to the Het group. This indicates that the number of ORFs related to the function and structure of heterocyst increases in cluster 3. In all clusters, the ratio of ORFs belonging to groups shared by all six species of cyanobacteria (All) was greater than that on the chromosome. The responses to nitrogen deprivation would be, at least partially, common in the cyanobacteria.

### 3.5 Genomic regions upregulated in heterocysts

Heterocysts are specialized cells for nitrogen fixation and differ from vegetative cells both physiologically and structurally. As described above, changes in gene expression occurred during the differentiation of heterocysts. We tried to identify the genomic regions that were upregulated in heterocysts by comparing the transcripts from heterocyst-enriched fraction with that from whole filaments consisting of vegetative cells and heterocysts 48 hr after the start of nitrogen deprivation. We chose 48 hr, because heterocysts are mature at this time and are resistant enough to the treatment used for heterocyst isolation. Another reason is that the cells recover normal growth on dinitrogen by 48 hr. In other words, the vegetative cells within the filaments containing heterocysts no longer suffer from nitrogen starvation.

DNA segments upregulated or downregulated in the heterocysts are displayed in a GenoMap in Figure 3D and are listed in Table 3. Sixty-six DNA segments, which contained 191 ORFs, were upregulated in the heterocysts. Genes known to be expressed mainly in the heterocysts, such as nifHDK, devA, hupL, and coxBACII.\(^{51,59-61}\) were included in the upregulated DNA segments. Three expressed islands around positions 3.02, 1.70, and 6.39 Mb were found to be upregulated in heterocysts (No. 9, 17, and 18 in Table 3 and Figure 3D, respectively). The expressed island around position 3.45 Mb (No. 10 in Table 1 and Figure 3C) was not upregulated in heterocysts. As this island is thought to be related to the synthesis of the polysaccharide layer of the heterocyst envelope, it would be expressed mainly in the heterocysts or cells in the process of heterocyst differentiation. Its expression was maximal at the eighth hour and then decreased by 4 to 7-fold at 24\(^{th}\) hour (No. 10 in Table 1), namely, its expression is higher in cells in the process of heterocyst differentiation, which is not resistant to the treatment used for heterocysts isolation, than in the heterocysts. So, it was not judged to be upregulated in heterocysts compared with whole filaments which included cells in the process of heterocyst differentiation. Forty-three of the 66 DNA segments upregulated in heterocysts were found to be upregulated at one
or more time points during the 24 hour growth without combined nitrogen (Figure 6A). All but one of them were upregulated at the eighth hour or later after the start of nitrogen deprivation. This suggests that the genes upregulated in the heterocysts had been tightly repressed until the time when they were required, which is later than the eighth hour of nitrogen deprivation. Eight DNA segments, which contained 29 ORFs, were downregulated in the heterocysts. Those included psbV (encoding the cytochrome c550), rbcL's and psbA1 (encoding the D1 protein of photosystem II) (No. 27, 28, and 30 in Table 3 and Figure 3D, respectively). This is consistent with the results of the previous study that the activities of ribulose 1,5-bisphosphate carboxylase and photosystem II in heterocysts are much lower than those in vegetative cells, and the promoter of rbcL's is inactive in heterocysts. DNA segment 6-3-1-7 was downregulated in the heterocysts by 0.4-fold (No. 29 in Table 3), but no ORF known to be downregulated in the heterocysts was contained.

Fifty-seven DNA segments were classified into cluster 3 (see above), of which 28 DNA segments were judged to be upregulated in heterocysts at a significance level of 1%. At a significance level of 5%, 40 DNA segments were judged to be upregulated in heterocysts. This suggests that 80%, at least 50%, of ORFs whose expression was maximal at the 24th hour of nitrogen deprivation were indeed expressed mainly in heterocysts and are likely to be related to the function or structure of heterocysts. The remaining DNA segments were not judged to be either upregulated or downregulated in heterocysts. This means that these DNA segments are upregulated in both vegetative cells and heterocysts by nitrogen deprivation and are related to the general responses to nitrogen starvation. No DNA segments classified into cluster 3 were downregulated in heterocysts.

