Microbes Environ. Vol. 25, No. 4, 295–301, 2010
http://wwwsoc.nii.ac.jp/jsme2/ doi:10.1264/jsme2.ME10125

Microbes and Environments

The Distribution of a Phage-Related Insertion Sequence Element in the Cyanobacterium, *Microcystis aeruginosa*

SOTARO KUNO[1], TAKASHI YOSHIDA[1]*, RYOMA KAMIKAWA[2], NAOHIKO HOSODA[3], and YOSHIHIKO SAKO[1]

[1]*Department of Agriculture, Kyoto University, Kitashirakawa-Oiwake, Sakyo-ku, Kyoto 606–8502, Japan;* [2]*Center for Computational Sciences and Institute of Biological Sciences, University of Tsukuba, 1–1–1 Tennoudai, Tsukuba, Ibaraki 305–8577, Japan; and* [3]*Department of Marine Bioscience, Fukui Prefectural University, 1–1 Gakuen-cho, Obama, Fukui 917–0003, Japan*

(Received April 12, 2010—Accepted July 27, 2010—Published online August 21, 2010)

The cyanophage Ma-LMM01, specifically-infecting *Microcystis aeruginosa*, has an insertion sequence (IS) element that we named IS607-cp showing high nucleotide similarity to a counterpart in the genome of the cyanobacterium *Cyanothece* sp. We tested 21 strains of *M. aeruginosa* for the presence of IS607-cp using PCR and detected the element in strains NIES90, NIES112, NIES604, and RM6. Thermal asymmetric interlaced PCR (TAIL-PCR) revealed each of these strains has multiple copies of IS607-cp. Some of the ISs were classified into three types based on their inserted positions; IS607-cp-1 is common in strains NIES90, NIES112 and NIES604, whereas IS607-cp-2 and IS607-cp-3 are specific to strains NIES90 and RM6, respectively. This multiplicity may reflect the replicative transposition of IS607-cp. The sequence of IS607-cp in Ma-LMM01 showed robust affinity to those found in *M. aeruginosa* and *Cyanothece* spp. in a phylogenetic tree inferred from counterparts of various bacteria. This suggests the transfer of IS607-cp between the cyanobacterium and its cyanophage. We discuss the potential role of Ma-LMM01-related phages as donors of IS elements that may mediate the transfer of IS607-cp; and thereby partially contribute to the genome plasticity of *M. aeruginosa*.

Key words: cyanobacteria, *Microcystis*, cyanophage, insertion sequence, horizontal gene transfer

Cyanobacteria are principal primary producers in marine and freshwater ecosystems. They are susceptible to infection by a variety of viruses (cyanophages) that significantly affect their mortality and population structure; and influence the aquatic food web and biogeochemical cycles (11, 30, 37). Recent studies show cyanophages may also play a role in the vast gene pool of cyanobacterial communities where many of the marine cyanophages have genes essential for photosynthesis (27, 34); and these 'host-like' genes (28) facilitate energy and carbon production during long infections (23). It is believed the cyanophages have co-evolved with their hosts through horizontal gene transfer (HGT) events and have shaped part of the cyanobacterial genomes (7, 23).

Microcystis aeruginosa is a major bloom-forming unicellular cyanobacterium found in eutrophic freshwater environments. Some strains of this species produce cyclic peptide toxins called microcystins that cause acute poisoning of humans and livestock (6, 40). Some of us previously isolated a lytic myovirus Ma-LMM01 infecting a toxic strain of *M. aeruginosa* (38, 43, 44). The genome of Ma-LMM01 has 162,109 bp and contains 184 predicted protein-coding genes. The majority of the predicted genes have no detectable homologues in the present databases and the genome showed no co-linearity with previously sequenced genomes of other *Myoviridae* including cyanomyoviruses; and thus Ma-LMM01 was assigned as a member of a new lineage of the *Myoviridae* family (45). Unlike most of the marine cyanophages, Ma-LMM01 has no homologues of the photosynthesis genes. Instead, the genome harbors several 'host-like' genes including *nblA* involved in the maintenance of the photosynthetic reaction center (45). Of the 'host-like' genes in Ma-LMM01, a region of ca. 1,800 bp including ORF135 and ORF136 shows high nucleotide sequence similarity (82% identity) with a partial sequence of the *Cyanothece* sp. CCY0110 genome (GenBank: NZ_AAXW01000043) suggesting this region has experienced HGT between Ma-LMM01-related phages and cyanobacteria (45).

