Nitrogen fixation in microbial mats: complexity and dynamics of a small-scale ecosystem
Severin, I.

Citation for published version (APA):
Chapter 7

Horizontal transfer of the nitrogen fixation gene cluster in the cyanobacterium Microcoleus chthonoplastes

Henk Bolhuis\textsuperscript{1}, Ina Severin\textsuperscript{1}, Veronique Confurius-Guns\textsuperscript{1}, Ute I. A. Wollenzien\textsuperscript{1} and Lucas J. Stal\textsuperscript{1,2}

\textsuperscript{1}Department of Marine Microbiology, Netherlands Institute of Ecology, NIOO-KNAW; \textsuperscript{2}Department of Aquatic Microbiology, Institute of Biodiversity and Ecosystem Dynamics, University of Amsterdam

Published in: The ISME Journal
Running title: nitrogen fixation in Microcoleus chthonoplastes
Keywords: Cyanobacteria, Microcoleus, nitrogen fixation, horizontal gene transfer, microbial mat
Abstract

The filamentous, non-heterocystous cyanobacterium *Microcoleus chthonoplastes* is a cosmopolitan organism, known to build microbial mats in a variety of different environments. Although most of these cyanobacterial mats are known for their capacity to fix dinitrogen, *M. chthonoplastes* has not been assigned as a diazotrophic organism. None of the strains that were correctly identified as *M. chthonoplastes* has been shown to fix dinitrogen and it has repeatedly been reported that these organisms lacked the cyanobacterial *nifH*, the structural gene for dinitrogenase reductase. Here we show that a complete *nif*-gene cluster is present in the genome of *M. chthonoplastes* PCC 7420 and that the three structural nitrogenase genes, *nifHDK*, are present in a collection of axenic strains of *M. chthonoplastes* from distant locations. Phylogenetic analysis of *nifHDK* revealed that they cluster with the Deltaproteobacteria and that they are closely related to *Desulfovibrio*. The *nif* operon is flanked by typical cyanobacterial genes suggesting that it is an integral part of the *M. chthonoplastes* genome. Here we provide evidence that the *nif* operon of *M. chthonoplastes* is acquired through horizontal gene transfer. Moreover, the presence of the same *nif*-cluster in *M. chthonoplastes* isolates derived from various sites around the world suggests that this horizontal gene transfer event must have occurred early in the evolution of *M. chthonoplastes*. We have been unable to express nitrogenase in cultures of *M. chthonoplastes*, but we show that these genes were expressed under natural conditions in the field.
Introduction

Many cyanobacterial mats are formed by the cosmopolitan filamentous non-heterocystous cyanobacterium *Microcoleus chthonoplastes* (Stal, 2000). This species is easily recognizable by its occurrence in bundles of trichomes, enveloped by a common polysaccharide sheath, although this property may be lost in culture, which has sometimes led to the misidentification of cultured isolates of morphological similar opportunists (Garcia-Pichel et al., 1996; Siegesmund et al., 2008).

The fixation of atmospheric dinitrogen is a property of all cyanobacterial mats that have been investigated for it (Severin & Stal, 2010). The first isolate of an aerobic dinitrogen-fixing filamentous non-heterocystous cyanobacterium was originally assigned to *M. chthonoplastes* (Malin & Pearson, 1988; Pearson et al., 1979), but was later re-assigned to *Symplocca* sp. (Janson et al., 1998). Similar, *M. chthonoplastes* ‘strain 11’, isolated from a diazotrophic microbial mat from the Wadden Sea (Southern North Sea), was reported to be capable of anaerobic dinitrogen fixation (Stal & Krumbein, 1985) but was also re-assigned, in this case to the genus *Geitlerinema* (Garcia-Pichel et al., 1996; Siegesmund et al., 2008), a genus to which also the famous anoxygenic phototrophic Solar Lake strain ‘*Oscillatoria limnetica*’ belongs that can fix dinitrogen anaerobically. Other reports of dinitrogen fixing *M. chthonoplastes* and a *Microcoleus* sp. were by Dubinin et al. (1992) and Sroga (1997), respectively, but it is unclear whether their assignments were correct.