Table 3. DNA segments that are upregulated or downregulated in heterocysts. DNA segments in which expression increased by more than 3.0-fold and decreased by more than 2.0-fold in heterocysts are listed.

| No. | DNA segment | start | end | Relative ratio | ORF | Cluster<sup>b</sup> | Ref. |
|-----|-------------|-------|-----|---------------|-----|----------------------|-----|
| 3   | 3-3-3-4     | 3016872 | 3020382 | 3.4 ± 1.4 | all2511, all2512(pcbB), assr2513(lichtB) | 3   |
| 4   | 2-4-3-9     | 3022830 | 3026249 | 3.2 ± 0.3 | all2515(cocAI), all2516(lichtB), all2517, all2518 | 2   |
| 5   | 2-1-3-7     | 3025983 | 3028742 | 3.3 ± 1.4 | all2518, assr2519, all2520, all2521 | 3   |
| 6   | 3-3-1-5     | 3026888 | 3030440 | 3.9 ± 1.0 | all2520, all2521, assr2522, assr2523, assr2524 | 3   |
| 7   | 15-3-4-5    | 795168 | 798517 | 4.2 ± 1.9 | all0687(hupL), all0688(hupS), assr0689, assr0690, assr0691 | 3   |
| 8   | 2-4-2-6     | 1694646 | 1697534 | 3.7 ± 1.5 | allH436(nfX), allH437(nfN), allH438(nfE) | 3   |
| 9   | 5-2-5-5     | 1696082 | 1699389 | 4.0 ± 1.5 | allH437(nfN), allH438(nfE), allH439, allH440(nfK) | 3   |
| 10  | 5-3-3-4     | 1699214 | 1702192 | 4.1 ± 2.1 | allH440(nfK), allH445(nfD), allH442(xia) | 3   |
| 11  | 3-4-1-2     | 1711421 | 1715052 | 3.4 ± 1.3 | allH454(nfD), allH455(nfH), allH456(nfL) | 3   |
| 12  | 8-1-2-11    | 337296 | 3330863 | 5.6 ± 1.6 | all2729, all2730, all2731(cocAI), all2732(cocAI) | -   |
| 13  | 5-4-4-14    | 6374252 | 6377651 | 3.1 ± 0.9 | allL541, allL542, allL543 | 3   |
| 14  | 2-4-4-7     | 6376550 | 6379644 | 5.2 ± 1.3 | allL543, allL544, allL545, allL546 | -   |
| 15  | 7-2-1-2     | 6378843 | 6382399 | 4.2 ± 1.8 | allL545, allL546, allL547 | 3   |
| 16  | 2-3-5-6     | 6378200 | 6390176 | 5.4 ± 2.6 | allL551(hglE) | -   |
| 17  | 5-1-5-13    | 6389246 | 6392141 | 3.7 ± 1.8 | allL551(hglE), allL5352, allL5353 | 3   |
| 18  | 1-2-5-12    | 6392481 | 6397616 | 7.2 ± 2.1 | allL5354(hglD), allL5355(hglC) | 3   |
| 19  | 8-2-4-2     | 6372753 | 6399905 | 3.1 ± 1.3 | allL553(hglC), allL5356 | -   |
| 20  | 4-2-1-13    | 6398882 | 6402454 | 4.5 ± 2.2 | allL556, allL557(hetM), allL558(hetN) | 3   |
| 21  | 1-2-4-9     | 6403144 | 6405603 | 3.8 ± 1.0 | allL557(hetM), allL558(hetN), allL559(hetL), allL560 | 3   |
| 22  | 8-3-3-11    | 1870784 | 1901249 | 2.9 ± 0.9 | alr0177, alr0178 | 3   |
| 23  | 3-5-3-13    | 398853 | 401942 | 3.2 ± 0.8 | alr0348(nidD), alr0349, all0350 | 2   |
| 24  | 5-3-6-14    | 433657 | 436879 | 3.7 ± 1.3 | alr0370, all0371, alr0372, all0373 | -   |
| 25  | 3-3-1-51    | 805531 | 853733 | 3.4 ± 1.4 | allL573, all0730, all0731 | 3   |
| 26  | 2-4-5-14    | 1669919 | 1672474 | 3.5 ± 1.1 | alr1404, assr1405, assr1406, alr1407(nfV1) | 3   |
| 27  | 1-2-2-2     | 1672119 | 1674867 | 3.6 ± 1.7 | alr1407(nfV1), assr1408(nfZ), alr1409(nfT), alr1410, all1411, assr1412 | 3   |
| 28  | 5-1-4-1     | 1930908 | 1934296 | 2.8 ± 0.9 | allL635, allL636 | -   |
| 29  | 2-1-4-13    | 3785305 | 3788017 | 3.1 ± 1.3 | allL5326, allL5327 | 3   |
| 30  | 8-2-2-7     | 277339 | 280046 | 0.5 ± 0.1 | alr0258(petE), alr0259(psbV), all0260, all0261 | (62) |
| 31  | 6-2-6-13    | 1786516 | 1790163 | 0.4 ± 0.3 | alr1524(rbchi), alr1525(rbxi), alr1526(rbces), alr1527 | (59) |
| 32  | 6-3-1-7     | 2462070 | 2465599 | 0.4 ± 0.2 | alr2054, alr2055, all0205, all2057(aroK) | TS |
| 33  | 8-1-1-8     | 5794576 | 5797024 | 0.5 ± 0.1 | alr4863, alr4864, asr4865, asr4866(psbA1) | TS |

*a No. corresponds to the number in Figure 3D. Expressed islands are indicated by bold face.