To gain insights into the co-evolution of cyanophages and the host cyanobacteria, the presence of the genomic region of Ma-LMM01 including ORF135 and ORF136 was examined within 21 strains of the host cyanobacterium *M. aeruginosa* using PCR. Further, we assessed the possibility of HGT of the region by constructing a phylogenetic tree based on nucleotide sequences of one of the two ORFs encoded in each of the regions.

Materials and Methods

Maintenance of cyanobacterial strains and DNA extraction

The strains of *M. aeruginosa* used in this study are listed in Table 1. All strains with the prefix NIES were obtained from the National Institute for Environmental Studies (Tsukuba, Japan). Strains LMM- and MMY- were isolated from Lake Mikata in previous studies (19, 41). Strains RMK1 and RM6 were isolated from Lake Reisenji in 2006. All the strains except MMY52, RMK1 and RM6 were axenic. The strains were maintained in CB medium (16) under 12:12 LD conditions (ca. 40 μmol photons $m^{-2}s^{-1}$) at 30°C. Exponentially growing cultures were mildly sonicated to remove gas vesicles; and then cells were harvested by centrifugation at 3,000

* Corresponding author. E-mail: yoshiten@kais.kyoto-u.ac.jp; Tel: +81–75–753–6218; Fax: +81–75–753–6226.

Table 1. *M. aeruginosa* strains used in this study

Strains	Origins	Morphotypes[1]	ST[2]	IS607-cp[3]
NIES90	Lake Kawaguchi	aer	5	+
NIES91	Lake Kasumigaura	aer	77	−
NIES98	Lake Kasumigaura	aer	6	−
NIES99	Lake Suwa	aer	7	−
NIES101	Lake Suwa	aer	9	−
NIES102	Lake Kasumigaura	vir	10	−
NIES112	Lake Suwa	wes	16	+
NIES298	Lake Kasumigaura	aer	60	−
NIES604	Lake Kasumigaura	wes	17	+
NIES843	Lake Kasumigaura	aer	18	–[4]
NIES1058	Lake Kasumigaura	vir	28	−
NIES1067	Chikatou Pond	wes	27	−
NIES1070	Rokusuke Pond	aer	36	−
NIES1086	Koshi Pond	aer	ND[5]	−
NIES1098	Shigure Dam	aer	38	−
NIES1105	Lake Barato	aer	31	−
LMM9508-11	Lake Mikata	unidentified	ND	−
LMM9509-2	Lake Mikata	unidentified	ND	−
RM6	Lake Reisenji	unidentified	ND	+
RMK1	Lake Reisenji	unidentified	ND	−
MMY52	Lake Mikata	unidentified	ND	−

[1] Morphotypes were defined in accordance with a previous study (42) as follows: aer, aeruginosa complex-type, vir, viridis-type and wes, wesenbergii-type.
[2] ST, sequence types determined in a previous study (39).
[3] Results of PCR are indicated as: +, detected, −, not detected.
[4] We determined the IS sequence is absent from the complete genome.
[5] ND: not determined.

×*g* for 10 min. The total genomic DNA was extracted using the cetyltrimethylammonium bromide (CTAB) method (14); and the DNA was suspended in 50 μL of TE buffer.

