High variability of the *gvpA-gvpC* region in *Microcystis* *

Xu Min, Xu Xudong, Gao Hong and Kong Renqiu **

(The State Key Laboratory of Freshwater Ecology and Biotechnology, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan 430072, China)

Accepted on May 23, 2007

Abstract Gas vesicles provide buoyancy to *Microcystis* and other common cyanobacterial bloom-forming species. *gvpA* and *gvpC* are structural genes encoding gas vesicle proteins. Phylogenetic analyses of 10 *Microcystis* strains/uncultured samples showed that *gvpC* and each intergenic segment of the *gvpA-gvpC* region can be divided into two types. The combination of different types of *gvpC* and intergenic segments is an important factor that diversifies this genomic region. Some *Microcystis* strains isolated in China possess a 172 to 176 bp sequence tag in the intergenic segment between *gvpA* and *gvpC*. The *gvpA-gvpC* region in *Microcystis* can be divided into at least 4 classes and more numbers of subclasses. Compared to *rbcLX* and other regions, the high variability of the *gvpA-gvpC* region should be more useful in identifying geographical isolates or ecotypes of *Microcystis*.

Keywords: *Microcystis*, *gvpA-gvpC*, diversity.

Gas vesicles are hollow, gas-filled structures that provide buoyancy in many aquatic microorganisms, including cyanobacteria. Many gas vesicles aggregate with each other to form gas vacuoles, which are found as refractile bodies in cells under a phase-contrast microscope. Gas vesicles are made of proteins GvpA and GvpC in a shape of cylindrical tubes with conical end caps[1]. GvpA is a small hydrophobic protein that assembles to form the ribbed wall of the gas vesicle[2—4]. GvpC, located on the outer surface of gas vesicle, contains contiguous and highly conserved 33-residue repeats (33RR) and strengthens the entire structure[5—10]. The width of gas vesicles is determined by the number of 33RR in GvpC and inversely related to the critical pressure of gas vesicles[9—11]. Apart from GvpA and GvpC, there are several other proteins involved in the formation of gas vesicles. Genes encoding these proteins are often arranged in a cluster[12—15].

Microcystis is a genus of cyanobacteria widely distributed all over the world. Some species of this genus are the most often found bloom-formers in eutrophic lakes, causing serious environmental problems. Gas vesicles, as the structure enhancing the buoyancy of cells, play an important role in the formation of "surface" water blooms of *Microcystis*. Complete sequences of *Microcystis gvp* gene cluster have been reported in the strains PCC 7806[15] and BC 8401[10]. These two *gvp* clusters from *Microcystis* and the one from *Anabaena* / *Nostoc* sp. PCC 7120, a filamentous N-fixing cyanobacterium, contain 10 different *gvp* genes in the same order, except the difference in the copy number of *gvpA* and the intragenic structure of *gvpC*. Another *Microcystis* strain PCC 9354 probably contains a similar *gvp* cluster as that in PCC 7806, but the *gvpA* copy number was not determined[15]. In the present study, we analyzed the *gvpA-gvpC* portion from different strains or uncultured samples of *Microcystis*. Phylogenetic analyses indicated the role of interstrain recombination in diversifying this genomic region in this group of environmentally important cyanobacteria. Some *Microcystis* strains isolated in China possess a sequence tag of 172 to 176 bases in the intergenic segment between *gvpA* and *gvpC*.

1 Materials and methods

1.1 Strains and culture conditions

Microcystis sp. strains FACHB 854, 910, 916, 927, 929 and 930 were obtained from the Freshwater Algae Culture Collection of the Institute of Hydrobiology (FACHB), Chinese Academy of Sciences, but of different eventual origins (Table 1). All of them

* Supported by National Natural Science Foundation of China (Grant No. 30470339) and the Major State Basic Research Development Program of China (2002CB412306)

** To whom correspondence should be addressed. E-mail: kongr@ihb.ac.cn

were cultured in BG11 medium[16] in the light (~30 $\mu E\ m^{-2}s^{-1}$) at 30℃.