*b Cluster in which each DNA segment has been classified in Figure 6 is shown.

References indicate evidence for regulated gene expression. TS (this study) indicates DNA segments that have been confirmed to exhibit changed expression by RNA gel blot analysis.
4. Discussion

The present study reports an initial attempt to identify genes that are regulated after nitrogen deprivation within the whole genome of a cyanobacterium Anabaena sp. strain PCC 7120. To this end, we prepared a custom-made microarray containing 2407 DNA segments that corresponded to 90% of the whole chromosome of Anabaena (N. Sato et al., under submission). Then, this microarray was hybridized with various fluorescently labeled cDNAs prepared from the RNA isolated from the cells during the course of deprivation of combined nitrogen in the medium. The data indicate clearly that the expression of various genes known to be regulated by nitrogen deprivation was in fact regulated in a manner exactly as expected. In addition, the same data suggest that a number of hitherto unknown genes are also regulated by the same treatment. The functions of most of these newly identified genes are still unknown and are targets of future research. The number of genes that were downregulated during the course of nitrogen deprivation, including various genes mostly involved in photosynthesis, was half of the number of upregulated genes. In addition, the extent of decrease in the expression of these genes was not as large as the extent of increase in the expression of the genes that were upregulated by nitrogen deprivation.

4.1. Overview of gene expression during deprivation of combined nitrogen

Of the 2380 DNA segments, 340 DNA segments, containing 1018 ORFs, were judged as being upregulated at one or more time points during the 24 hour growth without combined nitrogen. Based on the results of RNA blot analyses that were performed for 27 ORFs contained in 9 upregulated DNA segments, about 60% of the selected ORFs were upregulated. In other words, about 600 ORFs are estimated to be upregulated by nitrogen deprivation. This value is almost comparable to the value of 400–500 ORFs estimated by Curtis and Hebar (2001) who carried out a screen for sequences upregulated during heterocyst differentiation using a library representing approximately 2% of the Anabaena genome. Genes upregulated by nitrogen deprivation were not evenly distributed throughout the genome but formed expressed islands. So our estimated value is a little higher than that of previous report, in which the existence of expressed islands was not taken into consideration. We also identified 171 DNA segments, containing 520 ORFs, that were downregulated. These include genes involved in photosynthesis and carbon dioxide fixation, and those encoding ATP synthase, transporters of sulfate, phosphate and iron, and ribosomal proteins. This indicates that the activity of various metabolic pathways are decreased to maintain a balance of various nutrients under the nitrogen-limited condition.

The time-course experiment showed that the response to nitrogen deprivation in Anabaena can be divided into several phases (Figure 6). First, genes that have been repressed by ammonium ion are derepressed within 1 hr of nitrogen deprivation (cluster 1 in Figure 6A and ‘1–3 hr’ in Figure 6B). Transporters and metabolic enzymes are induced to use alternate nitrogen sources, such as nitrate. Second, heterocyst differentiation is triggered by nitrogen starvation at about the eighth hour of nitrogen deprivation. Marked changes in gene expression occurred at this phase (Figure 3C). Genes required for heterocyst differentiation and related to the function and structure of heterocysts are upregulated, perhaps in the proheterocysts (cluster 2 in Figure 6A and ‘8 hr’ in Figure 6B). The eighth hour of nitrogen deprivation corresponds to the timing of the commitment to complete heterocyst differentiation. ORFs that play an important role in the determination of the cell fate might be included in the 138 ORFs belonging to the homologue groups shared by the heterocyst-forming cyanobacteria, Anabaena and Nostoc punctiforme (Table 2). A lot of ORFs contained in cluster 2 might also be involved in the responses to nitrogen starvation. It is, however, difficult to classify which ORFs are involved in the responses to nitrogen starvation or heterocyst differentiation. A role in heterocyst differentiation cannot be excluded a priori for those ORFs expressed transiently at only the eighth hour of nitrogen deprivation, since the expression of hetK, hetC, and hetP (Figure 6) was found only at the eighth hour. Analysis of the responses to nitrogen starvation in non nitrogen-fixing cyanobacteria such as Synechocystis sp. PCC 6803 would help solve this problem. Finally, genes related to heterocyst function are upregulated in the heterocysts at the 24th hour of nitrogen deprivation (cluster 3 in Figure 6A and ‘24 hr’ in Figure 6B). Fifty to 80% of ORFs whose expression was maximal at the 24th hour were preferentially expressed in heterocysts. Expressed islands were also upregulated in heterocysts. At the 24th hour of nitrogen deprivation, some ORFs were upregulated in both vegetative cells and heterocysts. These ORFs might play a role in diazotrophic growth.

