AEM
Journals.ASM.org

From Green to Red: Horizontal Gene Transfer of the Phycoerythrin Gene Cluster between *Planktothrix* Strains

Ave Tooming-Klunderud,[a] Hanne Sogge,[a,b] Trine Ballestad Rounge,[a,c] Alexander J. Nederbragt,[a] Karin Lagesen,[a] Gernot Glöckner,[d,e] Paul K. Hayes,[f] Thomas Rohrlack,[g,h] Kjetill S. Jakobsen[a,b]

University of Oslo, Centre for Ecological and Evolutionary Synthesis (CEES), Department of Biosciences, Oslo, Norway[a]; University of Oslo, Microbial Evolution Research Group (MERG), Department of Biosciences, Oslo, Norway[b]; Cancer Registry of Norway, Oslo, Norway[c]; Institute for Biochemistry I, Medical Faculty, University of Cologne, Cologne, Germany[d]; Leibniz Institute of Freshwater Ecology and Inland Fisheries, IGB, Berlin, Germany[e]; Faculty of Science, University of Portsmouth, Portsmouth, United Kingdom[f]; NIVA, Norwegian Institute for Water Research, Oslo, Norway[g]; Norwegian University of Life Sciences, Department of Plant and Environmental Sciences, Ås, Norway[h]

Horizontal gene transfer is common in cyanobacteria, and transfer of large gene clusters may lead to acquisition of new functions and conceivably niche adaption. In the present study, we demonstrate that horizontal gene transfer between closely related *Planktothrix* strains can explain the production of the same oligopeptide isoforms by strains of different colors. Comparison of the genomes of eight *Planktothrix* strains revealed that strains producing the same oligopeptide isoforms are closely related, regardless of color. We have investigated genes involved in the synthesis of the photosynthetic pigments phycocyanin and phycoerythrin, which are responsible for green and red appearance, respectively. Sequence comparisons suggest the transfer of a functional phycoerythrin gene cluster generating a red phenotype in a strain that is otherwise more closely related to green strains. Our data show that the insertion of a DNA fragment containing the 19.7-kb phycoerythrin gene cluster has been facilitated by homologous recombination, also replacing a region of the phycocyanin operon. These findings demonstrate that large DNA fragments spanning entire functional gene clusters can be effectively transferred between closely related cyanobacterial strains and result in a changed phenotype. Further, the results shed new light on the discussion of the role of horizontal gene transfer in the sporadic distribution of large gene clusters in cyanobacteria, as well as the appearance of red and green strains.

Horizontal gene transfer (HGT), the exchange of genetic information between two organisms that do not share a recent ancestor-descendant relationship, is now recognized as a major force shaping the evolutionary history of prokaryotes (e.g., references 1 to 4). HGT is considered to be common in cyanobacteria (5). Through the availability of bacterial genome sequences, it has become clear that HGT can occur throughout the genome and that a substantial fraction of genes have been horizontally transferred (5, 6). The quantity of genetic material that can be horizontally transferred may range from small gene fragments (e.g., references 7 to 9) to fragments spanning complete genes (e.g., references 10 to 12) and whole operons encoding complex biochemical pathways (e.g., references 13 to 15). As even the transfer of a single or a few genes can give recipient organisms the opportunity to implement a new function and exploit new ecological niches, HGT contributes to the rapid creation of biological novelty that otherwise, through mutations and gene duplications, might have taken millions of years to appear.

According to Andam and coworkers (1), HGT is the norm and not the exception, while others call the transfer of genes between bacteria "both rare and promiscuous" (4). Successful HGT depends on transfer of genetic material to the cell (via transformation, conjugation, transduction, or gene transfer agents), survival of the DNA in the cell, integration of foreign DNA via recombination, and finally fixation of the integrated DNA in the population (involving, for example, selection). Since the rate of recombination decreases with increased sequence dissimilarity (16, 17), HGT events are more common among close relatives, as shown by a recent analysis of 657 sequenced prokaryotic genomes (18).

For fixation of a newly transferred gene in the population, it should provide a relevant function and this function must operate within the native machinery of the host cell. Since bacterial genomes are subject to deletional bias (19), genes that do not contribute to fitness of the organism will eventually be removed from the genome. Integration of new genes into existing cellular networks can be facilitated by acquisition of an operon containing all genes and regulatory regions required for function (20). For single-gene acquisitions, the fate of new genes depends largely upon the existing genes in the new host. Experimental studies have shown that most HGT events are deleterious (21, 22). However, rare HGT events and mutations can be selected for under particular conditions and thus contribute to bacterial adaptation and evolution (23–25).

Horizontal gene transfer events have also been demonstrated for the filamentous cyanobacterium *Planktothrix* (e.g., references 26 to 29), which occurs in deep and stratified lakes in temperate regions of the Northern Hemisphere. Traditionally, *Planktothrix* organisms isolated from different lakes have been classified into

Received 3 May 2013 **Accepted** 25 August 2013
Published ahead of print 30 August 2013
Address correspondence to Ave Tooming-Klunderud, ave.tooming-klunderud@ibv.uio.no, or Kjetill S. Jakobsen, kjetill.jakobsen@ibv.uio.no.
Supplemental material for this article may be found at http://dx.doi.org/10.1128/AEM.01455-13.
Copyright © 2013, American Society for Microbiology. All Rights Reserved.
doi:10.1128/AEM.01455-13

species according to morphological characteristics, such as cell dimension and pigmentation. Following the first description of the genus *Planktothrix* including 14 distinct species by Anagnostidis and Komárek (30), the number of different species has been heavily disputed. Studies based on molecular data, such as sequences of gas vesicle genes and 16S rRNA, have suggested that the whole *Planktothrix* genus is monospecific (31, 32), while Suda and coworkers (33) described four *Planktothrix* species based on several genetic and phenotypic properties.

Planktothrix strains isolated from Norwegian lakes and classified as distinct species at the Algal Culture Collection of the Norwegian Institute for Water Research (NIVA) cannot be separated by 16S rRNA. Recently, Rohrlack and coworkers (34, 35) reported that strains of *Planktothrix* showing >99% 16S rRNA gene sequence similarity may produce distinct cellular patterns of oligopeptides, bioactive secondary metabolites synthesized mostly by nonribosomal peptide synthetases. Using the oligopeptide profiles produced by each strain as markers, they grouped strains into distinct chemotypes (Cht). Based on field studies of the Norwegian Lake Steinsfjorden, four coexisting *Planktothrix* chemotypes differing considerably in seasonal dynamics, depth distribution, and participation in loss processes were identified (34). Since the production of oligopeptides is facilitated by several large and independently evolving operons (36, 37), strains associated with a distinct chemotype are assumed to be more closely related. This hypothesis is also supported by data showing that *Planktothrix* strains associated with the same chemotype generally have the same color, either red or green (35). However, in Lake Steinsfjorden, one chemotype was shown to comprise both red and green strains (34, 35). The red and green appearance of *Planktothrix* strains is associated with the content of accessory light-harvesting pigments, the phycobiliproteins, involved in the photoautotrophic machinery. Phycobilisomes, the macromolecular complexes formed from phycobiliproteins, have an allophycocyanin core that links to the photosystems and peripheral light-harvesting rods that comprise either phycocyanin (PC) or phycocyanin and phycoerythrin (PE) (for a review, see, e.g., reference 38). Phycocyanin, common to all cyanobacteria, imparts a green appearance to the cell and absorbs red light (620 to 630 nm). Phycoerythrin absorbs green light (560 to 570 nm) and imparts a dominant red color when present. The coexistence of red and green strains within the same chemotype can be explained by acquisition or loss of genes coding for phycoerythrin as suggested earlier for *Synechococcus* and other picocyanobacteria (39–41).

The aim of this study was to investigate the genome arrangements leading to the co-occurrence of red and green strains within the same oligopeptide chemotype. For that purpose, the genomes of eight different *Planktothrix* strains classified as four different species were sequenced, four red and four green strains, including one red and two green strains from the same chemotype. We address the following questions: (i) how similar are the genomes of closely related *Planktothrix* strains and is there any evidence for genetic substructuring according to color or chemotype; (ii) are the structure and chromosomal location of genes encoding phycocyanin and phycoerythrin pigments the same in all strains; (iii) in the light of results from the first two questions, can the co-occurrence of red and green strains within the same chemotype be explained by altered phycoerythrin genes and is this because of (a) an acquisition of the phycoerythrin gene cluster by the red strain or (b) mutations leading to nonfunctional phycoerythrin genes in two green strains.