Table 1. A list of *Microcystis* strains / uncultured samples and NCBI GenBank accession numbers of DNA sequences

Strains/uncultured samples	Location of original collection	GenBank accession numbers	
		gvp genes	*rbcLX*
FACHB 854	Lake Dianchi, China	AY254709	DQ888811
FACHB 910	Lake Wudalianchi, China	DQ888806	DQ888812
FACHB 916/ NIES 42	Lake Kasumigaura, Japan	DQ888807	DQ888813
FACHB 927/ PCC 7820	Loch Balgavies, Scotland	DQ888808	DQ888814
FACHB 929/ NIES 107	Lake Kawaguchi, Japan	DQ888809	DQ888815
FACHB 930	Lake Dianchi, China	DQ888810	DQ888816
CHAO	Lake Chao, China	DQ889221	No
TAI	Lake Tai, China	EF035170	No

1.2 Extraction of genomic DNA

Thirty milliliters of *Microcystis* cells were collected by centrifugation, washed with 10 mL of TE (10 mmol/L Tris.Cl, 1 mmol/L EDTA, pH 8.0) and resuspended in 1 mL of TE. The cells were lysed with 1% SDS at 37℃ for 1 to 2 hrs. After digestion with 100 $\mu g \cdot mL^{-1}$ proteinase K at 37℃ for additional 2 hrs, the cell lysate was extracted with phenol-chloroform twice and chloroform once. Nucleic acids were precipitated with ethanol and RNA was removed with RNase A.

1.3 Cloning and sequencing of the *gvp* and *rbcLX* regions

The *gvp* and *rbcLX* regions were amplified *in vitro* by polymerase chain reactions (PCR) with DNA extracted from cultured cells or water bloom samples using primers listed in Table 2. A mixture of *Taq* and *Pfu* (Fermentas, Vilnius, Lithuania) (1∶1) DNA polymerases was used in the PCR reactions. DNA fragments were cloned in a T-vector (Takara, Dalian, China) and sequenced. Sequences were deposited in the NCBI GenBank under accession numbers listed in Table 1.

Table 2. PCR primers in this study

Primer	Sequence (5′→3′)	Genomic region
gvpA1-down	AGCACAAAACTGAGAATTGACTAC	*gvpA* to *gvpC* (FACHB 854, 916, 927, 930)
gvpC-up	AAGCAGGATCAGCATGGTAGG	
gvp-18	CCTACCATGCTGATCCTGCTT	*gvpN* to *gvpX* (FACHB 854)
gvpX-1	TTGCCATCTCCCTTGCTCT	
gvpX-down	GAGTCCTCAAATGGTCAAGAAGT	*gvpX* to *gvpW* (FACHB 854)
gvpW-up	GATCAATCAAGTTAAGTAAAGTGATAGG	
910AC-2	TGGGTAGTTCGCCATCTCA	*gvpA* to *gvpC* (FACHB 910)
gvpCpri	GGAATTAACCTAGAGATTTATTTGCCA	
gvpA5a	GCG/CGAAGTGATGAAGGCA	*gvpA* to *gvpC* (FACHB 929)
gvpCpri	GGAATTAACCTAGAGATTTATTTGCCA	
gvpC-930F	AGGATCCCATGCCTGCTCTCATC	*gvpC* (FACHB930, CHAO, TAI)
gvpC-930R	ACTCGAGATCCTTCACCTGTTTGGCTC	
Mrbc-1	TCTTCGGTGACGATTCCTGCT	*rbcLX*
Mrbc-2	TCTTGATTCCTTATGGGTTGGGT	

1.4 Phylogenetic analyses

Sequence alignments and analyses were performed with the software packages Clustal W 1.83[17] and MEGA 3.1[18], and phylogenetic trees were reconstructed using the neighbor-joining (NJ) method[19] with 1000 bootstrap replicates. In addition to the sequences listed in Table 1, the *gvpA-gvpC* regions of *Microcystis* sp. PCC 7806 (AJ577136) and BC 8401 (AY965344) and *rbcLX* of PCC 7806 (AM157793) were used in the analyses.