PCR detection and sequencing of ORF 135 and ORF136 from M. aeruginosa

To amplify the fragments of approx. 1,300 bp corresponding to part of ORF135 and ORF136 in the Ma-LMM01 genome, primers 136F and 135R were designed (Fig. 1 and Table 2). The PCR was performed in a total volume of 20 μL containing 2 ng of template DNA, 3 μM primers, 250 μM each dNTP, 1×*Ex Taq* Buffer and 0.4 U of Takara *Ex Taq* polymerase (Takara Bio, Otsu, Japan). The reaction conditions were: 4 min of initial denaturing at 94°C followed by 30 cycles: 94°C for 30 s, 58°C for 1 min and 72°C for 1.5 min; and a final extension at 72°C for 7 min. The PCR products were then separated by electrophoresis on 1.0% (w/v) agarose gels. The gels were stained using ethidium bromide and visualized using the Gel Doc XR system (Bio-Rad Laboratories, Hercules, CA, USA). Visually confirmed bands were excised and purified with a Wizard Miniprep Purification Kit (Promega, Madison, WI, USA) according to the manufacturer's instructions. The purified PCR products were cloned into the pGEM-T Easy Vector (Promega) and both strands were sequenced at the Dragon Genomics Center, Takara Bio. A total of nine clones per strain were sequenced.

TAIL-PCR amplification of flanking sequences to determine the location in the genome

TAIL-PCR effectively amplifies DNA fragments adjacent to a known sequence using combinations of specific outward primers and arbitrary nonspecific primers (25). Based on the sequence determined above, we designed a total of nine specific primers for the TAIL-PCR (Fig. 1 and Table 2). Additionally, we used six arbitrary primers, AD1-AD6 (Table 2), based on previous studies (24, 25). The detailed three-step reactions of TAIL-PCR are shown in Table 3. Primary PCR was performed in 20 μL containing 30 ng

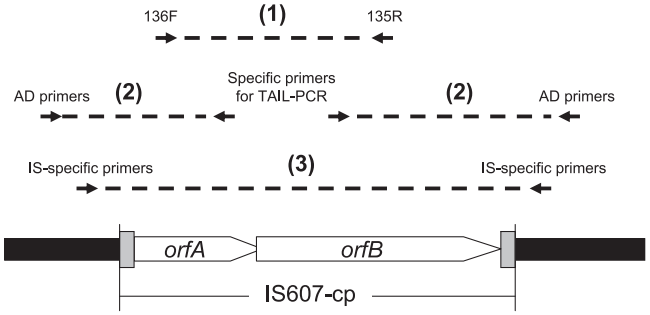

Fig. 1. Schematic representation of the PCR amplification strategy. Open arrows indicate *orfA* and *orfB* of IS607-cp. The closed boxes indicate the part of the genomic sequence of *M. aeruginosa* where IS607-cp is inserted. The terminal regions of IS607-cp are represented by gray boxes. (1) The presence of IS607-cp was tested using PCR amplification with primers specific for ORF135 and OF136. (2) Flanking sequences including a terminus for IS607-cp were amplified using TAIL-PCR with combinations of specific primers and AD primers. (3) IS607-cp-1 and IS607-cp-2 were amplified from the flanking sequence where the whole lengths were sequenced.

of template DNA, 0.2 μM specific primer, 3.5 μM arbitrary primer, 200 μM dNTPs, 1×*Ex Taq* Buffer and 1 U of Takara *Ex Taq* polymerase. One microliter of a 50-fold dilution of the primary PCR products was added to the secondary PCR solution (20 μL) containing 0.2 μM specific primer, 2.5 μM arbitrary primer used in the primary reaction, 200 μM dNTPs, 1×*Ex Taq* Buffer and 0.8 U of Takara *Ex Taq* polymerase. The tertiary reaction was performed in 100 μL containing 1 μL of a 10-fold dilution of the secondary PCR products, 0.3 μM specific primer, 4 μM arbitrary primer used in the former PCR, 200 μM dNTPs, 1×*Ex Taq* Buffer and 3 U of Takara *Ex Taq* polymerase. The resulting PCR products were purified, cloned and sequenced as described above. A total of 94 clones were sequenced including 14, 5, 7 and 10 upstream sequences flanking *orfA* and 16, 10, 18 and 14 downstream sequences flanking *orfB* for strains NIES90, NIES112, NIES604 and RM6, respectively. The flanking sequences obtained were then subjected to a similarity search in the genomic sequence of *M. aeruginosa* strains NIES843 (GenBank: AP009552) and PCC7806 (Genbank: AM778931) using BLASTN (1). Each region was amplified in full using combinations of specific primers (IS1F and IS1R, IS2F and IS2R and IS3F and IS3R; Fig. 1 and Table 2); and then sequenced using the primer walking strategy (primers not shown). Note that the primer IS3R was designed to match a part of *orfB* since the downstream flanking sequence of IS607-cp-3 was unknown.