Our results show that all eight *Planktothrix* genomes are highly similar and that strains associated with the same chemotype are the most closely related, regardless of color. Furthermore, we reveal that a red strain from a chemotype dominated by green strains has acquired the 19.7-kb phycoerythrin gene cluster. Our data indicate that the DNA fragment containing phycoerythrin operon originated from a strain associated with a "red" chemotype.

MATERIALS AND METHODS

***Planktothrix* strains and DNA isolation.** Eight *Planktothrix* strains isolated from lakes Steinsfjorden and Kolbotnvatnet (Norway) have been investigated. All strains have been kept in continuous, nonaxenic cultures at the Algal Culture Collection of the Norwegian Institute for Water Research (NIVA), in Z8 medium and light at a photon flux density of 10 μmol m^{-2} s^{-1}, and a light-dark cycle of 12:12 h. Prior to genomic DNA isolation, cells were centrifuged and resuspended in Tris-EDTA (TE) buffer. DNA was extracted by the following procedure: cells were treated with lysozyme (final concentration, 15 mg/ml), followed by RNase A and proteinase K (5 mg/ml, 1% lithium dodecyl sulfate [LDS]) treatment. Samples were then incubated at 60°C (shaking at 300 rpm) for 60 min. In cases where the solution did not clear after 60 min, the incubation time was prolonged with additional proteinase K. Subsequently, 1 volume phenol-chloroform-isoamyl alcohol (25:24:1, vol/vol) was added. The solution was mixed by inversion at 37°C for 30 min to remove pigments and proteins. After centrifugation, the upper layer was treated twice with 1 volume chloroform-isoamyl alcohol (24:1, vol/vol). DNA was precipitated using 0.1 volume 3 M sodium acetate and 2.5 volumes ice-cold 96% ethanol on ice for 1 h. The DNA pellet was washed twice with ice-cold 70% ethanol, dried at room temperature, and dissolved in Tris-HCl buffer (pH 8.0). All DNA samples were purified using Amicon Ultra-0.5 ml 50K centrifugal columns to ensure high-quality DNA for paired-end library preparation using the 454 Life Sciences protocols. DNA was concentrated according to the manufacturer's instructions and washed twice using Tris-HCl buffer, pH 8.0.

Sequencing. Seven out of eight *Planktothrix* genomes were sequenced using 454 pyrosequencing at the Norwegian Sequencing Centre (http://www.sequencing.uio.no/). Strain NIVA-CYA 34 was initially sequenced using Sanger sequencing and 454 pyrosequencing at the Max Planck Institute for Chemical Ecology. For Sanger sequencing, the DNA sample of NIVA-CYA 34 was amplified using the REPLI-g kit (Qiagen). The resulting DNA was randomly sheared, and the fragment size range from 2,500 to 3,000 kb was selected for cloning into pUC18 standard vectors. The resulting clones were sequenced from both ends with the standard sequencing primers on an ABI 3700 machine generating 20,066 sequencing reads comprising 16 Mb. An additional 454 shotgun library was prepared and sequenced at the Norwegian Sequencing Centre to ensure satisfactory quality of the NIVA-CYA 34 genome assembly. For the remaining seven strains, both shotgun and paired-end libraries were prepared and sequenced to 23- to 41-fold coverage (see Table S1 in the supplemental material for information about libraries and number of reads/bases sequenced for each strain).

Assembly and annotation. The newbler program (gsAssembler; Roche-454, Branford, CT, USA) was used to assemble the 454 reads into scaffolds and contigs (the newbler version used for assembly of each strain together with assembly statistics is shown in Table S1 in the supplemental material). Since all strains had been cultivated nonaxenically, reads from cocultured bacteria were present in the data set. To decrease the chance of coassembling these contaminating reads with the genome of the strain of interest, stringent overlap settings with a minimum of 98% overlap identity and a minimum of 60-bp overlap length were used. However, the assemblies of most strains still contained some short scaffolds with low average read depth (below 5- to 10-fold) and/or GC percentages that diverged from high-read-depth scaffolds. These scaffolds were considered

TABLE 1 *Planktothrix* strains investigated

Algal Culture Collection ID	*Planktothrix* species	Color of filaments	Chemotype	Yr of isolation	Source of isolate	Genome size (Mb)	No. of protein-coding genes	COG %	KEGG pathways %	Accession no.
NIVA-CYA 98	*P. prolifica*[a]	Red	Cht 1	1982	Lake Steinsfjorden	5.74	4,984	65.43	19.64	AVFZ00000000
NIVA-CYA 406	*P. prolifica*	Red	Cht 1	1998	Lake Steinsfjorden	5.76	5,088	64.97	19.10	AVFV00000000
NIVA-CYA 407	*P. rubescens* var.	Red	Cht 5	1998	Lake Steinsfjorden	5.51	4,846	66.22	19.62	AVFW00000000
NIVA-CYA 56/3	*P. mougeotii*	Green	Cht 7	1978	Lake Steinsfjorden	5.51	4,892	66.36	19.50	AVFY00000000
NIVA-CYA 34	*P. agardhii*	Green	Cht 9	1976	Lake Kolbotnvatnet	5.51	4,954	65.96	20.36	AVFT00000000
NIVA-CYA 405	*P. mougeotii*	Green	Cht 9	1998	Lake Steinsfjorden	5.5	4,878	66.05	20.62	AVFU00000000
NIVA-CYA 540	*P. prolifica*	Red	Cht 10[b]	2004	Lake Steinsfjorden	5.57	4,901	66.05	20.39	AVFX00000000
NIVA-CYA 15	*P. agardhii*	Green	Cht 14	1968	Lake Kolbotnvatnet	5.49	4,790	66.99	20.92	AVFS00000000

[a] Species designation has been changed; it was previously known as *Planktothrix rubescens*.
[b] Chemotype 10 is a subgroup of chemotype 9, differing in the cyanopeptolin isoform produced.

to be derived from cocultured bacteria present in the sample used for DNA extraction (42). The low-read-depth scaffolds were compared to the nonredundant NCBI protein database using BLASTX, and all noncyanobacterial matching scaffolds were removed before annotation.

Assembly of the NIVA-CYA 34 genome was done using both the Sanger and 454 reads. Newbler was given a trimming file (-vt option) to remove pUC18 plasmid vector sequences from the Sanger reads. The "-stopjoin" and "-repfill" options were used for assembly. One low-coverage noncyanobacterial scaffold was removed from the newbler assembly. Four of the unscaffolded contigs were identified as containing cyanobacterial sequences; these were added to the scaffolds before annotation.

According to classification of Genome Sequence Standards (43), we consider the assemblies to be improved high-quality drafts. Annotation of all genome sequences was performed using the U.S. Department of Energy (DOE) Joint Genome Institute (JGI) integrated microbial genomes database and comparative analysis system (IMG) (44).

Comparison of genomes. The IMG system was used for pairwise comparison of genomes and calculation of Pearson correlation coefficients of COG (Cluster of Orthologous Groups) profiles. Homologous genes were defined as genes having a minimum of 80% sequence identity and identified using BLASTP.

Hierarchical clustering by COG profiles was performed by the IMG system using the tool Cluster (http://www.falw.vu/~huik/cluster.htm). For construction of the pairwise hierarchical tree, a function profile vector (gene counts per COG) was generated for each genome. The distance metrics between these profile vectors were calculated by means of uncentered correlation. Pairwise single-linkage clustering was performed, first grouping the two closest profile vectors to form a group and then grouping the next pair of closest groups or vectors until all genomes were incorporated into the hierarchical tree.

Phylogenetic analyses of genomes were performed using a set of core genes. Core genes were defined as genes present in all genomes and having a minimum of 90% sequence identity, resulting in a data set of 3,690 genes. The final data set (after removing 100% identical genes in all genomes together with transposases, oligopeptide synthetases, retron-type reverse transcriptases, phage-associated proteins, and proteins shorter than 50 amino acids [aa]) contained 2,914 genes. Ten subsets of genes for generating phylogenies were created by random sampling of 20 genes from the core gene set (2,914 genes) repeated 10 times using R (R Development Core Team [http://www.R-project.org]). Sequences were aligned using MEGA 5 software (45) and concatenated for each of the 10 data sets. Lists of genes for each data set are shown in Table S2 in the supplemental material. ML analyses were carried out using RAxML (46) with 1,000 bootstrapped resamplings with the GTRCAT model. The resulting individual phylogenetic trees are shown in Fig. S1 in the supplemental material. Finally, the aligned sequences of all 200 genes were concatenated, resulting in a 217,546-bp-long alignment (available through http://dx.doi.org/10.6084/m9.figshare.719100), and ML analyses were carried out using RAxML (46) with 1,000 bootstrapped resamplings with the GTRCAT model.