Pure culture studies with different true *M. chthonoplastes* strains did not reveal nitrogenase activity or cyanobacterial *nifH* gene expression. Rippka et al. (1979) were unable to detect nitrogenase activity in the type strain of *M. chthonoplastases* PCC7420, not even under strictly anaerobic conditions. Villbrandt & Stal (unpublished results) confirmed this result and also showed that none of the ‘true’ *M. chthonoplastases* of the collection of Garcia-Pichel et al. (1996) possessed nitrogenase activity, even when assayed under anaerobic conditions (according to the method described in Rippka & Waterbury, 1977). Moreover, Steppe et al. (1996) were unable to detect cyanobacterial *nifH*, the gene coding for dinitrogenase reductase, in 4 strains of *M. chthonoplastases*, including the type strain PCC7420. Instead, they amplified a non-cyanobacterial *nifH* from these cultures which they attributed to contaminants. The amplified *nifH* genes were closely related to *Gammaproteobacteria*, *Clostridium pasteurianum* and *Desulfovibrio gigas*, sequences that were also found in microbial mats dominated by *M. chthonoplastases*. Studies in *M. chthonoplastases* dominated microbial mats revealed similar results, amplifying *Gamma*- or *Deltaproteobacterial nifH* sequences rather than cyanobacterial *nifH* genes (Zehr et al., 1995; Olson et al., 1999; Omorogie et al., 2004a). In contrast, mats containing another filamentous non-heterocystous cyanobacterium, *Lyngbya* sp., did reveal cyanobacterial *nifH* sequences, showing that it was not impossible to retrieve cyanobacterial *nifH* sequences from environmental samples. These results were generally interpreted as that *M. chthonoplastases* is not capable of fixing dinitrogen because it lacked the required nitrogen fixation genes. Moreover, finding bacterial *nifH* genes in *M. chthonoplastases* dominated mats rather than cyanobacterial genes was dedicated to the presence of diazotrophic heterotrophic bacteria living in a consortium with *M. chthonoplastases*. It was
also concluded that *Cyanobacteria* might be less important for dinitrogen fixation in these mats. In another marine *M. chthonoplastes* mat it was shown, by combining metabolic inhibitor studies with the analyses of mRNA by reverse transcriptase PCR (RT-PCR), that sulfate-reducing bacteria were potentially important dinitrogen fixers and that most of the *nifH* genes expressed clustered with sulfate-reducing bacteria of the *Deltaproteobacteria* and other anaerobic bacteria (Steppe & Paerl, 2002). It has been suggested that these dinitrogen-fixing heterotrophic bacteria occur as epiphytes on the mucilaginous sheaths of *M. chthonoplastes* with which they presumably interact, and this interaction has been conceptualized as consortial dinitrogen fixation (Paerl et al., 2000; Steppe et al., 1996). Here, we report the existence of a complete nitrogenase operon in the genome of *M. chthonoplastes* PCC7420 that has been sequenced through the Moore Foundation Marine Microbiology project (http://www.moore.org/microgenome/).

**Material and Methods**

**Strains and growth conditions**

The strains used in this study were obtained from the Culture Collection Yerseke (CCY) and are listed in Table 1. All strains were grown in artificial seawater medium ASN3 (Rippka et al., 1979). The dinitrogen-fixing *Symploca* strains were grown in medium lacking combined nitrogen (NO\textsubscript{3}\textsuperscript{−}) (ASN3º). *Microcoleus* strains were grown at 23°C and a photon flux density of 70 µmol m\textsuperscript{−2} s\textsuperscript{−1} in an orbital shaker at 120 rpm and a light/dark cycle of 14 hours light and 10 hours dark. *Symploca* strains were grown in an orbital shaker at 120 rpm and 18°C, with a photon flux density of 20 µmol m\textsuperscript{−2} s\textsuperscript{−1} and a light/dark cycle of 16 hours light and 8 hours dark. Nitrogenase activity in the other strains was induced following the method of Rippka & Waterbury (1977). Briefly, a growing culture was transferred to medium devoid of combined nitrogen (ASN3º) and incubation was continued until the first sign of bleaching was observed. Bleaching is caused by the degradation of the phycobiliproteins and indicates the onset of nitrogen starvation. Nitrogen starved cultures were subsequently concentrated by centrifugation and re-suspended in 10 ml of ASN3º medium to which 10\textsuperscript{−5} M DCMU (3-(3,4-dichlorophenyl)-l,l-dimethyl urea) was added in order to prevent oxygen evolution by the inhibition of photosystem II. The suspension was flushed with dinitrogen in order to achieve anaerobic conditions. The 30 ml flask was sealed and 15% acetylene gas was added with a syringe as overpressure and the suspensions were incubated in front of 2 × 18 Watt daylight fluorescent tubes with a photon flux density of 50 to 60 µmol m\textsuperscript{−2} s\textsuperscript{−1}. Ethylene was assayed gas chromatographically (Stal, 1988) at regular intervals, starting 3 hours after establishing anaerobic conditions. After 24 hours the incubation was terminated and the cells were mixed immediately with RNAlater (Ambion Inc., Austin, TX, USA) following the manufacturer's instructions. These samples were subsequently used for extraction of RNA and RT-PCR of *nifHDK*. Natural samples were taken in the summer of 2007 from a *M. chthonoplastes* containing microbial mat located
Nitrogen fixation in *Microcoleus chthonoplastes* at the North Sea beach of Schiermonnikoog, The Netherlands (Severin & Stal, 2008). These samples were flash frozen in liquid nitrogen and kept at -80°C until use.