2 Results and discussion

2.1 Classes of *gvpA-gvpC* regions in *Microcystis* strains

To analyze the diversity of the *gvp* gene cluster in *Microcystis*, we attempted to sequence this region in different strains. Using PCR, we first cloned and sequenced the entire *gvp* cluster of strain FACHB 854 (*gvpA1A2CNJXKFGVW*, GenBank accession

No. AY254709). A comparison of entire *gvp* gene clusters from *Microcystis* strains PCC 7806, BC 8401, FACHB 854 and *Anabaena* / *Nostoc* sp. PCC 7120 showed that the *gvpA-gvpC* region is variable among species or strains, while the rest part (*gvpN-JXKFGVW*) is relatively conserved. Therefore, we cloned the *gvpA-gvpC* regions of *Microcystis* sp. strains FACHB 910, 916, 927, 929 and 930 by PCR using primers listed in Table 1. We also cloned the *gvpA-gvpC* region directly with *Microcystis* water-bloom samples from the Lake Chao and Lake Tai in China. We denoted the uncultured *Microcystis* cells carrying the corresponding sequences as CHAO (Lake Chao) and TAI (Lake Tai) respectively.

The *gvpA* genes from *Microcystis* strains are identical to each other at the amino acid level, but differ in copy number: 2 copies in BC 8401[10], FACHB 854, CHAO and FACHB 910, 3 copies in PCC 7806[15], TAI, FACHB 916, 927 and 930 and 4 copies in FACHB 929. The *gvpC* genes are divided into 2 types according to their sizes: those of CHAO, TAI and strains FACHB 910, 916, 927 and 930 are 214-codon long with one more 33RR-encoding sequence than those of BC 8401, FACHB 854 and 929 (Fig. 1). Moreover, there are several substitutions among predicted GvpC proteins of *Microcystis* strains.

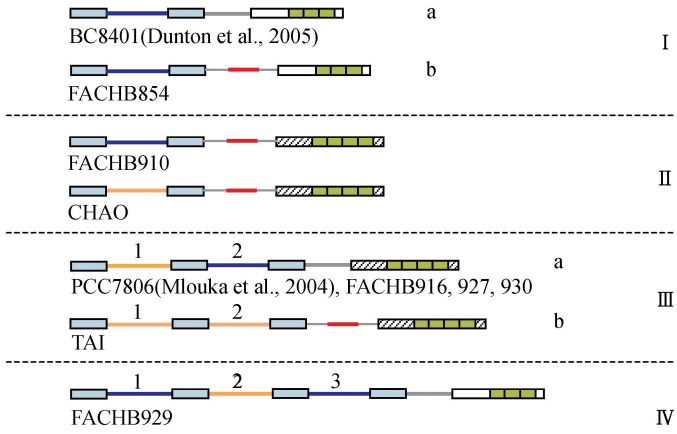

Fig. 1. Four structural types of the *gvpA-gvpC* regions from *Microcystis* strains. Blue-grey boxes, *gvpA* genes; white boxes hatched or not, *gvpC*; green boxes in *gvpC*, the 33RR-encoding sequences; red thick solid lines stand for the extra sequence tag of 172—176 bp. The two types of A-A intergenic segments are indicated by orange or blue lines, and the two types of A-C intergenic segments are indicated by thick or thin lines. Different types of the intergenic segments and *gvpC* correspond to cluster A or B in Figs. 2—4.