4.2. Expressed islands

We found that many of the upregulated genes were physically clustered in islands on the Anabaena chromosome (Figure 3B). We developed a tool called GenoMap, which can conveniently illustrate expression data over the circular chromosome. This tool enabled us to visualize large expressed islands. These regions that we call expressed islands are as long as 30 kb in size and include multiple transcription units (putative operons) (Table 1). Each expressed island contains functionally related ORFs. Namely, ORFs contained in the expressed island around position 1.70 Mb are related to nitrogen fixation (No. 17 in Table 1). ORFs...
around 3.45 Mb are related to the synthesis of the polysaccharide layer of the heterocyst envelope (No. 10 in Table 1) and ORFs around 6.39 Mb are related to the synthesis of the glycolipid layer of the heterocyst envelope (No. 19 in Table 1). The expressed island around position 3.02 Mb contains coxBACII, which are expressed specifically in heterocysts,51 but most of the remaining ORFs have unknown function. As these ORFs belong to the ‘Het’ category of homologue groups based on the analysis in Table 2, these would be also related to the heterocyst function. There are a few reports that genes showing a similar expression pattern are clustered on the chromosomes of eukaryotes.65,66 The mechanism underlying this regulation remains unknown.

One idea comes from the well-documented mechanism of the regulation of the hepA and hepC genes included in the expressed island around position 3.45 Mb.60,67 A mutation in hepK, encoding a histidine kinase, blocks induction of hepA and mutations in abp2 and abp3 block expression of hepA and hepC. A large cascade of transcriptional activation might account for the induction of many other genes in the cluster that includes hepA and hepC. However, it is not clear whether this regulatory cascade extends over the whole expressed island around position 3.45 Mb.

Another simple but plausible mechanism regulating the expressed islands involves structural changes in the chromosome. Changes in the chromatin structure associated with activation and silencing of gene expression are of paramount importance during development in eukaryotes. In bacteria, some DNA-binding proteins, such as HU, which is known as heat-unstable nucleoid protein, but is heat stable,68 IHF (integration host factor) and Dps (DNA-binding protein from starved cells), form a nucleoprotein complex with the chromosome, called a nucleoid, and affect the local as well as global structure of the chromosome.69 In Anabaena, only two kinds of nucleoid protein are present, namely, four ORFs each encoding putative Dps (all0458, all1173, ab3808, and all4145) and one ORF encoding HU (asr3935) which has been named hana.70 Two of the four dps-like genes were upregulated by nitrogen deprivation, but their expression was not confined to heterocysts (Figure 4B and data not shown). HU was first reported to be present in vegetative cells but not in heterocysts, and was shown to be able to condense the DNA.71 HU was then found to be essential for the initiation of heterocyst differentiation.70 Based on these findings, the following mechanism could explain the observed expressed islands. In vegetative cells, the chromosome around expressed islands is packed by HU and the expression of these regions is structurally repressed. During heterocyst differentiation, HU is degraded and the structure of the chromosome around expressed islands is destroyed. This structural change enables RNA polymerase and transcription factors to interact with their recognition sequences. Thus, expression of expressed islands is induced in a coordinated manner.

Another possible explanation of expressed islands, among others, is DNA methylation, which provides a mechanism for stably altering the local structure of the chromosome. The degree of methylation is tightly controlled and plays an important role in the cell-specific regulation of gene expression in the cells of higher eukaryotes.72 In E. coli, it is indicated using microarray that deoxyadenosine methyltransferase (Dam) controls the expression of a large number of genes.73 DNA of Anabaena is resistant to many restriction endonucleases, perhaps due to methylation.74 and it has been deduced that Anabaena has nine DNA methyltransferases (MTases).75 Four of them are classified as solitary MTases, which are excluded from a restriction/modification system, and might be involved in the regulation of gene expression. It is shown that mutation of the avaMV gene, encoding a DNA methyltransferase M.AvaV, appears to alter the frequency of heterocysts.75 In vegetative cells, expressed islands might be methylated and their expression is repressed, and this possibility is being tested in our laboratory.

4.3. Concluding remarks

In the present study, various genes upregulated by nitrogen deprivation have been identified with a custom-made microarray. The use of microarrays of DNA segments was successful in screening out DNA regions in which gene expression changes drastically during nitrogen deprivation in Anabaena. The genes within the selected DNA segments are individually checked by RNA blot analysis to determine if they are really regulated. This approach largely works well in the identification of nitrogen-regulated genes in Anabaena, as shown in the Results section. There are, however, a number of fortunate situations that led to success. The detection of changes in the expression level of 1 out of 3 or 4 genes within a DNA segment of 3–4 kbp was feasible for the following reasons: (1) The background level was very low in our microarray, mostly due to the tight binding of DNA that allowed stringent washing after hybridization. (2) The genes in Anabaena genome are, unlike the genes in Synechocystis sp. PCC 6803, polycistronic in most cases. (3) The changes in transcript levels caused by nitrogen deprivation were very large, and the changes were mostly upregulation. However, our approach might have failed in detecting positive signals in the following cases. (1) If two genes within a single DNA segment are differently regulated, e.g., one gene is upregulated while another is downregulated, the DNA segment might be recognized as being neither upregulated nor downregulated. (2) The detection limit of weakly expressed genes, such as those encoding regulatory proteins, is likely to be higher in a microarray with DNA segments than in a microarray with gene-specific probes, but we have no definite esti-
mate of this disadvantage.