<table>
<thead>
<tr>
<th>Species</th>
<th>CCY strain ID</th>
<th>Place of Isolation</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Symphocca atlantica</em> PCC 8002</td>
<td>CCY9617</td>
<td>UK</td>
</tr>
<tr>
<td><em>Symphocca</em> sp.</td>
<td>CCY0030</td>
<td>Rottnest Island, Australia</td>
</tr>
<tr>
<td><em>Geitlerinema sp</em></td>
<td>CCY9619</td>
<td>Mellum, Germany</td>
</tr>
<tr>
<td><em>Microcoleus chthonoplastes</em> PCC7420*</td>
<td>CCY9604</td>
<td>Woods Hole, USA</td>
</tr>
<tr>
<td><em>Microcoleus</em> sp. *</td>
<td>CCY0002</td>
<td>Schiermonnikoog, Netherlands</td>
</tr>
<tr>
<td><em>Microcoleus</em> sp. *</td>
<td>CCY9602</td>
<td>Chile.</td>
</tr>
<tr>
<td><em>Microcoleus</em> sp. *</td>
<td>CCY9603</td>
<td>Solar Lake, Egypt</td>
</tr>
<tr>
<td><em>Microcoleus</em> sp. *</td>
<td>CCY9605</td>
<td>Mellum, Germany</td>
</tr>
<tr>
<td><em>Microcoleus</em> sp. *</td>
<td>CCY9606</td>
<td>St. Peter-Ording, Germany</td>
</tr>
<tr>
<td><em>Microcoleus</em> sp. *</td>
<td>CCY9707</td>
<td>North Carolina, USA</td>
</tr>
<tr>
<td><em>Microcoleus</em> sp. *</td>
<td>CCY9608</td>
<td>Wismar, Germany</td>
</tr>
<tr>
<td><em>Geitlerinema sp.</em></td>
<td>CCY9412</td>
<td>Schiermonnikoog, Netherlands</td>
</tr>
<tr>
<td><em>Geitlerinema sp.</em></td>
<td>CCY0102</td>
<td>Krim, Russia.</td>
</tr>
<tr>
<td><em>Microcoleus</em> sp. *</td>
<td>CCY9602</td>
<td>Negev Desert, Israel</td>
</tr>
</tbody>
</table>

* Strains proposed to be renamed to *Coleofasculus* according to Siegesmund et al. (2008)

**Isolation of nucleic acids**

DNA was isolated from cultures and natural mixed mat samples using the UltraClean™ Soil DNA Isolation Kit (MoBio laboratories, inc. Carlsbad, CA USA) according to the manufacturer’s instructions for maximal yields. RNA was extracted from cultures and natural mixed mat samples using the RNeasy Mini Kit (Qiagen Inc., Valencia, CA, USA) following the manufacturer’s instructions. RNA samples were treated with DNase I (Invitrogen Corp., Carlsbad, CA, USA) to remove residual DNA and cDNA was generated using Super-Script II reverse transcriptase (Invitrogen Corp.) following the manufacturer’s instructions and using random hexamer primers.

**PCR amplification, fragment cloning and DNA sequencing**

Oligonucleotide primers designed and used in this study are listed in Table 2. Twenty five μl of PCR mixture consisted of 200 μM of dNTP’s (Roche Applied Science, Indianapolis, IN, USA), 200 nM of primers, 1x HotStarPCR buffer (Qiagen Inc.), 10% v/v of dimethyl sulfoxide (Sigma-Aldrich, Munich, Germany), 0.2 mg/ml of bovine serum albumin (Fermentas, Hanover, MD, USA) and 0.6 units of HotStarTaq DNA polymerase (Qiagen Inc.). The reactions were run on a thermal cycler (Thermal Cycler 2720, Applied Biosystem,
Foster City, CA, USA). For the MC-nifHDK-primers the following program was used: 15 min at 94ºC; 35 cycles of 30 s at 94ºC, 30 s at 48ºC, 60 s at 72ºC; and a final extension step for 7 min at 72 ºC. For the amplification of the 16S rRNA gene, primers B27F and U1492R (Table 2) were used with the following program: 15 min at 94ºC; 35 cycles of 30 s at 94ºC, 30 s at 55ºC, 110 s at 72ºC and a final extension step for 7 min at 72ºC. PCR products were separated by electrophoresis on a 1% w/v agarose (Sigma Aldrich) gel and stained with SYBR Gold (Invitrogen Corp.). Amplicon size was estimated by comparison with a MassRuler DNA Ladder (Fermentas). Amplicons were cloned using the TOPO-TA cloning Kit (Invitrogen Corp.) following the manufacturer’s instructions. White colonies were selected and suspended in 10 μl of sterile MilliQ water, boiled for 10 min and used as template in a PCR with the vector primers T7 and T3 to amplify the inserted gene fragments. The PCR mixture (25 μl) consisted of 200 μM of dNTP’s (Roche), 200 nM of primers, standard Taq buffer (1x, New England BioLabs Inc., Ipswich, MA, USA), and 0.6 units of Taq DNA polymerase (New England Biolabs Inc.). The PCR reaction was run at: 95ºC for 2 min; 35 cycles 30 s at 94ºC, 30 s at 55ºC, 60 s at 72ºC; and finally 10 min at 72ºC. Amplicons containing insert DNA of the appropriate size were purified using Sephadex G-50 Superfine (Sigma-Aldrich) and DNA concentrations were determined spectrophotometrically. Amplicons were sequenced using the BigDye Terminator chemistry (Big Dye Terminator v3.1 Cycle Sequencing Kit, Applied Biosysytem) according to the manufacturer’s instructions. The sequence products were analyzed with a 3130 Genetic Analyzer (Applied Biosystem).