Phylogenetic analysis

Twenty-one of the *orfA* sequences located in IS607-cp (see below) and its homologues were aligned where the ambiguously aligned sites were excluded and this resulted in 399 sites. Prepared dataset was subjected to phylogenetic analyses. The ML phylogenetic analysis was performed by using RAxML 7.2.1 (36) under the general time reversible (GTR) model with a gamma distribution with four categories (GTR+Γ model). The initial tree search was started from 10 distinct parsimony starting trees and topologies were evaluated by ML with the GTR+Γ model. The optimal tree among the 10 final trees was then selected as the ML tree. A bootstrap (BP) analysis (100 replicates) was conducted using RAxML 7.2.1 as described above. A Bayesian analysis was performed by using MrBayes 3.1.2 (31) under the GTR+Γ model with a proportion of invariable sites. One cold and three heated Markov chain Monte Carlo chains with default chain temperatures were run for 1,000,000 generations, with the sampling of log-likelihood values and trees at 100-generation intervals. The first 10,000 generations (*i.e.*, 100 trees) were discarded as burn-in, and the consensus tree with branch lengths and Bayesian posterior probabilities (BPP) were calculated from the rest of the sampling points.

Table 2. Primers used in this study

Primer names	Sequences (5' to 3')	Purposes
136F	AGACGTTGGGAGTCTGAAGG	detection of IS607-cp
135R	GTGGTTCTTGGCTAAATAAGTGGTG	detection of IS607-cp
136R	CATCTTGCGCCAAATCCTCCTC	detection of IS607-cp
leftA2	GTAGTTAAGTCCTGAACCTAAGCG	specific primer for TAIL-PCR
leftB	CAACTTGCCAACCGTTTTGAGC	specific primer for TAIL-PCR
leftC	TCTTTTACTCCAAGCAACTGAG	specific primer for TAIL-PCR
leftC2	ACCTTCAGACTCCCAACGTC	specific primer for TAIL-PCR
rightA	CTTACGAGCGGCTACCTGATGG	specific primer for TAIL-PCR
rightB1	GCTACCTGATGGATACCAGC	specific primer for TAIL-PCR
rightB2	CTCATGATGTAGTTGGTTGTGATCTGG	specific primer for TAIL-PCR
rightB3	CTTTAGCTACTTTATCGACAGG	specific primer for TAIL-PCR
rightC	GAAGGCTCTAAAAGTTATCG	specific primer for TAIL-PCR
AD1	NGTCGASWGANAWGAA	arbitrary primer for TAIL-PCR
AD2	GTNCGASWCANAWGTT	arbitrary primer for TAIL-PCR
AD3	WGTGNAGWANCANAGA	arbitrary primer for TAIL-PCR
AD4	TGWGNAGWANCASAGA	arbitrary primer for TAIL-PCR
AD5	AGWGNAGWANCAWAGG	arbitrary primer for TAIL-PCR
AD6	NTCGASTWTSGWGTT	arbitrary primer for TAIL-PCR
IS1F	CACTTGACCAATAGCCTGAG	amplification of IS607-cp-1
IS1R	CATTTTACCGAAACTTCGCC	amplification of IS607-cp-1
IS2F	GTCAGATTCGATAATAGGGTTC	amplification of IS607-cp-2
IS2R	AAATTACCCAACGGGGAAC	amplification of IS607-cp-2
IS3F	CCGTTCTGTTAAACTTACTCAAAT	amplification of IS607-cp-3
IS3R	GTCGGATGTCTTGGTTAAAATG	amplification of IS607-cp-3

Table 3. Conditions for the three-step TAIL-PCR.