The function of most of the genes that were identified to be regulated by nitrogen deprivation remains unknown. It is important to identify the function of these genes for further understanding of the mechanisms of heterocyst differentiation. A large-scale disruption of genes will be helpful in estimating the function of genes that are upregulated during nitrogen deprivation. In addition, whole-genome expression analysis of various mutants that are known to affect the process of heterocyst formation as well as mutants that are related to nitrogen stress will shed further light on the function of nitrogen-regulated genes that were identified in the present study.

Acknowledgements: We thank Dr. C. Peter Wolk (Michigan State University) for critical reading of the manuscript. This work was in part supported by Grants-in-Aid for Scientific Research from Japanese Society for Promotion of Science to NS (Nos. 13440234 and 12874104) and a Grant-in-Aid for Priority Areas (Genome Biology) from the Ministry of Education, Science, Culture, Sports and Technology, Japan, to MO.

References

1. Wolk, C. P., Ernst, A., and Elhai, J. 1994, In: Bryant, D. A. (ed) The molecular biology of cyanobacteria, Kluwer Academic Publishers, Dordrecht, pp. 769–823.
2. Golden, J. W. and Yoon, H. S. 1998, Heterocyst formation in Anabaena, Curr. Opin. Microbiol., 1, 623–629.
3. Wolk, C. P. 2000, In: Brun, Y. V. and Shimkets, L. J. (eds) Prokaryotic development, American Society for Microbiology, Washington, D.C., pp. 83–104.
4. Kroos, L., Zhang, B., Ichikawa, H., and Yu, Y. T. 1999, Control of sigma factor activity during Bacillus subtilis sporulation, Mol. Microbiol., 31, 1285–1294.
5. Brahamsha, B. and Haselkorn, R. 1991, Isolation and characterization of the gene encoding the principal sigma factor of the vegetative cell RNA polymerase from the cyanobacterium Anabaena sp. Strain PCC 7120, J. Bacteriol., 173, 2412–2450.
6. Brahamsha, B. and Haselkorn, R. 1992, Identification of multiple RNA polymerase sigma factor homologs in the cyanobacterium Anabaena sp. Strain PCC 7120: cloning, expression, and inactivation of the sigB and sigC genes, J. Bacteriol., 174, 7273–7282.
7. Khudyakov, I. Y. and Golden, J. W. 2001, Identification and inactivation of three group 2 sigma factor genes in Anabaena sp. Strain PCC 7120, J. Bacteriol., 183, 6667–6675.
8. Fogg, G. E. 1951, Growth and heterocyst production in Anabaena cylindrica Lemm. III. The cytology of heterocysts, Ann. Bot. n.s., 15, 23–35.
9. Fleming, H. and Haselkorn, R. 1974, The program of protein synthesis during heterocyst differentiation in nitrogen-fixing blue-green algae, Cell, 3, 159–170.
10. Lynn, M. E., Bantle, J. A., and Ownby, J. D. 1986, Estimation of gene expression in heterocysts of Anabaena variabilis by using DNA-RNA hybridization, J. Bacteriol., 167, 940–946.
11. Buikema, W. J. and Haselkorn, R. 1991, Isolation and complementation of nitrogen fixation mutants of the cyanobacterium Anabaena sp. Strain PCC 7120, J. Bacteriol., 173, 1879–1885.
12. Ernst, A., Black, T., Cai, Y., Panoff, J. M., Tiwari, D. N., and Wolk, C. P. 1992, Synthesis of nitrogenase in mutants of the cyanobacterium Anabaena sp. Strain PCC 7120 affected in heterocyst development or metabolism, J. Bacteriol., 174, 6025–6032.
13. Buikema, W. J. and Haselkorn, R. 1991, Characterization of a gene controlling heterocyst differentiation in the cyanobacterium Anabaena 7120, Genes Dev., 5, 321–330.
14. Buikema, W. J. and Haselkorn, R. 2001, Expression of the Anabaena hetR gene from a copper-regulated promoter leads to heterocyst differentiation under repressing conditions, Proc. Natl. Acad. Sci. USA, 98, 2729–2734.
15. Frías, J. E., Flores, E., and 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.
16. Muro-Pastor, A. M., Valladares, A., Flores, E., and Herrero, A. 1999, The hetC gene is a direct target of the NtcA transcriptional regulator in cyanobacterial heterocyst development, J. Bacteriol., 181, 6664–6669.
17. Valladares, A., Muro-Pastor, A. M., Fillat, M. F., Herrero, A., and Flores, E. 1999, Constitutive and nitrogen-regulated promoters of the petH gene encoding ferredoxin:NADP+ reductase in the heterocyst-forming cyanobacterium Anabaena sp., FEBS Lett., 449, 159–164.
18. Fiedler, G., Muro-Pastor, A. M., Flores, E., and Maldener, I. 2001, NtcA-dependent expression of the devBCA operon, encoding a heterocyst-specific ATP-binding cassette transporter in Anabaena spp., J. Bacteriol., 183, 3795–3799.
19. Wei, T.-F., Ramasubramanian, T. S., and Golden, J. W. 1994, Anabaena sp. Strain PCC 7120 ntcA gene required for growth on nitrate and heterocyst development, J. Bacteriol., 176, 4473–4482.
20. Cai, Y. and Wolk, C. P. 1997, Anabaena sp. Strain PCC 7120 responds to nitrogen deprivation with a cascade-like sequence of transcriptional activations, J. Bacteriol., 179, 267–271.
21. Schena, M., Shalon, D., Davis, R. W., and Brown, P. O. 1995, Quantitative monitoring of gene expression patterns with a complementary DNA microarray, Science, 270, 467–470.
22. Hihara, Y., Kamei, A., Kanehisa, M., Kaplan, A., and Ikuchi, M. 2001, DNA microarray analysis of cyanobacterial gene expression during acclimation to high light, Plant Cell, 13, 793–806.
23. Suzuki, I., Kanesaki, Y., Mikami, K., Kanehisa, M., and Murata, N. 2001, Cold-regulated genes under control of the cold sensor Hik33 in Synechocystis, Mol. Microbiol., 40, 235–244.
24. Gill, R. T., Katsoulakis, E., Schmitt, W., Taroncher-Oldenburg, G., Misra, J., and Stephanopoulos, G. 2002, Genome-wide dynamic transcriptional profiling of the light-to-dark transition in Synechocystis sp. Strain PCC 6803, J. Bacteriol., 184, 3671–3681.
25. Kanekyo, Y., Suzuki, I., Allakhverdiev, S. I., Mikami, K., and Murata, N. 2002, Salt stress and hyperosmotic stress regulate the expression of different sets of genes in Synechocystis sp. PCC 6803, Biochem. Biophys. Res. Commun., 290, 339–348.