Based on the copy numbers of *gvpA* genes and the size of *gvpC*, the *gvpA-gvpC* regions of the 8 purified strains and 2 uncultured samples of *Microcystis* could be divided into 4 classes: (I) 2*gvpA* + *gvpC*(3); (II) 2*gvpA* + *gvpC*(4); (III) 3*gvpA* + *gvpC*(4); (IV) 4*gvpA* + *gvpC*(3). The numerals before *gvpA* are the copy numbers, and the ones in parentheses following *gvpC* are numbers of 33RR-encoding sequence. Furthermore, in *gvp* sequences of CHAO and TAI and those of strains FACHB 854 and 910, the intergenic segments between *gvpA* and *gvpC* contain an extra sequence of 172 to 176 bp, which was not found in other strains (Fig. 1). Accordingly each of classes I and III could be further divided into two subclasses based on the absence (subclass Ia, BC 8401; subclass IIIa, PCC 7806, FACHB 916, 927 and 930) or presence (subclass Ib, FACHB 854; subclass IIIb, TAI) of this extra sequence (Fig. 1). We noticed that all strains with this extra sequence were collected from China, namely the Lake Dianchi in southern China, the Lake Chao and Lake Tai both in eastern China and the Lake Wudalianchi in northern China.

2.2 Phylogenetic analyses of the A-A and A-C intergenic segments and *gvpC*

To analyze the relationships of different structural classes, we reconstructed phylogenetic trees with each of *gvpC* genes and the intergenic segments between *gvpA* and *gvpA* (A-A) or between *gvpA* and *gvpC* (A-C). Although *gvp* genes are quite conserved at the amino acid level in different groups of cyanobacteria, they show no significant similarity at the nucleotide sequence level. Hence, no outgroup was included in the A-A, A-C and *gvpC* trees.

A-A segments of *Microcystis* are grouped in two major clusters A and B (Fig. 2). The two A-A segments of subclass IIIa belong to the two clusters respectively, while both of subclass IIIb belong to clus-

ter B. In the class IV strain, the first A-A segment is closely related to the third, which could have resulted from a duplication event during the evolution. Similarly, a duplication of *gvpA* could have been involved in the origin of the *gvpA-gvpC* region of TAI from that of CHAO. To reconstruct the phylogeny of A-C intergenic segments, we deleted the extra sequences of 172 to 176 bp from those of CHAO, TAI and strains FACHB854 and 910. Even so, these four segments form a separate cluster from others in the phylogenetic tree (Fig. 3). To reconstruct the phylogeny of *gvpC*, we deleted an extra 33RR-encoding sequence from those of class II and class III strains. As shown in Fig. 4, *gvpC* genes of these two classes form cluster A, while those of classes I and IV form cluster B.

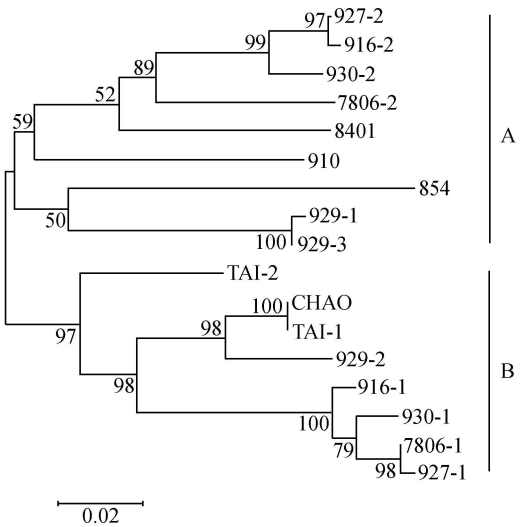

Fig. 2. A dendrogram of *gvpA-gvpA* intergenic segments. Segment types 1, 2 and 3 in some strains are indicated in Fig. 1. Local bootstrap support for branches present in more than 50% of 1000 resamplings is indicated at the relevant nodes. The scales are in units of nucleotide substitutions per site.