Reactions[1]	Program nos.	Numbers of cycles	Cycle parameters
Primary PCR	1	1	94°C, 2 min
(primers A)	2	5	94°C, 1 min; 65°C, 1 min; 72°C, 3 min
	3	1	94°C, 1 min; 30°C, 3 min; 72°C, 3 min
	4	15	94°C, 30 s; 68°C, 1 min; 72°C, 3 min
			94°C, 30 s; 68°C, 1 min; 72°C, 3 min
			94°C, 30 s; 44°C, 1 min; 72°C, 3 min
	5	1	72°C, 5 min
Secondary PCR	6	12	94°C, 30 s; 64°C, 1 min; 72°C, 3 min
(primers B)			94°C, 30 s; 64°C, 1 min; 72°C, 3 min
			94°C, 30 s; 44°C, 1 min; 72°C, 3 min
	7	1	72°C, 5 min
Tertiary PCR	8	20	94°C, 1 min; 44°C, 1 min; 72°C, 3 min
(primers C)	9	1	72°C, 5 min

[1] Specific primers used in each reaction are indicated in parentheses.

Nucleotide sequence accession numbers

The nucleotide sequences determined in this study were deposited in the DDBJ/EMBL/GenBank database. The accession numbers are as follows: AB543092 to AB543096 for the whole lengths of IS607-cp-1 to -3, and AB543097 to AB543109 for each type of IS607-cp-flanking sequence.

Results and Discussion

Re-annotation of the two ORFs in the Ma-LMM01 genome

A homologous region shared between the complete genomes of *M. aeruginosa* phage Ma-LMM01 and *Cyanothece* sp. CYY0110 includes two ORFs corresponding to ORF135 and ORF136 of Ma-LMM01 that have been previously annotated as a transposase and a putative site-specific recombinase, respectively (45). Nevertheless, we considered the region to be an IS element because ORF136 contains features of a transposase encoded in IS607 family elements: a DNA-binding domain with a helix-turn-helix motif and a catalytic domain belonging to the IS607-like transposase subfamily at the N-terminal and C-terminal, respectively (35). We performed a survey using BLASTP (1) with the ORF136 and ORF135 amino acid sequences as queries in the IS Finder database (http://www-is.biotoul.fr/is.html). As expected, ORF136 and ORF135 showed the highest sequence similarity to OrfA and OrfB of the IS607 family members—up to 78% and 50%, respectively. IS607 is a chimeric IS element originally found in certain strains of *Helicobacter pylori* (18). In general, IS607 contains two genes; the first one (*orfA*) encodes a putative transposase classified into the serine recombinase family of site-specific recombinase; and the second (*orfB*) is a putative transposase gene. Thus, we re-annotated the region containing ORF135 and ORF136 as a variant of the IS607 family members although its transposition activity remains unclear. Curiously, the complete genome sequence of *M. aeruginosa* NIES843 contains several putative IS607 family elements (15). However, the IS607 family elements in Ma-LMM01 and *Cyanothece* sp. CCY0110 exhibit weak amino acid sequence similarity to those of *M. aeruginosa* NIES843 (at most 31% sequence

identity for OrfA). Thus, we here designate IS607 of Ma-LMM01 and its closely-related elements as "IS607-cp" to discriminate them from other IS607 family elements.