Analyses of phycobilisome genes. All gene sequences analyzed (phycoerythrin, phycocyanin, and flanking genes) were downloaded from IMG. Sequences were aligned using MEGA 5 software (45). Nucleotide diversity was analyzed using the computer program DNA Sequence Polymorphism (DnaSP) (47). Maximum likelihood analyses of *cpc* genes were carried out using RAxML (46) with 1,000 bootstrapped resamplings. GTRCAT was determined to be the best evolutionary model using ModelTest (48).

Detection of recombination events. Recombination events were detected by visual analyses of informative sites (variable sites where each variant occurs in at least two sequences) as described by Rudi and coworkers (7). In order to detect the recombination breakpoints, concatenated sequences of *cpcBA* (reverse complement) and CHAP domain genes were analyzed using the RDP 4 software (49). Sequences from strains NIVA-CYA 406 and NIVA-CYA 15 were discarded as these are identical with sequences from strains NIVA-CYA 98 and NIVA-CYA 34, respectively. Recombination signals were accepted if at least three different methods detected statistically significant ($P < 0.05$) evidence of recombination.

Nucleotide sequence accession numbers. All *Planktothrix* genomes have been deposited as Whole Genome Shotgun projects at DDBJ/EMBL/GenBank under the accession numbers AVFS00000000 to AVFZ00000000 (see Table 1 for details). The versions described in this paper are versions AVFS01000000 to AVFZ01000000. The 454 sequencing data are available from the NCBI Sequence Read Archive (NIVA-CYA 98, SRP028838; NIVA-CYA 406, SRP028560; NIVA-CYA 407, SRP028561; NIVA-CYA 56/3, SRP028564; NIVA-CYA 34, SRP028840; NIVA-CYA 405, SRP028540; NIVA-CYA 540, SRP028566; NIVA-CYA 15, SRP028535). Annotated genomes are publicly available at IMG (https://img.jgi.doe.gov/cgi-bin/er/main.cgi).

RESULTS

Comparison of genomes reveals high similarity. Genomes of eight *Planktothrix* strains were sequenced: four red and four green strains, including a red (NIVA-CYA 540) and two green (NIVA-CYA 34 and NIVA-CYA 405) strains associated with Cht 9/10 (Table 1).

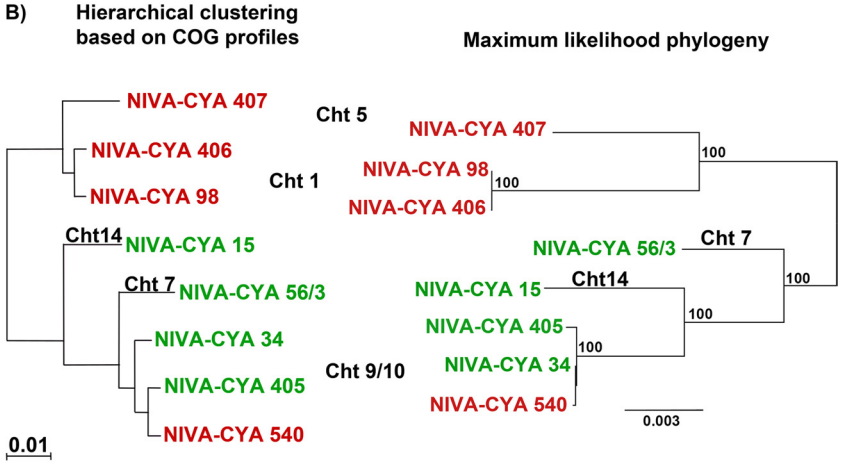
All eight genomes were of similar size, ranging from 5.49 Mb to 5.71 Mb, with a GC content of 39.5% (Table 1). The IMG annotation system predicted between 4,790 and 5,088 protein-coding genes in each genome (Table 1). Of the predicted protein-coding genes, 65 to 67% could be annotated using the COG functional annotation database, and 19 to 20% were annotated to a KEGG pathway (Table 1).

Analyses of gene content (both COG-annotated and all pro-

A)

	NIVA-CYA 98	NIVA-CYA 406	NIVA-CYA 407	NIVA-CYA 540	NIVA-CYA 405	NIVA-CYA 34	NIVA-CYA 15	NIVA-CYA 56/3
NIVA-CYA 98	1.00	1.00	0.99	0.97	0.97	0.97	0.97	0.97
NIVA-CYA 406	1.00	1.00	0.98	0.98	0.98	0.98	0.97	0.98
NIVA-CYA 407	0.99	0.98	1.00	0.98	0.98	0.98	0.97	0.98
NIVA-CYA 540	0.97	0.98	0.98	1.00	1.00	1.00	0.99	0.99
NIVA-CYA 405	0.97	0.98	0.98	1.00	1.00	1.00	0.99	0.98
NIVA-CYA 34	0.97	0.98	0.98	1.00	1.00	1.00	0.98	0.98
NIVA-CYA 15	0.97	0.97	0.97	0.99	0.99	0.98	1.00	0.98
NIVA-CYA 56/3	0.97	0.98	0.98	0.99	0.98	0.98	0.98	1.00

B)



C)

	NIVA-CYA 98	NIVA-CYA 406	NIVA-CYA 407	NIVA-CYA 540	NIVA-CYA 405	NIVA-CYA 34	NIVA-CYA 15	NIVA-CYA 56/3
NIVA-CYA 98	100%	98.6%	91.2%	86.1%	86.4%	84.6%	87.3%	84.7%
NIVA-CYA 406	99.1%	100%	91.1%	86.3%	86.7%	84.7%	87.5%	84.6%
NIVA-CYA 407	88.9%	88.4%	100%	87.7%	87.9%	86.0%	88.7%	88.4%
NIVA-CYA 540	84.2%	83.8%	88.4%	100%	99.3%	97.5%	91.1%	90.2%
NIVA-CYA 405	83.7%	83.2%	87.8%	98.3%	100%	97.2%	90.9%	90.4%
NIVA-CYA 34	83.7%	83.3%	87.7%	98.7%	99.3%	100%	90.8%	88.7%
NIVA-CYA 15	82.9%	82.2%	86.4%	88.3%	89.0%	87.2%	100%	90.4%
NIVA-CYA 56/3	86.2%	86.3%	87.5%	89.6%	89.1%	89.5%	87.1%	100%

FIG 1 Comparison of eight *Planktothrix* genomes. Strains are color coded according to red (red font) and green (green font) phenotypes. (A) Correlation matrix of pairwise comparison of COG profiles. Pearson correlation coefficients range from 1.00 (highest correlation, dark blue) to 0.97 (dark red). (B) Hierarchical clustering of strains based on COG profiles and maximum likelihood phylogeny of 200 randomly selected core genes (see Materials and Methods). The proximity of grouping in the hierarchical tree indicates the relative degree of similarity of genomes to each other. ML bootstrap values at all branches are shown. (C) Percentage of homologous protein-coding genes in pairwise analyses. Values range from 100% (dark blue) to 82.2% (red). Homologous genes were detected using BLASTP and a minimum of 80% sequence identity.

tein-encoding genes) and phylogenetic analyses were performed for comparison of eight sequenced genomes. Nucleotide diversity analysis of 200 randomly selected core genes revealed high sequence similarity between eight compared *Planktothrix* genomes, with only 6,243 polymorphic sites detected within the 217,546-bp-long alignment and nucleotide diversity (Pi) of 0.01222. Pairwise comparisons of COG profiles suggest that the 8 *Planktothrix* genomes are also functionally highly similar (Pearson correlation coefficients range from 0.97 to 1.00 [Fig. 1A]). However, the hierarchical clustering of COG-annotated genes (Fig. 1B) reveals that some genomes are more closely related than others. Proximity of groupings indicates the relative degree of similarity between the COG-annotated gene complements and shows that strains within a chemotype are more closely related, as seen in, for instance, Cht 1 (NIVA-CYA 98 and NIVA-CYA 406) and Cht 9/10 (NIVA-CYA 34, NIVA-CYA 405, and NIVA-CYA 540), than are strains of different chemotypes. The amount of homologous genes in 8 genomes was detected by examining all protein-coding genes, including the 33 to 35% not assigned to COGs. Pairwise gene complement comparisons revealed that at least 82% of protein-encoding genes have a homolog in all eight genomes (Fig. 1C). Furthermore, 98 to 99% of the protein-encoding genes in the red strain NIVA-CYA 540 have a homolog in the green strains NIVA-CYA 34 and NIVA-CYA 405. Phylogenetic analyses of 200 core genes support the finding that strains within a chemotype (Cht 1 and Cht 9/10) are more closely related (Fig. 1B; see also Fig. S1 in the supplemental material), essentially showing the same topology as the COG clustering. All together, these results suggest that the occurrence of red strains within Cht 9/10, which mostly consists of green strains, is probably caused by acquisition of phycoerythrin-encoding genes by green strains.