Table 2. List of oligonucleotides used in this paper

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence 5’-- 3’</th>
<th>Gene</th>
<th>Position within gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>MC-nifH-f</td>
<td>TTTACGGTAAAGGTGGAATCG</td>
<td>nifH</td>
<td>17-37</td>
</tr>
<tr>
<td>MC-nifH-r</td>
<td>TGCAATAATCAAACCGCTAAACG</td>
<td>nifH</td>
<td>550-529</td>
</tr>
<tr>
<td>MC-nifD-f</td>
<td>CAAGCCGTTCAGGAAGGCTA</td>
<td>nifD</td>
<td>424-443</td>
</tr>
<tr>
<td>MC-nifD-r</td>
<td>GAGCGGTGACACATGACCAA</td>
<td>nifD</td>
<td>873-854</td>
</tr>
<tr>
<td>MC-nifK-f</td>
<td>GTCAGCCTATTGCTGTCAATG</td>
<td>nifK</td>
<td>68-87</td>
</tr>
<tr>
<td>MC-nifK-r</td>
<td>CAGAACAGAGACGGAACCTT</td>
<td>nifK</td>
<td>811-792</td>
</tr>
<tr>
<td>MC-nifH-QRTr</td>
<td>ATCTAACACACTTTTGGTGTC</td>
<td>nifH</td>
<td>153-174</td>
</tr>
<tr>
<td>B27F</td>
<td>AGAGTTTGATCMTGGCTGAG</td>
<td>Bact - 16S</td>
<td>8-27</td>
</tr>
<tr>
<td>U1492R</td>
<td>GGTTACCTGTGTACGCCCT</td>
<td>Univ -16S</td>
<td>1510-1492</td>
</tr>
</tbody>
</table>

MC oligonucleotides are identical to their matching sequences in the targeted genes of *Microcoleus chthonoplastes* PCC7420
Gene expression analysis by qRT-PCR (quantitative reverse transcriptase PCR)

Based on the nifH sequence information of *M. chthonoplastes* CCY0002, a reverse primer (MC-nifH-QRTr; Table 2) was designed manually that could be combined with the MC-nifH-f forward primer in a qRT-PCR reaction. The specificity of this primer pair was checked in silico as well as by PCR amplification, cloning and sequencing 48 amplification products from environmental samples. All sequenced amplicons were 100% identical to part of the nifH gene of *M. chthonoplastes* CCY0002 confirming the specificity of the chosen primer set for qRT-PCR analysis. Quantitative RT-PCR was run on a Corbett Rotor-Gene 6000™ (Corbett Life Science, NSW 2137 Sydney, Australia). *M. chthonoplastes* CCY0002 was isolated from the microbial mats of Schiermonnikoog from which the environmental DNA was obtained. Environmental DNA as well as reverse-transcribed cDNA reversed transcribed from environmental RNA was used as a template. Cycling conditions were the following: 15 min at 95°C followed by 40 cycles of 15 s at 95°C, 20 s at 55°C and 20 s at 72°C. The reaction was directly followed by a melting curve analysis (72°C – 95°C). Environmental samples were run in triplicate and each run included two independently dilution series of linearized plasmid containing the target nifH insert as standard curves as well as non-template and non-target controls. The 15 μl qPCR reactions consisted of 7.5 μl of a commercially available reaction mixture (Absolute™ QPCR SYBR® Mix, Thermo Fisher Scientific, Waltham, MA, USA) containing all the components necessary to perform quantitative PCR, with the exception of template and primers, 333 nM of the forward and reverse primer, 5.5 μl of MilliQ water and 1 μl of template. The standard curves obtained from the linearized nifH containing plasmids allowed us to determine the number of gene copies and the number of transcripts per ng of extracted nucleic acid and expression levels were calculated as number of transcripts per number of nifH copies. The number of nifH copies in the standard curve ranged from $2.2 \times 10^4$ to $2.2 \times 10^7$ per reaction. Comparison of the ct values of the transcripts and nifH gene copies with that of the standard curve by least square linear regression analyses allowed us to quantify the original amount of target RNA or DNA molecules in the sample. Efficiencies for all reactions were 85 ± 2 %. NifH expression levels were compared to a 24h cycle of nitrogenase activity.