Distribution of IS607-cp in M. aeruginosa

Many IS elements are reported to have patchy distribution in bacterial lineages (13, 26, 33). Therefore, although the *M. aeruginosa* NIES843 genome lacks IS607-cp as described above, we investigated the possibility that other *M. aeruginosa* strains may possess homologous elements. Of the 21 strains examined (Table 1), the PCR-based surveys for IS607-cp with the primers 136F and 135R (Table 2) yielded a single band of expected size (ca. 1,300 bp) for four strains (NIES90, NIES112, NIES604 and RM6); whereas those for the other 17 strains failed to produce any observable bands on 1.0% agarose gel electrophoresis (data not shown). The same result was obtained using another combination of primers (136F and 136R; Table 2). The sequences of the PCR products showed a high level of nucleotide identity to the corresponding part of IS607-cp in the Ma-LMM01 genome (98%) indicating these four strains have homologues of IS607-cp in their genomes. The nine PCR clones originating from strain NIES90 apparently bear single nucleotide polymorphisms at several positions suggesting this strain has multiple copies of IS607-cp in its genome. To confirm the locations and numbers of IS607-cp in the genomes of the four strains, we determined the flanking sequences of IS607-cp using TAIL-PCR. The flanking sequences obtained were classified into types on the basis of sequence similarity as follows: the upstream sequences obtained from NIES112 and RM6 each fell into one type, and those from NIES90 and NIES604 were each separated into two types. Downstream flanking sequences were separated into 2, 1, 1 and 3 type(s) for NIES90, NIES112, NIES604 and RM6, respectively. This suggested the four strains contain at least 2, 1, 2 and 3 copies of IS607-cp, respectively. Kersulyte *et al.* (18) suggested a model for the transposition mechanism for IS607 which is frequently replicative; thus the multiplicity of IS607-cp observed may represent the signature of transposition events following the invasion of IS607-cp into the genomes. The discrepancy in the number of sequence types between the two directions obtained from one strain was possibly due to the bias of the TAIL-PCR amplification using the arbitrary primers. Several types of flanking sequences showed significant similarity to the genomic sequence of *M. aeruginosa* strains NIES843 and PCC7806, therefore we could determine the insertion loci for a subset of IS607-cp. They were separated into three distinct ISs based on the putative insertion loci: IS607-cp-1, IS607-cp-2 and IS607-cp-3 (Fig. 2A). IS607-cp-1 were amplified from the strains NIES90, NIES112 and NIES604; and located at a homologous site between genes encoding a hypothetical protein and NAD-dependent epimerase/dehydratase (corresponding to IPF_1689 and IPF_1688 of *M. aeruginosa* PCC7806, respectively). The IS607-cp-2 from NIES90 was found within a gene encoding a hypothetical protein (MAE_25900 of *M. aeruginosa* NIES843). IS607-cp-3 from the strain RM6 was detected 15 bp downstream from the Holliday junction resolvase-like protein gene (MAE_50370); although we could not determine the 3' flanking sequence (see below). These sequences surrounding the insertion sites for the three ISs were not found in the Ma-LMM01 genome. One of the upstream flanking sequences obtained from strain NIES604 contained a sequence nearly identical to a part of IS607-cp itself suggesting the possibility that an IS607-cp inserted into the genome was broken into by another IS607-cp. Unfortunately, other sequences including downstream of IS607-cp in RM6 had no significant similarity to the GenBank database possibly due to the high plasticity of the genomic structure of *M. aeruginosa* (10, 15). The distributions of IS607-cp-1 to -3 were further confirmed using a PCR assay with primers specific to each of the IS607-cp (IS1F and IS1R, IS2F and IS2R and IS3F and IS3R; Table 2; data not shown). By comparing their flanking sequences with the intact sequences of strain NIES843 and PCC7806, IS607-cp-1 and IS607-cp-2 were determined to be 1,853-bp and 1,852-bp elements, respectively. Both IS607-cp are found between adjacent G residues (Fig. 2B), suggesting a similar insertion target sequence for IS607 (18). Both elements contain *orfA* (202 a.a.) and *orfB* (390 a.a.) as well as the terminal regions including 11-bp incomplete inverted repeats (IRs) that are shorter than IS607 (18). Note that *orfA* of IS607-cp-1 in NIES90 and *orfB* of IS607-cp-3 may be pseudogenes where a mutation resulted in an in-frame termination codon within each ORF.

Phylogenetic relationships of IS607-cp

We tried to estimate the evolutionary history of the IS607-cp sequences using phylogenetic analyses. In a preliminary analysis with divergent IS607 family elements based on amino acid sequences deduced from *orfA* which is indispensable for transposition (18), the sequence of Ma-LMM01 was robustly nested within a monophyletic clade composed of some cyanobacterial counterparts and those from bacteria belonging to diverse phyla including *H. pylori* (data not shown). Henceforth, we focused on the clade including Ma-LMM01 and performed further phylogenetic analysis using the nucleotide sequence of *orfA* from Ma-LMM01 and closely related sequences shown by the preliminary analysis in addition to sequences from the four strains of *M. aeruginosa* (Fig. 3). The *orfA* sequence of Ma-LMM01 showed phylogenetic affinity with *M. aeruginosa* counterparts (BP=82%; BPP0.95); and this group was then connected to the sequences from *Cyanothece* sp. strains CCY0110 and PCC7424. *Nodularia spumigena* CCY9414 and *Crocosphaera watsonii* WH8501 were branched into the clade described above in this order (BP=83% and 90%, respectively; BPP0.95). Robust phylogenetic affinity was observed among the sequences of Ma-LMM01, *M. aeruginosa* and *Cyanothece* spp., which strongly suggests a recent HGT event for IS607-cp between the cyanobacteria and Ma-LMM01 relatives. Considering the fact that IS607-cp homologues are widespread in diverse cyanobacterial genera and that IS elements are rarely found in phages (32), IS607-cp of Ma-LMM01 may have been horizontally acquired from cyanobacteria although no direct evidence of its origin has been obtained.