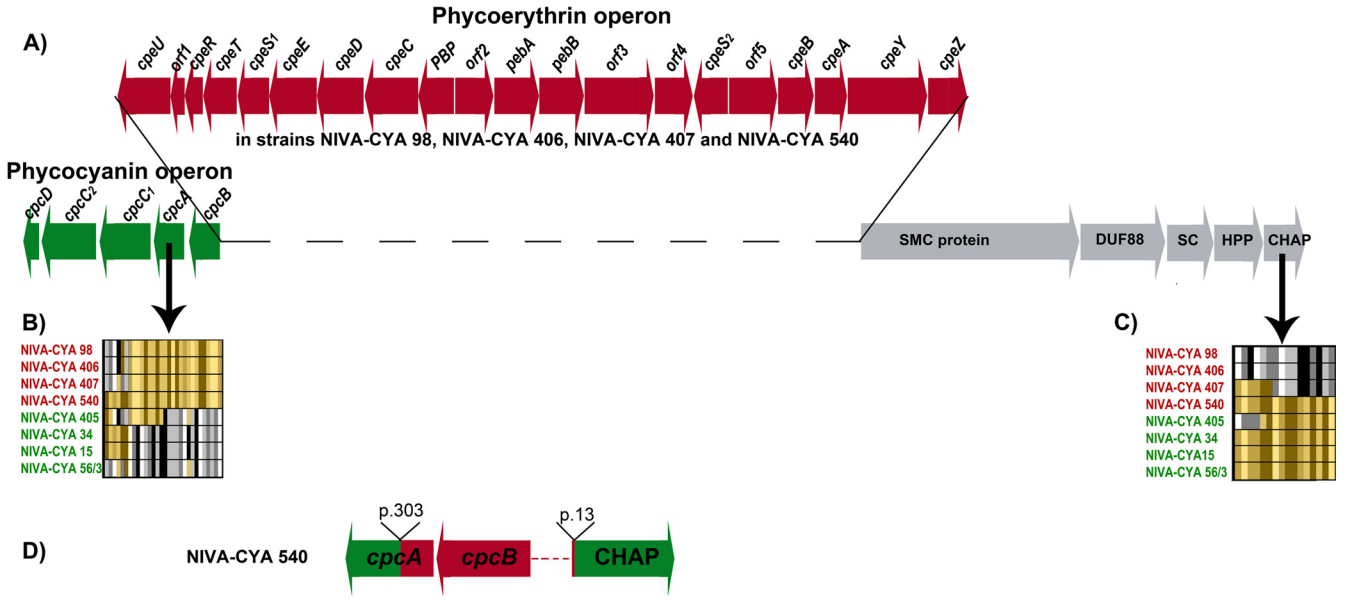


FIG 2 Phycocyanin and phycoerythrin operons in *Planktothrix*. (A) The figure shows organization of the phycocyanin (green) gene cluster with individual gene names, the phycoerythrin gene cluster (red) with individual gene names, and flanking genes downstream of the *cpeZ*/*cpcB* gene. Flanking open reading frames are indicated as follows: SMC protein, SMC chromosome segregation protein; DUF88, protein of unknown function DUF88; SC, short-chain alcohol dehydrogenase; HPP, 4-hydroxyphenylpyruvate dioxygenase; CHAP, CHAP domain. (B) Alignment of informative sites in *cpcA*. Areas marked with a gold/yellow shade indicate identical regions between NIVA-CYA 540 and other *Planktothrix* strains. (C) Alignment of informative sites of the gene encoding the CHAP domain. Areas marked with a gold shade indicate identical regions between NIVA-CYA 540 and other *Planktothrix* strains. (D) Recombination breakpoints detected by RDP in NIVA-CYA 540. Strain NIVA-CYA 34 (green) was detected as major parent, and NIVA-CYA 98 (red) was detected as minor parent. Positions of recombination breakpoints in both genes are shown.

Analyses of genes coding for PC. Eight different phycocyanin (PC)-encoding genes were identified in all *Planktothrix* genomes studied. Similar to other cyanobacteria, genes coding for alpha PC (*cpcA*) and beta PC (*cpcB*) subunits are clustered with three rod linker proteins ($cpcC_1$, $cpcC_2$, and *cpcD*) in a 3.2-kb $cpcBAC_1C_2D$ operon (Fig. 2A). Additionally, the genes *cpcE* and *cpcF* (encoding two subunits of phycobilin lyase) form another small gene cluster. Neither the *cpcEF* cluster nor the *cpcG* gene (coding for phycobilisome rod-core linker) is located in the vicinity of other photosynthesis genes. The *cpc* genes are highly conserved, with between 0.00085 and 0.03754 substitutions per base (Table 2).

Maximum likelihood phylogenetic analyses of concatenated *cpcA* and *cpcB* sequences revealed that NIVA-CYA 540 clusters with red strains NIVA-CYA 98, NIVA-CYA 406, and NIVA-CYA 407. However, NIVA-CYA 540 clusters together with green NIVA-CYA 34 and NIVA-CYA 405 strains in the $cpcC_1C_2$ tree (Fig. 3). Comparison of *cpcB* sequences showed that the NIVA-CYA 540 gene is identical to that of all red strains (see Fig. S2 in the supplemental material). Interestingly, examination of the *cpcA* gene located downstream of *cpcB* revealed that in strain NIVA-CYA 540, the 5′ end of the gene is identical with red strains while a small part of the 3′ end of the gene resembles the green strains (Fig. 2B). This finding indicates that in strain NIVA-CYA 540, the 5′ (upstream) part of the $cpcBAC_1C_2D$ operon was replaced during insertion of the phycoerythrin gene cluster using genic homology as the recombination locus.

Genes encoding phycoerythrin are nearly identical in all four red strains. The genes involved in synthesis of phycoerythrin (PE) are organized into one 19.7-kb bidirectional gene cluster located upstream of the phycocyanin gene cluster $cpcBAC_1C_2D$ (Fig. 2A). The phycoerythrin gene cluster comprises 20 genes in *Planktothrix* (Table 2; Fig. 2A), and comparisons with other cyanobacteria indicate that 18 of the genes have homologs in previously characterized phycoerythrin operons. The genes *pebAB* and *cpeAB* code for phycoerythrobilin (PEB) and PE alpha and beta subunits, respectively. Six genes code for the PE-specific phycobilisome rod linkers, and four genes encode lyases (Table 2). Two of the five genes encoding hypothetical proteins, ORF3 and ORF4 (Fig. 2A; Table 2), have no homologs in PE gene clusters in genomes available through the IMG system.

The phycoerythrin gene cluster is otherwise highly conserved at the sequence level. Only three genes (*cpeA*, *cpeB*, and *cpeU*) show any nucleotide diversity at all; the remainder of the genes are identical in all four red strains (Table 2). In these three genes, the strain NIVA-CYA 540 has a sequence identical to that of strain NIVA-CYA 407, which might indicate that the phycoerythrin gene cluster in NIVA-CYA 540 originates from this strain or a strain associated with the same chemotype.