Nitrogenase activity

Nitrogenase activity was measured in pure cultures and microbial mat samples using the acetylene reduction assay (ARA) and normalized per milligram of chlorophyll-a as previously described (Severin & Stal, 2008). Light response curves of nitrogenase activity and daily dinitrogen fixation were calculated from ethylene production rates in the microbial mat sample as previously described (Severin & Stal, 2008).
DNA sequence and phylogenetic analysis

The Molecular Evolutionary Genetics Analysis (MEGA 4.0) software package (Tamura et al., 2007) was used to analyze and correct ABI trace files and to conduct phylogenetic analysis. DNA and protein sequences were aligned using the ClustalW module of MEGA 4.0 using the default settings and phylogenetic trees were constructed by applying the neighbor joining (NJ) method with a 1000 replicates bootstrap analysis. DNA sequences described in this study were deposited in GenBank under accession numbers GQ397255-GQ397274 for the nifH, nifD and nifK gene sequences and GQ402014-GQ402026 for the 16S rRNA gene sequences.

Results

Genetic analysis

In June 2007 the partial sequenced and automatically annotated genome of M. chthonoplastes PCC 7420 became publicly available at https://moore.jcvi.org/research. Upon analysis of contig #1103659003591 we discovered a complete nif gene cluster. The nif gene cluster (Figure 1A) contains the genes encoding the structural proteins (nifHDK) and genes for biosynthesis of the iron-molybdenum cofactor (FeMo-co; nifB’SU and nifENB). In addition, the nif cluster contains two genes proposed to be involved in nitrogen regulation, P-II and P-II’, (Martin & Reinhold-Hurek, 2002) and a gene encoding a ferredoxin protein (fd) mediating electron transfer (Rubio & Ludden, 2008). Two copies of nifB were found of which the first appears to be a partial and possibly non-functional gene (nifB’). Further downstream of the nif cluster, genes were found that encode binding- and transport proteins involved in the uptake of the molybdate co-factor (mopl, modDABC) (Figure 1B).
Nitrogen fixation in *Microcoleus chthonoplastes*

**A**

**AV**

![Diagram A AV]

**AV'**

![Diagram A AV']

**CY**

![Diagram A CY]

**CT**

![Diagram A CT]

**SF**

![Diagram A SF]

**DV**

![Diagram A DV]

**MC**

![Diagram A MC]

**B**

![Diagram B]

**Figure 1.** Comparative analysis of *nif*-cluster organization in different organisms. (A) Cyanobacterial genes are presented as white arrows and *Deltaproteobacterial*/*Chlorobia* type *nif*-genes are presented as black arrows. AV; *A. variabilis* ATCC29413, AV'; *A. variabilis* ATCC29413 (anaerobic *nif* cluster including insertion element plus integrase), CY; Cyanothece PCC7426, CT *Chlorobium tepidum*, SF; *S. fumaroxidans* MPOB, DV; *D. vulgaris* DP4, MC; *M. chthonoplastes* PCC 7420. (B) G+C content of part of contig #1103659003591 of *M. chthonoplastes* PCC7420; revealing the depressions in G+C content (vertical arrows) flanking the predicted horizontally acquired *nif*-gene cluster (black arrows)
Phylogenetic analysis of the nif genes revealed that only the initial three genes (nifB’US) are related to cyanobacterial homologs (Figure 2A). Phylogenetic analysis of nifHDK, P-II, P-II’, fd and nifENB (hereafter called the nifHDKENB cluster), revealed clustering with nif genes from the family of Chlorobiaceae and the Deltaproteobacteria families of Syntrophobacteraceae and especially with the Desulfovibrionaceae (Figure 2B). The M. chthonoplastes nifH gene contains a 6 nucleotide deletion resulting in the absence of 2 amino acids after position 64 in the amino acid sequence, a feature not shared by other cyanobacterial nifH genes but common in deltaproteobacterial nifH. Despite clustering in the Deltaproteobacteria group, nifH has highest identity to Ava_4046, one of the five nifH genes of Anabaena variabilis ATCC 29413. This nifH is the only nif gene in A. variabilis that clusters with the Deltaproteobacteria/Chlorobia group. The other four nifH genes of A. variabilis cluster with cyanobacterial homologs. To exclude the possibility that this non-typical cyanobacterial gene clustering is caused by a contamination in the original DNA sample used for the genome sequencing, we studied genes directly flanking the nif cluster and found that they are typical cyanobacterial. Moreover, we PCR amplified DNA fragments of the nifH, nifD and nifK genes from M. chthonoplastes PCC 7420, that is kept axenic in our CCY culture collection, and from 8 other axenic strains of M. chthonoplastes originating from different regions around the world (Table 1), using specific primers designed on the basis of the sequences of the M. chthonoplastes genome. The resulting DNA fragments were of the expected length and their sequence revealed high similarity to the genome sequences of M. chthonoplastes PCC7420. Neighbor joining trees based on the nucleotide sequences of the three nif genes from the CCY strains showed good congruency with their respective 16S rRNA genes (Figure 3). At the amino acid sequence level the identity was 98%, 91% and 91% for NifH, NifD and NifK respectively. Moreover, most mutations consist of neutral substitutions not causing an amino acid change. Using these M. chthonoplastes specific nifHDK primers we failed to obtain products from Symploca or Geitlerinema strains. Since we can now exclude contamination of the genome sequence, horizontal gene transfer of the nifHDKENB cluster from a Deltaproteobacteria/Chlorobi donor is the most likely explanation of this observation.
A Figure 2A. Phylogenetic analysis of NifU of *M. chthonoplastes* PCC7420, revealing its cyanobacterial descent. The neighbor joining tree was constructed using bootstrap analysis with 1000 iterations. Shading indicates the phylogenetic Class wherein NifU of *M. chthonoplastes* (bold and underlined) clusters.
Figure 2B. Phylogenetic analysis of NifH of *M. chthonoplastes* PCC7420, revealing its non-cyanobacterial descent. The neighbor joining tree was constructed using bootstrap analysis with 1000 iterations. Shading indicates the phylogenetic Class wherein NifH of *M. chthonoplastes* (bold and underlined) clusters.
Figure 3. Phylogenetic analysis of 16S, nifH, nifD and nifK of CCY strains reveals congruency in tree topology. Neighbor joining trees were generated using maximum likelihood algorithm with 1000 bootstrap iterations (bootstrap values are shown at nodes). nifH copyDNA (cDNA) obtained by RT-PCR from a natural microbial mat is included in the nifH tree.
Evidence for horizontal gene transfer