Tanabe *et al.* (39) inferred a phylogenetic relationship in strains of *M. aeruginosa* based on multi-locus sequence typing analysis where most of the strains were grouped into

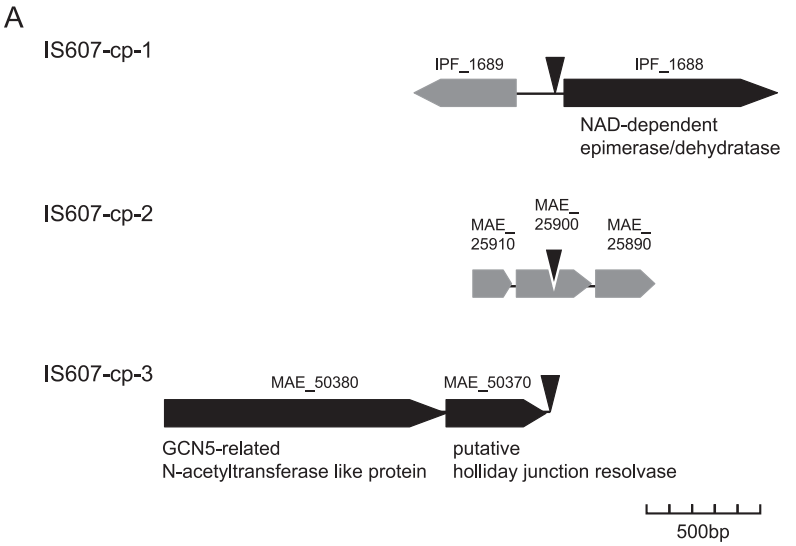

Fig. 2. Insertion loci of IS607-cp found in *M. aeruginosa* strains. (A) Partial genomic structures around the IS607-cp insertion loci. Putative functions and locus tags of the corresponding genes for each ORF are indicated. Genes encoding hypothetical proteins are indicated in gray. IS607-cp insertion sites are represented by closed triangles, with IS607-cp having a rightward orientation. (B) Nucleotide sequences of terminal regions of IS607-cp-1 (cp1) from strains NIES90, NIES112 and NIES604, and IS607-cp-2 (cp2) and their insertion sites. Terminal sequences of both IS607-cp are in lower-case letters, and flanking sequences are in capital letters. Insertion sites were determined based on the genomic sequences of *M. aeruginosa* strains NIES843 and PCC7806. Note that the insertion site of IS607-cp-3 is not fixed since we failed to obtain its 3' flanking sequence.