Identification of recombination breakpoints. Visual inspection of the phycocyanin gene cluster $cpcBAC_1C_2D$ revealed that one of the breakpoints of the homologous recombination event leading to insertion of the phycoerythrin operon in strain NIVA-CYA 540 is most likely located within the *cpcA* gene. In order to identify the second recombination breakpoint, an approximately 9-kb-long region downstream of *cpeZ* in the genomes of the red strains (corresponding to a 9-kb-long region upstream of *cpcB* in the genomes of the green strains) was examined. The genes encoding the following proteins flank *cpeZ*/*cpcB* in all eight genomes: SMC chromosome segregation protein, protein of unknown function DUF88, short-chain alcohol dehydrogenase, 4-hydroxy-

TABLE 2 Genes encoding phycocyanin and phycoerythrin

Gene	Length (bp)	Product name	KEGG orthology (KO) term	Nucleotide diversity (Pi)	No. of informative sites
Phycocyanin genes					
cpcB	519	Phycobilisome protein	cpcB phycocyanin beta chain	0.02746	30
cpcA	489	Phycobilisome protein	cpcA phycocyanin alpha chain	0.03754	21
cpcC1	816	Phycobilisome linker polypeptide	cpcC phycocyanin-associated rod linker protein	0.01619	24
cpcC2	870	Phycobilisome linker polypeptide	cpcC phycocyanin-associated rod linker protein	0.03173	46
cpcD	243	CpcD/allophycocyanin linker domain	cpcD phycocyanin-associated, rod	0.01760	1
cpcE	837	FOG: HEAT repeat	cpcE phycocyanobilin lyase alpha subunit	0.00196	3
cpcF	633	PBS lyase HEAT-like repeat	cpcF phycocyanobilin lyase beta subunit	0.00085	1
cpcG	753	Phycobilisome linker polypeptide	cpcG phycobilisome rod-core linker protein	0.00142	2
Phycoerythrin genes					
cpeA	495	Phycobilisome protein	cpeA, mpeA phycoerythrin alpha chain	0.00135	1
cpeB	555	Phycobilisome protein	cpeB, mpeB phycoerythrin beta chain	0.00961	8
cpeC	879	Phycobilisome linker polypeptide	cpeC, mpeC phycoerythrin-associated linker protein	0.00000	0
cpeD	768	Phycobilisome linker polypeptide	cpeD, mpeD phycoerythrin-associated linker protein	0.00000	0
cpeE	759	Phycobilisome linker polypeptide	cpeE phycoerythrin-associated linker protein	0.00000	0
cpeR	306	Hypothetical protein	cpeR phycoerythrin-associated linker protein	0.00000	0
cpeS	534	CpeS-like protein	cpeS phycoerythrin-associated linker protein	0.00000	0
cpeS2	588	CpeS-like protein		0.00000	0
cpeT	585	CpeT/CpcT family (DUF1001)	cpeT CpeT protein	0.00000	0
cpeU	900	FOG: HEAT repeat	cpeU, mpeU bilin biosynthesis protein	0.01852	25
cpeY	1,281	PBS lyase HEAT-like repeat	cpeY bilin biosynthesis protein	0.00000	0
cpeZ	603	Hypothetical protein	cpeZ bilin biosynthesis protein	0.00000	0
pebA	729	Ferredoxin-dependent bilin reductase	pebA 15,16-dihydrobiliverdin:ferredoxin oxidoreductase (EC 1.3.7.2)	0.00000	0
pebB	735	Ferredoxin-dependent bilin reductase	pebB phycoerythrobilin:ferredoxin oxidoreductase (EC 1.3.7.3)	0.00000	0
orf1	240	Hypothetical protein		0.00000	0
PBP	573	Phycobilisome protein		0.00000	0
orf2	660	Hypothetical protein		0.00000	0
orf3	1,161	Hypothetical protein		0.00000	0
orf4	621	Hypothetical protein		0.00000	0
orf5	798	Uncharacterized low-complexity proteins		0.00000	0

phenylpyruvate dioxygenase, and CHAP domain (Fig. 2A). Examination of informative sites (see Fig. S3 in the supplemental material) revealed several recombination events in all genes except for the gene coding for protein of unknown function DUF88, which is highly conserved. Interestingly, the 8.5-kb sequence comprising all genes except for the gene encoding the CHAP domain is almost identical in strains NIVA-CYA 540 and NIVA-CYA 407, differing in only one position. Visual examination of informative sites of the gene coding for the CHAP domain revealed that the strains NIVA-CYA 540 and NIVA-CYA 407 have different sequences at the 5′ end of the CHAP domain gene (Fig. 2C). We therefore suggest that putative recombination breakpoints are located in *cpcA* and the gene coding for the CHAP domain, indicating that the DNA fragment inserted through homologous recombination is approximately 30 kb long. This entire 30-kb fragment differs at only a single nucleotide between strains NIVA-CYA 540 and NIVA-CYA 407, which again supports the notion that the DNA fragment inserted into an originally green strain originates from strain NIVA-CYA 407 or a closely related strain associated with the same chemotype as NIVA-CYA 407.

The aligned and concatenated sequences of *cpcBA* (reverse complement) and the CHAP domain genes were examined for recombination signals with the RDP4 software. Four methods implemented in RDP (Geneconv, MaxChi, Chimaera, and 3 Seq) detected NIVA-CYA 540 as recombinant, with the green strain NIVA-CYA 34 as major parent and the red strain NIVA-CYA 98 as minor parent (Fig. 2D). Recombination breakpoints were detected in *cpcA* (position 303) and the CHAP gene (position 13) and correspond well with positions of the most 5′ red-like informative site in *cpcA* (position 356) and the most 5′ green-like informative site in the gene encoding the CHAP domain (position 141).

Presence of oligopeptide production genes. Since classification of *Planktothrix* strains into chemotypes is based on the profile of oligopeptide isoforms produced, all genomes were examined for the presence of the genes needed for oligopeptide production. Oligopeptide isoforms produced by each strain (according to the work of Rohrlack and coworkers [35]) and the presence of genes encoding biosynthetic pathways are shown in Table 3. In general, genes encoding synthesis of all oligopeptides but microginin are present in every genome examined, even though several oligopeptides are not produced by all strains. The microginin gene cluster

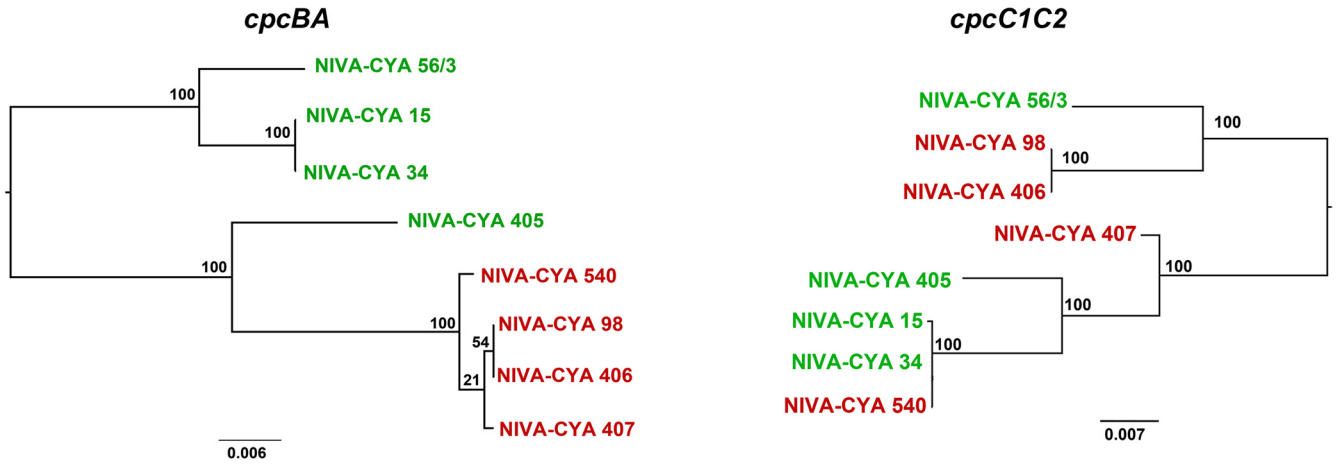


FIG 3 Maximum likelihood phylogenies of *cpcA* and *cpcB* versus *cpcC1* and *cpcC2*, all concatenated. ML bootstrap values on all branches are shown. *Planktothrix* strains are color coded according to red (red font) and green (green font) phenotypes.

is present only in the genome of microginin-producing strain NIVA-CYA 98.

DISCUSSION

Extensive genomic similarity between two strains with red and green pigmentation. Based on oligopeptide profiles, Rohrlack and coworkers (35) have classified 87 *Planktothrix* strains from Northern Europe into 17 distinct chemotypes. Here, we compare genomes of eight *Planktothrix* strains from six of these 17 chemotypes. However, Cht 10 is, according to Rohrlack and coworkers (35), closely related to Cht 9 and was considered a subgroup of the latter in the present study. Both comparison of protein-coding genes and phylogenetic analyses of 200 core genes revealed that the genome of NIVA-CYA 540, the only Cht 10 strain described by Rohrlack and coworkers (35), is highly similar to the genomes of the Cht 9 strains NIVA-CYA 34 and NIVA-CYA 405, supporting the hypothesis that Cht 10 is a subgroup of Cht 9. In the light of genome similarity and the fact that cyanopeptolin isoforms are the only oligopeptides which are different in Cht 9 and Cht 10, we expect that a mutation within the cyanopeptolin synthetase gene cluster leads to the formation of Cht 10. Additional analyses are also required to elucidate why several oligopeptide gene clusters are present in genomes of strains that do not produce (detectable amounts of) these oligopeptides; it may be that the operons have been inactivated by mutations, as shown for microcystin synthetase genes (9, 50).