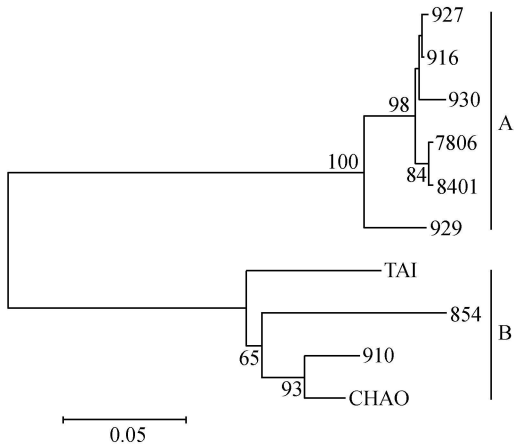

Fig. 3. A dendrogram of *gvpA-gvpC* intergenic segments. Local bootstrap support and the scales are as in Fig. 2.

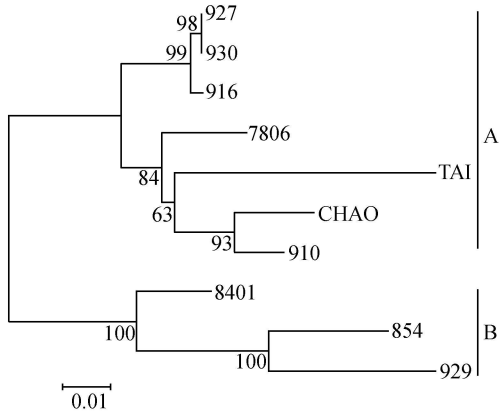

Fig. 4. A dendrogram of *gvpC* genes. Local bootstrap support and the scales are as in Fig. 2.

All the three phylogenetic trees possess two major clusters. Accordingly, all A-A and A-C segments and *gvpC* genes can be divided into two types. The diversity of *gvpA-gvpC* regions is mostly due to combinations of different types of the intergenic segments and *gvpC* (Fig. 1), which could be interpreted as results of interstrain recombinations. Taking two strains with the extra sequence tag, FACHB 854 and FACHB 910, for an example, they are located in the same cluster in A-A and A-C trees; however, the *gvpC* gene of FACHB 910 is very close to that of subclass IIIa strain rather than that of FACHB 854 (Fig. 4). Because the phylogenetic tree for *gvpC* was constructed after removal of a 33RR-encoding sequence from *gvpC* (4), the difference between *gvpC* genes of FACHB 854 and FACHB 910 was not due to an intragenic rearrangement, namely, a duplication of the 33RR-encoding sequence. Instead, the difference of *gvpC* type in the two strains should be attributed to an interstrain recombination, that is, the replacement of *gvpC* (3) in a subclass Ib strain by *gvpC* (4), resulting in FACHB 910, or the replacement of *gvpC* (4) in FACHB 910 by *gvpC* (3), resulting in a subclass Ib strain.

2.3 *Microcystis* strains are tightly clustered in the phylogenetic tree of *rbcLX* sequence

We cloned, sequenced a 790-bp fragment overlapping partial *rbcL*, the intergenic segment and partial *rbcX* from strains FACHB 854, 910, 916, 927, 929 and 930 and reconstructed the NJ tree of these *rbcLX* sequences and that of PCC 7806 using the *rbcLX* sequence of *Synechocystis* sp. PCC 6803 as an outgroup. We also reconstructed the maximum parsimony (MP), maximum-likelihood (ML) (using PHYLIP 3.65) and Bayesian (using MrBayes

3.1.2) trees with these sequences, or reconstructed the NJ tree using an alternative outgroup (such as the *rbcLX* sequence from *Anabaena* sp. PCC 7120), and found that the overall topology was similar to the one in Fig. 5 (data not shown). *Microcystis* strains are tightly clustered in all the *rbcLX* trees.