Analysis of average G+C content revealed two G+C poor stretches (29.5% and 34.3% G+C) exactly flanking the predicted horizontally acquired $nif$ cluster (Figure 1B). The average G+C content of the $nifHDKENB$ cluster is in the same range (47.1% G+C) as that of the whole contig (45.1% G+C). Codon usage frequencies for each gene in the contig was calculated and compared by cluster analysis using a correlation matrix in the statistical software package Past (Hammer et al., 2001). The resulting dendrogram (Figure 4) shows that based on their codon usage, the $nifHDKENB$ genes form a distinct cluster, suggesting that their codon usage is more similar to each other than to that of the other genes in the same contig. Putative promoter regions are not well resolved for $M$. chthonoplastes and no obvious deviations were found from the promoter regions of the typical cyanobacterial genes in the contig (data not shown).

![Figure 4](image-url)

**Figure 4.** Cluster analysis dendrogram of codon usage frequencies of the genes within contig #1103659003591 of $M$. chthonoplastes PCC7420. Genes form clusters when codon usage frequencies are similar. The genes of the horizontally transferred $nif$-gene cluster are in bold. The other open reading frames in the contig are arbitrary numbered.
Expression of nitrogenase

Eight strains of *M. chthonoplastes* (Table 1) were cultivated and subsequently starved for nitrogen. In so-called anaerobically dinitrogen-fixing *Cyanobacteria*, nitrogenase is normally induced when nitrogen-deprived cultures are incubated under anaerobic conditions while oxygenic photosynthesis is blocked through inhibition of PS-II by DCMU (Lavergne, 1982). However, none of the eight *M. chthonoplastes* strains showed any nitrogenase activity after 24 hours of incubation. In addition, gene expression studied by RT-PCR analysis revealed no synthesis of *nifH*, *nifK* or *nifD* specific mRNA in any of the strains.

In contrast to the pure cultures, gene expression was observed in environmental samples of the *M. chthonoplastes* dominated microbial mat from which strain CCY0002 was originally isolated. Twelve samples were taken from a 24 hr sampling campaign from which RNA was extracted. RT-PCR using the *M. chthonoplastes* specific *nifH* primers revealed positive bands for each sample. Bands were of the expected size and sequence analysis on two clones revealed 100% identity with the *nifH* gene from strain CCY0002 (Figure 3, cDNA 26-5-19 and cDNA 26-5-4). Quantitative RT-PCR revealed that the number of *nifH* transcripts per copy varied between the time points (Figure 5; vertical bars). In general, higher ratios were observed at daytime and ranged between virtually zero to 0.06 *nifH* transcripts / copy. The pattern was in agreement with the daily variations of nitrogenase activity. Peaks in nitrogenase activity were observed throughout the day but were higher during the day than at night time reaching 0.23 and 0.13 µmol C₂H₄ per mg chlorophyll per hour, respectively.

**Figure 5.** Natural 24h cycle of nitrogenase activity (filled circles) and *nifH* gene expression (columns) measured in microbial mat samples. The standard deviation of three independent experiments is indicated by error bars. The horizontal bar at the top of the graph depicts the natural light (white bar) and dark (black bar) period of the 24 h cycle.
Discussion

Analysis of nifH sequences is widely used as measure for the potential of a system to fix dinitrogen and for the diversity of diazotrophic organisms in that ecosystem. The assumption is that the phylogeny of a gene can predict the phylogeny of its host. This would be true if all genes are only transferred to its successor by linear descent. However, there are many cases known in which DNA is acquired from other species via horizontal gene transfer and subsequently becomes fixed in the genome of a host-organism via adaptive evolution. The acquisition of foreign DNA may occur via specific DNA uptake systems, the active uptake of free extracellular DNA, via conjugation with other microorganisms followed by active transfer of DNA or after transduction via phages (Koonin et al., 2001). Horizontal gene transfer of DNA or RNA may occur between unrelated organisms and has been shown in several members of every domain of life.