Transfer of phycoerythrin gene cluster between *Planktothrix* strains. The most significant finding of this study is the transfer of the entire 19.7-kb phycoerythrin gene cluster between *Planktothrix* strains isolated from Lake Steinsfjorden. We propose that the DNA fragment containing the phycoerythrin gene cluster was transferred by transformation or transduction and integrated into the chromosome through homologous recombination.

Transformation and transduction are considered the most likely mechanisms for DNA uptake, since no gene transfer agent systems have been documented in cyanobacteria (18) and no genes coding for cell-to-cell junctions were found in any of the *Planktothrix* genomes. Cyanobacteria have been shown to be naturally competent (e.g., references 51 and 52); therefore, it is conceivable that *Planktothrix* strains have the ability to take up naked DNA from the environment. Free DNA is released into the aquatic environment from decomposing cells or viral particles or through excretion from living cells and can persist for days to years (53) in fragments exceeding 10 kb in size (54). As several chemotypes of *Planktothrix* have been shown to coexist in Lake Steinsfjorden (34, 35), DNA from different chemotypes is most likely continuously available for natural transformation. Cyanophage-mediated transduction has also been shown to be important (e.g., references 55 and 56), and the majority of marine cyanophages have been reported to carry one or more photosynthetic genes (57, 58). Notably, no cyanophages containing the entire phycoerythrin gene cluster have been reported so far, and no obvious traces of cyanophages were found in the vicinity of phycocyanin and phycoerythrin gene clusters.

Following the physical transfer of exogenous DNA into a new cell, foreign DNA is usually integrated into the recipient genome through recombination. Functionally related genes (islands) that are frequently transferred between strains are often located in proximity to tRNAs or transposons (e.g., references 59 and 60). In this case, however, we did not find such hallmarks of island shuttles. For homologous recombination, the incoming sequences must contain a region of sufficient length that is similar to sequences in the recipient genome (61). The *Planktothrix* genomes investigated in this study show a high degree of sequence identity, and therefore, homologous recombination events between these eight strains could be expected to occur frequently. Moreover, homologous recombination can facilitate the transfer of heterologous genes, particularly when these are located between regions of sequence similarity shared by the donor and the recipient (62). In strain NIVA-CYA 540, the DNA fragment containing phycoerythrin genes is inserted at the same location as in other red genomes, suggesting homology-based incorporation of DNA. In the genomes of red strains, the phycoerythrin genes are located between the phycocyanin operon and genes upstream of the operon (putative gene encoding a CHAP domain). Transfer and incorporation of such a DNA fragment from a red genome into a green genome will lead to transfer of the phycoerythrin gene cluster and production of the pigment.

The DNA fragment inserted to NIVA-CYA 540 is nearly identical to the equivalent genome region in strain NIVA-CYA 407

TABLE 3 Oligopeptide isoforms produced by *Planktothrix* strains (25) and presence of oligopeptide production genes in genomes

Oligopeptide isoform (mol mass, $[M + H]^+$)	Presence for strain and chemotype[a]							
	N-C 98, Cht 1	N-C 406, Cht 1	N-C 407, Cht 5	N-C 56/3, Cht 7	N-C 34, Cht 9	N-C 405, Cht 9	N-C 540, Cht 10	N-C 15, Cht 14
Microcystins								
Dm-Mcyst LR (981.6)	X	X	X	X	X	X	X	
[Asp3,Dha7]Mcyst-RR (1,010.5)								X
Dm-Mcyst RR (1,024.7)	X	X	X	X	X	X	X	
Dm-Mcyst YR (1,031.7)				X				
Dm-Mcyst HtyrR (1,045.6)			X					
Gene cluster present	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cyanopeptolins								
Cyanopeptolin (1,029.7)					X	X		
Cyanopeptolin (1,034.7)			X					
Cyanopeptolin (1,064.5)								X
Cyanopeptolin (1,084.7)			X					
Cyanopeptolin (1,093.7)					X	X		
Cyanopeptolin (1,098.7)			X					
Cyanopeptolin (1,109.6)							X	
Oscillapeptin G (1,112.7)	X	X						
Cyanopeptolin (1,142.7)				X				
Cyanopeptolin (1,160.7)				X				
Gene cluster present	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Aeruginosins								
Aeruginosin (559.5), 1 and 2					X	X	X	
Aeruginosin (583.5), 1 and 2					X	X	X	
Aeruginosin (593.5)	X	X						
Aeruginosin A (617.5)	X	X						
Gene cluster present	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Microginins								
Oscillaginin B (581.5)	X	X						
Oscillaginin A (615.5)	X	X						
Gene cluster present	Yes	Yes	No	No	No	No	No	No
Anabaenopeptids								
Anabaenopeptin C (809.6)			X	X				
Me-anabaenopeptin C (823.6)			X	X				
Anabaenopeptin B (837.6)	X	X	X	X	X	X	X	X
Anabaenopeptin A (844.6)	X	X			X	X	X	X
Anabaenopeptin F (851.6)	X	X	X	X	X	X	X	
Oscillamide Y (858.6)	X	X			X	X	X	
Gene cluster present	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Microviridins								
Putative microviridin (1,854.8)			X	X				
Putative microviridin (1,971.8)	X	X						
Putative gene cluster present	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Oscillatorin								
Oscillatorin (1,240.4)	X	X						
oscA gene present	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

[a] N-C, NIVA-CYA.

(Cht 5), suggesting that the DNA fragment containing the phycoerythrin operon originates from a strain associated with Cht 5 and that the transfer event is recent. The latter is also supported by long-term field study of Lake Steinsfjorden since the red type of Cht 9 was first reported in 2004 (34).

Fixation of phycoerythrin genes within strains associated with Cht 9/10. Bacteria tend to lose nonfunctional or otherwise unneeded DNA from their genomes (63, 64); therefore, fixation of the acquired DNA within the genome is highly dependent on its functionality and/or advantage to the recipient under selectable environmental conditions (65–67). An operon, like the phycoerythrin operon transferred to NIVA-CYA 540 from a closely related donor, contains all genes and regulatory regions required for a function to be conferred and is therefore more likely to be fixed than a single gene or genes from a distant relative after HGT (20). The red color of NIVA-CYA 540 filaments supports the functionality of the horizontally transferred phycoerythrin gene cluster. Consequently, the fixation of the operon within the chemotype is likely to depend on the competitive gain associated with expressing the phycoerythrin gene cluster and the contribution to the ecological fitness of the given *Planktothrix* chemotype. Since phycocyanin and phycoerythrin absorb different colors of light, red and green cyanobacteria can coexist under white light as shown by competition studies (68). Previous studies have also shown that growth of *Planktothrix* in Lake Steinsfjorden is probably first of all controlled by irradiance (69). Therefore, coexistence of strains associated with the same chemotype (Cht 9/10) but producing different light-harvesting pigments might increase the range of niches in which this chemotype is able to thrive. At the same time, green- and red-colored *Planktothrix* strains have a similar depth distribution in Lake Steinsfjorden, indicating that growth under low-irradiance conditions and formation of blooms in the metalimnion do not necessarily require phycoerythrin (69). Moreover, production of phycoerythrin involves a metabolic cost and does

not necessarily led to a higher growth rate (70). Further sampling of *Planktothrix* strains from Lake Steinsfjorden or from other lakes considered to form the same metapopulation (29) and competition experiments under laboratory conditions are necessary to investigate whether the red type of Cht 9 has an advantage or disadvantage compared to the green type of Cht 9. Another important question is whether transfer of other genes or gene clusters is necessary to become a successful red *Planktothrix* strain or if transfer of a phycoerythrin gene cluster is all that is needed.

Based on the higher degree of sequence similarity detected in phycoerythrin genes than in phycocyanin genes, it is appealing to suggest that the ancestral state of *Planktothrix* was green and that, during evolution, some strains have acquired genes coding for phycoerythrin as previously suggested for other cyanobacteria (40, 41). However, analysis of both phycoerythrin and phycocyanin sequences from a larger number of red *Planktothrix* strains is necessary before conclusions here can be drawn.