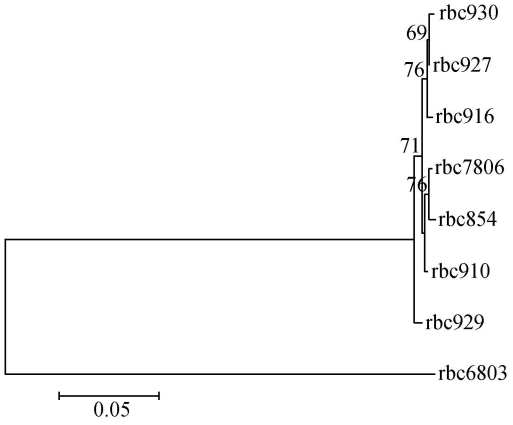

Fig. 5. A dendrogram of *rbcLX*. Sequences rbc854, rbc910, rbc916, rbc927, rbc929, rbc930, rbc6803 and rbc7806 are the partial *rbcLX* from strains FACHB 854, 910, 916, 929, 930, PCC 6803 and 7806 respectively. The *rbcLX* of *Synechocystis* PCC 6803 was used as an outgroup. Local bootstrap support and the scales are as in Fig. 2.

Compared to the sequences of 16S rRNA, 16S to 23S internal transcribed spacer, phycocyanin intergenic spacer[20, 21] and *rbcLX*, the *gvpA-gvpC* region is much more variable. Of the ten full length *gvpA-gvpC* region sequences in Fig. 1, only four possess the same structure (subclass IIIa), while all the other six differ from each other and from these four sequences. The diversity could have been caused by duplication, insertion and interstrain recombination of DNA fragments. Transposable elements-induced rearrangements were often found in the *gvp* cluster of *Microcystis* under laboratory conditions[15, 22], but it has not been reported for an ecotype in the wild. Based on the 16S rRNA sequences, the 5 to 6 morphospecies of *Microcystis*[23, 24] should probably be integrated into one species[25]. Based on the 16S-23S ITS sequences, however, 47 strains of these morphospecies formed 3 clusters in phylogenetic trees, and the tree topologies were not necessarily correlated with morphospecies distinction[26]. More recently, all the morphospecies were unified into one under the rules of the Bacteriological Code[27]. To facilitate strain identification in ecological or bio-geographical studies of *Microcystis*, a powerful molecular tool should be developed. Our results showed that the *gvpA-gvpC* region could provide sufficient diversity for this purpose.