For the study of the diversity of nifH, often the degenerated primers are used that were developed by Zehr & McReynolds (1989). These primers were initially designed to target nifH genes from marine dinitrogen-fixing organisms, particularly Cyanobacteria but in several studies using these degenerated many non-cyanobacterial nifH genes are targeted. Using these primers, M. chthonoplastes dominated microbial mats appear to be low in cyanobacterial specific nifH genes. As described in the introduction, this observation, in combination with the failure to detect nitrogenase activity in pure cultures of M. chthonoplastes led to the general assumption that this organism lacked the genes for dinitrogen fixation and, hence, lacked the ability to fix dinitrogen (de Wit et al., 2005; Steppe & Paerl, 2002; Zehr et al., 1995)

Our results however show that all strains of M. chthonoplastes tested by us contained nifH, nifD and nifK and the genome sequence of M. chthonoplastes PCC7420 shows that a complete nif-gene cluster is present (Rubio & Ludden, 2008). Since M. chthonoplastes forms large bundles of filaments enclosed in a thick extracellular polysaccharide sheath it was difficult to grow this organism axenically. Other prokaryotes have been found tightly associated with or even within the polysaccharide sheath, including members of the family of Desulfovibrionaceae (D’Amelio et al., 1987). A putative contamination of the DNA used for genome sequencing can therefore not be excluded. However, since the same sequence was found in eight different strains of M. chthonoplastes isolated from various regions around the world and were brought into pure culture by different laboratories, the chance of catching the same contamination in all strains is highly unlikely. Therefore, the most likely explanation is that M. chthonoplastes acquired the nif gene cluster via horizontal gene transfer. This is substantiated by the following data. Phylogenetic analysis shows that the genes of the nifHDKENB cluster all group with Deltaproteobacteria rather than with Cyanobacteria (Figure 2). The nifHDKENB cluster is flanked by two steep depressions in average G+C content which is often found to be involved in DNA integration (Kleckner, 1990). Possible, here an A+T rich sequence functioned as site of integration for the horizontally acquired DNA fragment. In fact, given the presence of the typical cyanobacterial nifB’SU genes directly upstream of the acquired nifHDKENB genes it is likely that a putative original nif-gene cluster may have been replaced by the new nif-cluster. Such a replacement could be mediated by insertion sequences. However, except for the
Nitrogen fixation in *Microcoleus chthonoplastes*

A+T rich sequence, no remnants of an IS-element transfer was detected. The fact that the insertion itself does not deviate in G+C content suggests that either the donor organism had a similar G+C content or that the G+C content of DNA insert adapted to that of the host organism to ensure smooth replication and transcription. Also the codon usage can be a marker of putative horizontal gene transfer events. Overall the average codon usage is species or even strain specific and is regarded as an adaptation to its pool of tRNA molecules (Ikemura, 1985). The preference for certain codons over others encoding the same amino acid in a protein depends on the relative expression profiles of the different tRNA’s giving rise to optimal protein synthesis machinery. In the case of the *Microcoleus nifHDKENB* cluster the codon usage of the genes in the suspected DNA insert differs from that of the flanking genes in the contig (Figure 4) which is in support of acquisition of the these genes through horizontal gene transfer.

The filamentous cyanobacterium *A. variabilis* contains 5 NifH homologs, one of which (Ava_4046) is at the amino acid level 72% identical to NifH from *M. chthonoplastes* and also clusters in the *Deltaproteobacteria*/Chlorobia group. However, this particular nifH is not part of a nif gene cluster in *A. variabilis*. This organism possesses three nif clusters, one of which is interrupted by an insertion sequence including an integrase (Brusca et al., 1989). Under anaerobic conditions this insertion sequence is excised, leaving an intact nif operon that functions in vegetative cells (Thiel et al., 1995). The third nif cluster in *A. variabilis* encodes an alternative, vanadium dependent, nitrogenase that functions under aerobic conditions in the heterocyst when molybdate is unavailable. Otherwise the regular molybdate nitrogenase is expressed in heterocysts. Possibly, nifH (Ava_4046) of *A. variabilis* is also obtained through horizontal gene transfer but its function in this organism is not known. We did not encounter any other putative nif genes in the genome of *M. chthonoplastes*.