In summary, we show that the filamentous cyanobacterium *Planktothrix* is able to take up and integrate large (up to 30-kb-long) DNA fragments. Our results demonstrate that through a single HGT event, highly similar *Planktothrix* strains can acquire a complex gene cluster coding for synthesis of an additional photosynthetic pigment and thus an opportunity to exploit a new ecological niche(s).

ACKNOWLEDGMENTS

We thank the Norwegian Institute for Water Research (NIVA) for providing cultured NIVA-CYA strains, the Norwegian Sequencing Centre (http://www.sequencing.uio.no/) for 454 sequencing of *Planktothrix* strains, Heiko Vogel from the Max Planck Institute for Chemical Ecology in Jena for Sanger and 454 sequencing of strain NIVA-CYA 34, and Kjetil L. Voje for discussions and for random sampling of core genomes using R.

The work was supported by Norwegian Research Council grant no 183360/S30 to Thomas Rohrlack and by University of Oslo Project no. 460960 to Kjetill S. Jakobsen.

REFERENCES

1. **Andam CP, Fournier GP, Gogarten JP.** 2011. Multilevel populations and the evolution of antibiotic resistance through horizontal gene transfer. FEMS Microbiol. Rev. **35:**756–767.
2. **Koonin EV, Makarova KS, Aravind L.** 2001. Horizontal gene transfer in prokaryotes: quantification and classification. Annu. Rev. Microbiol. **55:**709–742.
3. **Skippington E, Ragan MA.** 2011. Lateral genetic transfer and the construction of genetic exchange communities. FEMS Microbiol. Rev. **35:**707–735.
4. **Wiedenbeck J, Cohan FM.** 2011. Origins of bacterial diversity through horizontal genetic transfer and adaptation to new ecological niches. FEMS Microbiol. Rev. **35:**957–976.
5. **Zhaxybayeva O, Gogarten JP, Charlebois RL, Doolittle WF, Papke RT.** 2006. Phylogenetic analyses of cyanobacterial genomes: quantification of horizontal gene transfer events. Genome Res. **16:**1099–1108.
6. **Shi T, Falkowski PG.** 2008. Genome evolution in cyanobacteria: the stable core and the variable shell. Proc. Natl. Acad. Sci. U. S. A. **105:**2510–2515.
7. **Rudi K, Skulberg OM, Jakobsen KS.** 1998. Evolution of cyanobacteria by exchange of genetic material among phyletically related strains. J. Bacteriol. **180:**3453–3461.
8. **Tooming-Klunderud A, Fewer DP, Rohrlack T, Jokela J, Rouhiainen L, Sivonen K, Kristensen T, Jakobsen KS.** 2008. Evidence for positive selection acting on microcystin synthetase adenylation domains in three cyanobacterial genera. BMC Evol. Biol. **8:**256. doi:10.1186/1471-2148-8-256.
9. **Tooming-Klunderud A, Mikalsen B, Kristensen T, Jakobsen KS.** 2008. The mosaic structure of the *mcyABC* operon in *Microcystis*. Microbiology **154:**1886–1899.
10. **Caro-Quintero A, Deng J, Auchtung J, Brettar I, Höfle MG, Klappenbach J, Konstantinidis KT.** 2011. Unprecedented levels of horizontal gene transfer among spatially co-occurring *Shewanella* bacteria from the Baltic Sea. ISME J. **5:**131–140.
11. **Klassen JL.** 2009. Pathway evolution by horizontal transfer and positive selection is accommodated by relaxed negative selection upon upstream pathway genes in purple bacterial carotenoid biosynthesis. J. Bacteriol. **191:**7500–7508.
12. **Mikalsen B, Boison G, Skulberg OM, Fastner J, Davies W, Gabrielsen TM, Rudi K, Jakobsen KS.** 2003. Natural variation in the microcystin synthetase operon *mcyABC* and impact on microcystin production in *Microcystis* strains. J. Bacteriol. **185:**2774–2785.
13. **Bolhuis H, Severin I, Confurius-Guns V, Wollenzien UI, Stal LJ.** 2010. Horizontal transfer of the nitrogen fixation gene cluster in the cyanobacterium *Microcoleus chthonoplastes*. ISME J. **4:**121–130.
14. **Martiny AC, Huang Y, Li W.** 2009. Occurrence of phosphate acquisition genes in *Prochlorococcus* cells from different ocean regions. Environ. Microbiol. **11:**1340–1347.
15. **Rae BD, Förster B, Badger MR, Price GD.** 2011. The CO2-concentrating mechanism of *Synechococcus* WH5701 is composed of native and horizontally-acquired components. Photosynth. Res. **109:**59–72.
16. **Gogarten JP, Doolittle WF, Lawrence JG.** 2002. Prokaryotic evolution in light of gene transfer. Mol. Biol. Evol. **19:**2226–2238.
17. **Papke RT, Zhaxybayeva O, Feil EJ, Sommerfeld K, Muise D, Doolittle WF.** 2007. Searching for species in haloarchaea. Proc. Natl. Acad. Sci. U. S. A. **104:**14092–14097.
18. **Popa O, Hazkani-Covo E, Landan G, Martin W, Dagan T.** 2011. Directed networks reveal genomic barriers and DNA repair bypasses to lateral gene transfer among prokaryotes. Genome Res. **21:**599–609.
19. **Mira A, Ochman H, Moran NA.** 2001. Deletional bias and the evolution of bacterial genomes. Trends Genet. **17:**589–596.
20. **Lawrence J.** 1999. Selfish operons: the evolutionary impact of gene clustering in prokaryotes and eukaryotes. Curr. Opin. Genet. Dev. **9:**642–648.
21. **Elena SF, Ekunwe L, Hajela N, Oden SA, Lenski RE.** 1998. Distribution of fitness effects caused by random insertion mutations in *Escherichia coli*. Genetica **102-103:**349–358.
22. **Remold SK, Lenski RE.** 2004. Pervasive joint influence of epistasis and plasticity on mutational effects in *Escherichia coli*. Nat. Genet. **36:**423–426.
23. **Barrett RD, MacLean RC, Bell G.** 2006. Mutations of intermediate effect are responsible for adaptation in evolving *Pseudomonas fluorescens* populations. Biol. Lett. **2:**236–238.
24. **Imhof M, Schlötterer C.** 2001. Fitness effects of advantageous mutations in evolving *Escherichia coli* populations. Proc. Natl. Acad. Sci. U. S. A. **98:**1113–1117.
25. **Orr HA.** 2005. The genetic theory of adaptation: a brief history. Nat. Rev. Genet. **6:**119–127.
26. **D'Alelio D, Gandolfi A.** 2012. Recombination signals in the *rpoC1* gene indicate gene-flow between *Planktothrix* (Cyanoprokaryota) species. J. Phycol. **48:**1424–1432.
27. **Kurmayer R, Gumpenberger M.** 2006. Diversity of microcystin genotypes among populations of the filamentous cyanobacteria *Planktothrix rubescens* and *Planktothrix agardhii*. Mol. Ecol. **15:**3849–3861.
28. **Rounge TB, Rohrlack T, Decenciere B, Edvardsen B, Kristensen T, Jakobsen KS.** 2010. Subpopulation differentiation associated with nonribosomal peptide synthetase gene cluster dynamics in the cyanobacterium *Planktothrix* spp. J. Phycol. **46:**645–652.
29. **Sogge H, Rohrlack T, Rounge TB, Sønstebø JH, Tooming-Klunderud A, Kristensen T, Jakobsen KS.** 2013. Gene flow, recombination, and selection in cyanobacteria: population structure of geographically related *Planktothrix* freshwater strains. Appl. Environ. Microbiol. **79:**508–515.
30. **Anagnostidis K, Komárek J.** 1988. Modern approach to the classification system of cyanophytes. 3. Oscillatorales. Arch. Hydrobiol. Suppl. **80:**327–472.
31. **Beard SJ, Handley BA, Hayes PK, Walsby AE.** 1999. The diversity of gas vesicle genes in *Planktothrix rubescens* from Lake Zürich. Microbiology **145:**2757–2768.
32. **Humbert J, Le Berre B.** 2001. Genetic diversity in two species of freshwater cyanobacteria, *Planktothrix* (*Oscillatoria*) *rubescens* and *P. agardhii*. Arch. Hydrobiol. **150:**197–206.
33. **Suda S, Watanabe MM, Otsuka S, Mahakahant A, Yongmanitchai W, Nopartnaraporn N, Liu Y, Day JG.** 2002. Taxonomic revision of water-bloom-forming species of oscillatorioid cyanobacteria. Int. J. Syst. Evol. Microbiol. **52:**1577–1595.