26. Yoshimura, H., Yanagisawa, S., Kamehisa, M., and Ohmori, M. 2002, Screening for the target gene of cyanobacterial cAMP receptor protein SYCRP1, Mol. Microbiol., 43, 843–853.

27. Wang, R., Guegler, K., LaBrie, S. T., and Crawford, N. M. 2000, Genomic analysis of a nutrient response in Arabidopsis reveals diverse expression patterns and novel metabolic and potential regulatory genes induced by nitrate, Plant Cell, 12, 1491–1509.

28. Wang, Y. H., Garvin, D. F., and Kochian, L. V. 2001, Nitrate-induced genes in tomato roots. Array analysis reveals novel genes that may play a role in nitrogen nutrition, Plant Physiol., 127, 345–359.

29. Che, P., Gingerich, D. J., Lall, S., and Howell, S. H. 2002, Global and Hormone-Induced Gene Expression Changes during Shoot Development in Arabidopsis, Plant Cell, 14, 2771–2785.

30. Chu, S., DeRisi, J., Eisen, M. et al. 1998, The transcriptional program of sporulation in budding yeast, Science, 282, 699–705.

31. Mata, J., Lyne, R., Burns, G., and Bahler, J. 2002, The transcriptional program of meiosis and sporulation in fission yeast, Nat. Genet., 32, 143–147.

32. Kaneko, T., Nakamura, Y., Wolk, C. P. et al. 2001, Complete genomic sequence of the filamentous nitrogen-fixing cyanobacterium Anabaena sp. Strain PCC 7120, DNA Res., 8, 205–213.

33. Rippka, R., Deruelles, J., Waterbury, J. B., Herdman, M., and Stanier, R. Y. 1979, Genetic assignments, strain characteristics, the evolutionary relationships of genera of purple nonsulfur bacteria, J. Gen. Microbiol., 111, 1–61.

34. Sato, N. 1995, A family of cold-regulated RNA-binding protein genes in the cyanobacterium Anabaena variabilis M3, Nucleic Acids Res., 23, 2161–2167.

35. Golden, J. W., Whorff, L. L., and West, D. R. 1991, Independent regulation of nif/HDK operon transcription and DNA rearrangement during heterocyst differentiation in the cyanobacterium Anabaena sp. Strain PCC 7120, J. Bacteriol., 173, 7098–7105.