One of the reasons why *M. chthonoplastes* thus far has been tagged as a non-diazotroph is that it does not grow without combined nitrogen and it lacks nitrogenase activity in pure isolates in the lab, even under fully anaerobic conditions. In this assay, *Cyanobacteria* that possess the genetic capacity of synthesizing nitrogenase but are unable to provide anaerobic conditions will express nitrogenase (Rippka & Waterbury, 1977). Although a highly identical nitrogenase cluster was found in all *M. chthonoplastes* strains in our culture collection, none showed nitrogenase activity or expressed any of the three structural nif genes. In contrast, we demonstrated the expression of nifH, nifD and nifK in natural samples taken from a *M. chthonoplastes* dominated marine microbial mat using primers that were specifically designed for the *M. chthonoplastes* strain that was isolated from that mat. It is possible that after maintaining the cultures of *M. chthonoplastes* for many years in the lab growing on combined nitrogen that it might have lost the capacity of expressing nitrogenase. However, this is not the case with many other species of *Cyanobacteria* that express nitrogenase only under anaerobic conditions. Another possibility is that *M. chthonoplastes* need different conditions in order to express nitrogenase. For instance, the inhibitor DCMU may have other negative side effects in this organism that prevents it from expressing nitrogenase. Sulfide or far red light are other means for inhibiting photosystem II while allowing photosystem I activity (Jørgensen et al., 1986). It is also possible that regulation of nitrogenase activity in *M. chthonoplastes* is
completely different from that in other *Cyanobacteria* due to the presence of the PII genes rather than *nifX* and *nifW* typically found in *Cyanobacteria*. Alternatively, nitrogenase expression might even depend on external signals from closely associated sulfate reducing bacteria to be activated. Quantitative RT-PCR revealed that *M. chthonoplastes* specific *nifH* expression in the microbial mat was highest during day-time and declined during night-time. Nitrogenase activity as measured by the acetylene reduction test reveals a similar pattern suggesting that the expressed genes indeed result in the translation into active nitrogenase proteins. Many non-heterocystous diazotrophic *Cyanobacteria* separate dinitrogen fixation temporally from oxygenic photosynthesis and confine the former to the night time. This was not obvious in this *M. chthonoplastes* mat which showed activity both during the day and night. Possibly, dinitrogen fixation occurs only at greater depths in the mat where the sulfide concentrations are higher and only far red light penetrates (Jørgensen et al., 1987). We have shown here that *M. chthonoplastes* contains a functional *nif*-gene cluster which is not typical cyanobacterial and which is most likely acquired via horizontal gene transfer from a member of the *Deltaproteobacteria*. Therefore several previous studies may have misinterpreted the observed microbial diversity and underestimated cyanobacterial dinitrogen fixation and overestimation the role of other bacteria in dinitrogen fixation in cyanobacterial mats (Wawer et al., 1997).

Members of the family of *Desulfovibrionaceae*, such as *D. vulgaris* or *D. gigas* are the most likely candidate donor organisms for the *nif*-genes found in *M. chthonoplastes*. First of all the gene and protein sequences of the *M. chthonoplastes* *nif*-cluster showed highest identity with that of the *Desulfovibrionaceae* and secondly these sulfate reducing bacteria are frequently found in *M. chthonoplastes* dominated microbial mats (Sigalevich et al., 2000; Zehr et al., 1995). Moreover, members of the *Desulfovibrionaceae* are known to be capable of transducing DNA fragments via phages (Rapp & Wall, 1987). Finding the highly conserved *nif* genes amongst *M. chthonoplastes* strains isolated from various locations around the world suggest that these genes were obtained early in the evolution of this species. 16S rRNA analysis of the strains used in this study revealed a clade consisting of species related to *Geitlerinema* sp, a clade with species related to *Symplca* sp. and one clade with species closely related to the *M. chthonoplastes* type strain (Fig 3). Recently, Siegesmund et al. (2008) studied 16S rRNA genes and ribosomal ITS sequences of a large number of *Cyanobacteria* originally assigned to *Microcoleus* and related genera were studied. They found that *Microcoleus* fell into two clades, one with taxa belonging to the *Oscillatoriaceae* and the other to the *Phormidiaceae*. Because the terrestrial type strain *M. vaginatus* for this genus belongs to the *Oscillatoriaceae*, a new genus *Coleofasciculus* was proposed for the strains belonging to the *Phormidiaceae* (Siegesmund et al., 2008). According to the proposed new nomenclature, the strains used in this study that are closely related to *M. chthonoplastes* belong to this new genus of *Coleofasciculus* (see table 1).

Congruency between the 16S rRNA gene tree with the *nifH*, *nifD* and *nifK* trees (Figure 3) of the CCY strains are in agreement with an early horizontally acquisition of the *nif*-gene cluster followed by strictly vertically inheritance and it is therefore tempting to speculate
that this horizontal gene transfer event may have formed the basis for the speciation of
*M. chthonoplastes*.

**Acknowledgements**

We thank Wolfgang Hess and Jan Mitschke of the University of Freiburg, Germany for their help in promoter search. Publication 4589 Netherlands Institute of Ecology (NIOO-KNAW)