34. **Rohrlack T, Edvardsen B, Skulberg R, Halstvedt CB, Utkilen HC, Ptacnik R, Skulberg OM.** 2008. Oligopeptide chemotypes of the toxic freshwater cyanobacterium *Planktothrix* can form subpopulations with dissimilar ecological traits. Limnol. Oceanogr. **53:**1279–1293.
35. **Rohrlack T, Skulberg R, Skulberg OM.** 2009. Distribution of oligopeptide chemotypes of the cyanobacterium *Planktothrix* and their persistence in selected lakes in Fennoscandia. J. Phycol. **45:**1259–1265.
36. **Rounge TB, Rohrlack T, Nederbragt AJ, Kristensen T, Jakobsen KS.** 2009. A genome-wide analysis of nonribosomal peptide synthetase gene clusters and their peptides in a *Planktothrix rubescens* strain. BMC Genomics **10:**396. doi:10.1186/1471-2164-10-396.
37. **Welker M, von Döhren H.** 2006. Cyanobacterial peptides—nature's own combinatorial biosynthesis. FEMS Microbiol. Rev. **30:**530–563.
38. **MacColl, R.** 1998. Cyanobacterial phycobilisomes. J. Struct. Biol. **124:** 311–334.
39. **Ernst A, Becker S, Wollenzien UI, Postius C.** 2003. Ecosystem-dependent adaptive radiations of picocyanobacteria inferred from 16S rRNA and ITS-1 sequence analysis. Microbiology **149:**217–228.
40. **Jasser I, Krolicka A, Karnkowska-Ishikawa A.** 2011. A novel phylogenetic clade of picocyanobacteria from the Mazurian lakes (Poland) reflects the early ontogeny of glacial lakes. FEMS Microbiol. Ecol. **75:**89–98.
41. **Six C, Thomas JC, Garczarek L, Ostrowski M, Dufresne A, Blot N, Scanlan DJ, Partensky F.** 2007. Diversity and evolution of phycobilisomes in marine *Synechococcus* spp.: a comparative genomics study. Genome Biol. **8:**R259. doi:10.1186/gb-2007-8-12-r259.
42. **Nederbragt AJ, Rounge TB, Kausrud KL, Jakobsen KS.** 2010. Identification and quantification of genomic repeats and sample contamination in assemblies of 454 pyrosequencing reads. Sequencing. doi:10.1155/2010/782465.
43. **Chain PSG, Grafham DV, Fulton RS, FitzGerald MG, Hostetler J, Muzny D, Ali J, Birren B, Bruce DC, Buhay C, Cole JR, Ding Y, Dugan S, Field D, Garrity GM, Gibbs R, Graves T, Han CS, Harrison SH, Highlander S, Hugenholtz1 P, Khouri HM, Kodira CD, Kolker E, Kyrpides NC, Lang D, Lapidus A, Malfatti SA, Markowitz V, Metha T, Nelson KE, Parkhill J, Pitluck S, Qin X, Read TD, Schmutz J, Sozhamannan S, Sterk P, Strausberg RL, Sutton G, Thomson NR, Tiedje JM, Weinstock G, Wollam A, Genomic Standards Consortium Human Microbiome Project Jumpstart Consortium, Detter JC.** 2009. Genomics. Genome project standards in a new era of sequencing. Science **326:**236–237.
44. **Markowitz VM, Mavromatis K, Ivanova NN, Chen IM, Chu K, Kyrpides NC.** 2009. IMG ER: a system for microbial genome annotation expert review and curation. Bioinformatics **25:**2271–2278.
45. **Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S.** 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. **28:**2731–2739.
46. **Stamatakis A.** 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics **22:**2688–2690.
47. **Librado P, Rozas J.** 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics **25:**1451–1452.
48. **Posada D, Crandall KA.** 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics **14:**817–818.
49. **Martin D, Rybicki E.** 2000. RDP: detection of recombination amongst aligned sequences. Bioinformatics **16:**562–563.
50. **Christiansen G, Kurmayer R, Liu Q, Börner T.** 2006. Transposons inactivate biosynthesis of the nonribosomal peptide microcystin in naturally occurring *Planktothrix* spp. Appl. Environ. Microbiol. **72:**117–123.
51. **Dittmann E, Neilan BA, Erhard M, von Döhren H, Börner T.** 1997. Insertional mutagenesis of a peptide synthetase gene that is responsible for hepatotoxin production in the cyanobacterium *Microcystis aeruginosa* PCC 7806. Mol. Microbiol. **26:**779–787.
52. **Johnsborg O, Eldholm V, Havarstein LS.** 2007. Natural genetic transformation: prevalence, mechanisms and function. Res. Microbiol. **158:** 767–778.
53. **Vlassov VV, Laktionov PP, Rykova EY.** 2007. Extracellular nucleic acids. Bioessays **29:**654–667.
54. **Corinaldesi C, Danovaro R, Dell'Anno A.** 2005. Simultaneous recovery of extracellular and intracellular DNA suitable for molecular studies from marine sediments. Appl. Environ. Microbiol. **71:**46–50.
55. **Kuno S, Yoshida T, Kamikawa R, Hosoda N, Sako Y.** 2010. The distribution of a phage-related insertion sequence element in the cyanobacterium, *Microcystis aeruginosa*. Microbes Environ. **25:**295–301.
56. **Weinbauer MG, Rassoulzadegan F.** 2004. Are viruses driving microbial diversification and diversity? Environ. Microbiol. **6:**1–11.
57. **Clokie MR, Shan J, Bailey S, Jia Y, Krisch HM, West S, Mann NH.** 2006. Transcription of a 'photosynthetic' T4-type phage during infection of a marine cyanobacterium. Environ. Microbiol. **8:**827–835.
58. **Lindell D, Jaffe JD, Johnson ZI, Church GM, Chisholm SW.** 2005. Photosynthesis genes in marine viruses yield proteins during host infection. Nature **438:**86–89.
59. **Larbig KD, Christmann A, Johann A, Klockgether J, Hartsch T, Merkl R, Wiehlmann L, Fritz HJ, Tümmler B.** 2002. Gene islands integrated into tRNA(Gly) genes confer genome diversity on a *Pseudomonas aeruginosa* clone. J. Bacteriol. **184:**6665–6680.
60. **Osborn AM, Boltner D.** 2002. When phage, plasmids, and transposons collide: genomic islands, and conjugative- and mobilizable-transposons as a mosaic continuum. Plasmid **48:**202–212.
61. **Thomas CM, Nielsen KM.** 2005. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. Nat. Rev. Microbiol. **3:**711–721.
62. **de Vries J, Wackernagel W.** 2002. Integration of foreign DNA during natural transformation of *Acinetobacter* sp. by homology-facilitated illegitimate recombination. Proc. Natl. Acad. Sci. U. S. A. **99:**2094–2099.
63. **Burke GR, MoranNA.** 2011. Massive genomic decay in *Serratia symbiotica*, a recently evolved symbiont of aphids. Genome Biol. Evol. **3:**195–208.
64. **Moran NA, McLaughlin HJ, Sorek R.** 2009. The dynamics and time scale of ongoing genomic erosion in symbiotic bacteria. Science **323:**379–382.
65. **Hao W, Golding GB.** 2006. The fate of laterally transferred genes: life in the fast lane to adaptation or death. Genome Res. **16:**636–643.
66. **Pal C, Papp B, Lercher MJ.** 2005. Adaptive evolution of bacterial metabolic networks by horizontal gene transfer. Nat. Genet. **37:**1372–1375.
67. **Zhaxybayeva O, Doolittle WF.** 2011. Lateral gene transfer. Curr. Biol. **21:**R242–R246.
68. **Stomp M, Huisman J, Vörös L, Pick FR, Laamanen M, Haverkamp T, Stal LJ.** 2007. Colourful coexistence of red and green picocyanobacteria in lakes and seas. Ecol. Lett. **10:**290–298.
69. **Halstvedt CB, Rohrlack T, Andersen T, Skulberg O, Edvardsen B.** 2007. Seasonal dynamics and depth distribution of *Planktothrix* spp. in Lake Steinsfjorden (Norway) related to environmental factors. J. Plankton Res. **29:**471–482.
70. **Raven JA.** 1984. A cost-benefit analysis of photon absorption by photosynthetic unicells. New Phytol. **98:**593–625.