References

1 Walsby AE. Gas vesicles. Microbiol Rev, 1994, 58: 94—144
2 Weathers PJ, Jost M and Lamport DT. The gas vacuole membrane of *Microcystis aeruginosa*. A partial amino acid sequence. Arch Biochem Biophys, 1977, 178: 226—244
3 Walker JE and Walsby AE. Molecular weight of gas-vesicle protein from the planktonic cyanobacterium *Anabaena flos-aquae* and implications for structure of the vesicle. Biochem J, 1983, 209: 809—815
4 Damerval T, Castets AM, Houmard J, et al. Gas vesicle synthesis in the cyanobacterium *Pseudanabaena* sp.: occurrence of a single photoregulated gene. Mol Microbiol, 1991, 5: 657—664
5 Hayes PK, Lazarus CM, Bees A, et al. The protein encoded by *gvpC* is a minor component of gas vesicles isolated from the cyanobacteria *Anabaena flos-aquae* and *Microcystis* sp. Mol Microbiol, 1988, 2: 545—552
6 Walsby AE and Hayes PK. The minor cyanobacterial gas vesicle protein, GVPc, is attached to the outer surface of the gas vesicle. J Gen Microbiol, 1988, 134: 2647—2657
7 Hayes PK, Buchholz B and Walsby AE. Gas vesicles are strengthened by the outer-surface protein, GvpC. Arch Microbiol, 1992, 157: 229—234
8 Kinsman R, Walsby AE and Hayes PK. GvpCs with reduced numbers of repeating sequence elements bind to and strengthen cyanobacterial gas vesicles. Mol Microbiol, 1995, 17: 147—154
9 Bright DI and Walsby AE. The relationship between critical pressure and width of gas vesicles in isolates of *Planktothrix rubescens* from Lake Zurich. Microbiology, 1999, 145: 2769—2775
10 Dunton PG and Walsby AE. The diameter and critical collapse pressure of gas vesicles in *Microcystis* are correlated with GvpCs of different length. FEMS Microbiol Lett, 2005, 247: 37—43
11 Beard SJ, Davis PA, Iglesias-Rodriguez D, et al. Gas vesicle genes in *Planktothrix* spp. from Nordic lakes: strains with weak gas vesicles possess a longer variant of *gvpC*. Microbiology, 2000, 146: 2009—2018
12 Damerval T, Houmard J, Guglielmi G, et al. A developmentally regulated *gvpABC* operon is involved in the formation of gas vesicles in the cyanobacterium *Calothrix* 7601. Gene, 1987, 54: 83—92
13 Kinsman R and Hayes PK. Genes encoding proteins homologous to halobacterial GvpN, J, K, F and L are located downstream of *gvpC* in the cyanobacterium *Anabaena flos-aquae*. DNA Seq, 1997, 7: 97—106
14 Albouy D, Castets AM and Tandeau de Marsac N. The gas vesicle gene (*gvp*) cluster of the cyanobacterium *Pseudanabaena sp.* Strain PCC 6901. DNA Seq, 2001, 12: 337—344
15 Mlouka A, Comte K, Castets AM, et al. The gas vesicle gene cluster from *Microcystis aeruginosa* and DNA rearrangements that lead to loss of cell buoyancy. J Bacteriol, 2004, 186: 2355—2365
16 Rippka R, Deruelles J, Waterbury JB, et al. Generic assignments, strain histories and properties of pure cultures of cyanobacteria. J Gen Microbiol, 1979, 111: 1—61
17 Thompson JD, Higgins DG and Gibson TJ. improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. Nucl Acids Res, 1994, 22: 4673—4680
18 Kumar S, Tamura K and Nei M. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform, 2004, 5: 150—163

19 Saitou N and Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol, 1987, 4: 406—425
20 Otsuka S, Suda S, Li R, et al. Phylogenetic relationships between toxic and non-toxic strains of the genus *Microcystis* based on 16S to 23S internal transcribed spacer sequence. FEMS Microbiol Lett, 1999, 172: 15—21
21 Tillett D, Parker DL and Neilan BA. Detection of toxigenicity by a probe for the microcystin synthetase A gene (*mcyA*) of the cyanobacterial genus *Microcystis*: comparision of toxicities with 16S rRNA and phycocyanin operon (phycocyanin intergenic spacer) phylogenies. Appl Environ Microbiol, 2001, 67: 2810—2818
22 Xu M, Kong R and Xu X. A miniature insertion element transposable in *Microcystis* sp. FACHB 854. Progress in Natural Science, 2006, 16: 486—491
23 Komarek J. A review of water-bloom forming *Microcystis* species, with regard to populations from Japan. Arch Hydrobiol Suppl Algological Studies, 1991, 64: 115—127
24 Watanabe M. Isolation, cultivation, and classification of bloom-forming *Microcystis* in Japan. In: Toxic Microcystis. FL: CRC Press, 1996, 13—34
25 Otsuka S, Suda S, Li R, et al. 16S rDNA sequences and phylogenetic analyses of *Microcystis* strains with and without phycoerythrin. FEMS Microbiol Lett, 1998, 164: 119—124
26 Otsuka S, Suda S, Li R, et al. Phylogenetic relationships between toxic and non-toxic strains of the genus *Microcystis* based on 16S to 23S internal transcribed spacer sequence. FEMS Microbiol Lett, 1999, 172: 15—21
27 Otsuka S, Suda S, Shibata S, et al. A proposal for the unification of five species of the cyanobacterial genus *Microcystis* Kutzing ex Lemmermann 1907 under the rules of the Bacteriological Code. Int J Syst Evol Microbiol, 2001, 51: 873—879