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

37. Sato, N. and Ehira, S. GenoMap, a circular genome data viewer, Bioinformatics, in press.

38. Doolittle, W. F. 1973, Postmaturational cleavage of 23s ribosomal ribonucleic acid and its metabolic control in the blue-green alga Anacystis nidulans, J. Bacteriol., 113, 1256–1263.

39. Sztukowska, M., Bugno, M., Potempa, J., Travis, J., and Kurtz, D. M., Jr. 2002, Role of rubrerythrin in the oxidative stress response of Porphyromonas gingivalis, Mol. Microbiol., 44, 479–488.

40. Zhu, J., Kong, R., and Wolk, C. P. 1998, Regulation of hepA of Anabaena sp. Strain PCC 7120 by elements 5′ from the gene and by hepK, J. Bacteriol., 180, 4233–4242.

41. Golden, J. W., Robinson, S. J., and Haselkorn, R. 1985, Rearrangement of nitrogen fixation genes during heterocyst differentiation in the cyanobacterium Anabaena, Nature, 314, 419–423.

42. Golden, J. W., Carrasco, C. D., Mulligan, M. E., Schneider, G. J., and Haselkorn, R. 1988, Deletion of a 55-kilobase-pair DNA element from the chromosome during heterocyst differentiation of Anabaena sp. Strain PCC 7120, J. Bacteriol., 170, 5034–5041.

43. Mulligan, M. E. and Haselkorn, R. 1989, Nitrogen fixation (nif) genes of the cyanobacterium Anabaena species strain PCC 7120. The nifB-nfxN-nifS-nifU operon, J. Biol. Chem., 264, 19290–19297.

44. Borthakur, D., Basche, M., Buikema, W. J., Borthakur, P. B., and Haselkorn, R. 1990, Expression, nucleotide sequence and mutational analysis of two open reading frames in the nif gene region of Anabaena sp. Strain PCC7120, Mol. Gen. Genet., 221, 227–234.

45. Böhme, H. and Haselkorn, R. 1988, Molecular cloning and nucleotide sequence analysis of the gene encoding for heterocyst ferredoxin from the cyanobacterium Anabaena sp. Strain PCC 7120, Mol. Gen. Genet., 214, 278–285.

46. Haselkorn, R. 1992, Developmentally regulated gene rearrangements in prokaryotes, Annu. Rev. Genet., 26, 113–130.

47. Black, T. A. and Wolk, C. P. 1994, Analysis of a Heterocyst mutation in Anabaena sp. Strain PCC 7120 implicates a secondary metabolite in the regulation of heterocyst spacing, J. Bacteriol., 176, 2282–2292.

48. Lambalot, R. H., Gehring, A. M., Flugel, R. S. et al. 1996, A new enzyme superfamily the phosphopantetheinyI transferases, Chem. Biol., 3, 923–936.

49. Bauer, C. C., Ramaswamy, K. S., Endley, S., Scappino, L. A., Golden, J. W., and Haselkorn, R. 1997, Suppression of heterocyst differentiation in Anabaena PCC 7120 by a cosmid carrying wild-type genes encoding enzymes for fatty acid synthesis, FEMS Microbiol. Lett., 151, 25–30.

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

51. Jones, K. M. and Haselkorn, R. 2002, Newly identified cytochrome c oxidase operon in the nitrogen-fixing cyanobacterium Anabaena sp. Strain PCC 7120, J. Bacteriol., 184, 2491–2499.

52. Frias, J. E., Flores, E., and Herrero, A. 1997, Nitrate assimilation gene cluster from the heterocyst-forming cyanobacterium Anabaena sp. Strain PCC 7120, J. Bacteriol., 179, 477–486.

53. Khudyakov, I. and Wolk, C. P. 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.

54. Fernández-Piñas, F., Leganes, F., and Wolk, C. P. 1994, A third genetic locus required for the formation of heterocysts in Anabaena sp. Strain PCC 7120, J. Bacteriol., 176, 5277–5283.
55. Black, K., Buikema, W. J., and Haselkorn, R. 1995, The hglK gene is required for localization of heterocyst-specific glycolipids in the cyanobacterium *Anabaena* sp. Strain PCC 7120, *J. Bacteriol.*, 177, 6440–6448.

56. Fiedler, G., Arnold, M., Hannus, S., and Maldener, I. 1998, The DevBCA exporter is essential for envelope formation in heterocysts of the cyanobacterium *Anabaena* sp. Strain PCC 7120, *Mol. Microbiol.*, 27, 1193–1202.

57. Zhang, C. C., Friry, A., and Peng, L. 1998, Molecular and genetic analysis of two closely linked genes that encode, respectively, a protein phosphatase 1/2A/2B homolog and a protein kinase homolog in the cyanobacterium *Anabaena* sp. Strain PCC 7120, *J. Bacteriol.*, 180, 2616–2622.

58. Sato, N. 2002, Comparative analysis of the genomes of cyanobacteria and plants, *Genome Informatics*, 13, 173–182.

59. Elhai, J. and Wolk, C. P. 1990, Developmental regulation and spatial pattern of expression of the structural genes for nitrogenase in the cyanobacterium *Anabaena*, *EMBO J.*, 9, 3379–3388.

60. Maldener, I., Fiedler, G., Ernst, A., Fernandez-Pinas, F., and Wolk, C. P. 1994, Characterization of *devA*, a gene required for the maturation of proheterocysts in the cyanobacterium *Anabaena* sp. strain PCC 7120, *J. Bacteriol.*, 176, 7543–7549.

61. Carrasco, C. D., Buettner, J. A., and Golden, J. W. 1995, Programmed DNA rearrangement of a cyanobacterial gene in heterocysts, *Proc. Natl. Acad. Sci. USA*, 92, 791–795.

62. Wolk, C. P. 1982, In: Carr, N. G. and Whitton, B. A. (eds) The biology of cyanobacteria, Blackwell Scientific Publications, Oxford, pp. 359–386.

63. Curtis, S. E. and Hebbard, P. B. 2001, A screen for sequences up-regulated during heterocyst development in *Anabaena* sp. Strain PCC 7120, *Arch. Microbiol.*, 175, 313–322.

64. Yoon, H. S. and Golden, J. W. 2001, Pats and products of nitrogen fixation control heterocyst pattern, *J. Bacteriol.*, 183, 2695–2613.

65. Cohen, B. A., Mitra, R. D., Hughes, J. D., and Church, G. M. 2000, A computational analysis of whole-genome expression data reveals chromosomal domains of gene expression, *Nat. Genet.*, 26, 183–186.

66. Spellman, P. T. and Rubin, G. M. 2002, Evidence for large domains of similarly expressed genes in the *Drosophila* genome, *J. Biol.*, 1, 5.

67. Koksharova, O. A. and Wolk, C. P. 2002, Novel DNA-Binding Proteins in the Cyanobacterium *Anabaena* sp. Strain PCC 7120, *J. Bacteriol.*, 184, 3931–3940.

68. Rouvière-Yaniv, J. and Gros, F. 1975, Characterization of a novel, low-molecular-weight DNA-binding protein from *Escherichia coli*, *Proc. Natl. Acad. Sci. USA*, 72, 3428–3432.

69. Pettijohn, D. E. 1996, In: Neidhardt, F. C. Curtis III, R. Ingraham, J. L. et al. (eds) *Escherichia coli and Salmonella*: cellular and molecular biology, American Society for Microbiology, Washington, D.C., pp. 158–166.

70. Khudyakov, I. and Wolk, C. P. 1996, Evidence that the *hanA* gene coding for HU protein is essential for heterocyst differentiation in, and cyanophage A-4(L) sensitivity of, *Anabaena* sp. strain PCC 7120, *J. Bacteriol.*, 178, 3572–3577.

71. Nagaraja, R. and Haselkorn, R. 1994, Protein HU from the cyanobacterium *Anabaena*, *Biochimie*, 76, 1082–1089.

72. Cedar, H., Arfin, S. M., Long, A. D. et al. 1988, DNA methylation and gene activity, *Cell*, 53, 3–4.

73. Oshima, T., Wada, C., Kawagoe, Y. et al. 2002, Genome-wide analysis of deoxyadenosine methyltransferase-mediated control of gene expression in *Escherichia coli*, *Mol. Microbiol.*, 45, 673–695.

74. Padhy, R. N., Hottat, F. G., Coene, M. M., and Hoet, P. P. 1988, Restriction analysis and quantitative estimation of methylated bases of filamentous and unicellular cyanobacterial DNAs, *J. Bacteriol.*, 170, 1934–1939.

75. Matveev, A. V., Young, K. T., Meng, A., and Elhai, J. 2001, DNA methyltransferases of the cyanobacterium *Anabaena PCC 7120*, *Nucleic Acids Res.*, 29, 1491–1506.

76. Cai, Y. and Wolk, C. P. 1997, Nitrogen deprivation of *Anabaena* sp. strain PCC 7120 elicits rapid activation of a gene cluster that is essential for uptake and utilization of nitrate, *J. Bacteriol.*, 179, 258–266.

77. Stricker, O., Masepohl, B., Klipp, W., and Böhm, H. 1997, Identification and characterization of the nifV-nifZ-nifT gene region from the filamentous cyanobacterium *Anabaena* sp. strain PCC 7120, *J. Bacteriol.*, 179, 2930–2937.