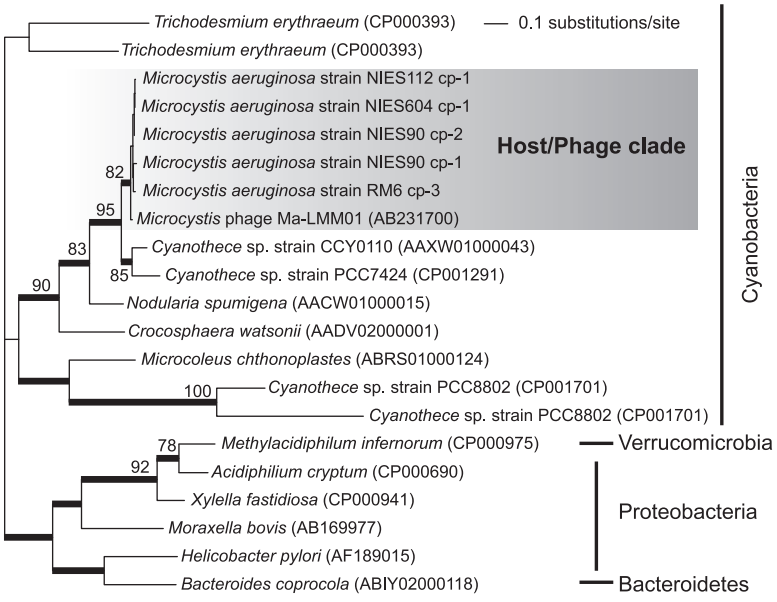

Fig. 3. A phylogenetic tree inferred from the nucleotide sequences of *orfA* of IS607-cp and its homologues. The maximum likelihood phylogenetic analysis was performed using RAxML 7.2.1 (36) with the general time reversible (GTR) model and a gamma distribution for four categories (GTR+Γ model). Only bootstrap supports equal to or more than 70% are indicated. Bayesian posterior probabilities equal to or more than 0.95 are represented by thick bars.

five major clades (Groups A–E). In the tree, the IS607-cp-containing strains NIES90, NIES112 and NIES604 belonged to Group A, Group E and Group C, respectively, whereas strains not containing IS607-cp were scattered across Group A, Group B and Group D (39). One possible explanation for this sporadic distribution of IS607-cp in *M. aeruginosa* strains may be vertical inheritance from ancestral cyanobacteria followed by the loss of IS607-cp in some strains. This is highly likely true for IS607-cp-1 found in strains NIES90, NIES112 and NIES604, since they are inserted at a homologous position (Fig. 2B). Alternatively, given the frequent HGT nature of IS elements (20, 26), there is a possibility of HGT of the other IS607-cp within certain strains of *M. aeruginosa*. Curiously, the branching patterns for IS607-cp *orfA* (Fig. 3; see above) are incongruent with cyanobacterial phylogeny inferred from 16S and 23S rRNA gene sequences, where *M. aeruginosa* is closely related to *Synechocystis* sp. PCC6803 and *C. watsonii* WH8501 in addition to two strains of *Cyanothece* spp.; whereas *N. spumigena* CCY9414 is relatively distantly related to *M. aeruginosa* (10). This observation also supports the possibility of HGT of IS607-cp particularly between distant cyanobacteria. Observations supporting the transfer of the IS elements between phylogenetically distant bacteria are also documented in a number of previous studies (3, 5, 17).

The horizontal transfer of IS elements has been explained largely by conjugation based on the observation that the elements are found in plasmids (3, 12, 17). However, since a number of observations with regard to frequent phage (virus)-to-host HGT events have been reported (4, 8, 9, 22),

which in some cases are attributed to virulent phages (8, 22), the Ma-LMM01 phage harboring IS607-cp shows the possibility of phage-mediated transfer of IS elements. Indeed, several bacteriophages have been reported to have IS elements (32). Occasionally IS elements lead to genomic rearrangements (29) and thereby drive chromosomal diversification within some prokaryote species (2, 21). The complete genomes of *M. aeruginosa* strains NIES843 and PCC7806 have a remarkably high number of transposable elements including more ISs than other cyanobacterial genomes sequenced thus far; and these elements are believed to bring plasticity into their genomes (10, 15). Although a conclusive answer requires further detailed research, Ma-LMM01-type phages may potentially contribute to the plasticity of the genomes of *M. aeruginosa* as donors of IS elements.

Acknowledgements

This study was supported by a Grant-in-Aid for Science Research (No. 20310045). SK, RK, and NH were research fellows supported by JSPS for Young Scientist (Nos. 224469, 210528, and 218877, respectively).

References

1. Altschul, S.F., W. Gish, W. Miller, E.W. Myers, and D.J. Lipman. 1990. Basic local alignment search tool. J. Mol. Biol. 215:403–410.
2. Bickhart, D.M., J.P. Gogarten, P. Lapierre, L.S. Tisa, P. Normand, and D.R. Benson. 2009. Insertion sequence content reflects genome plasticity in strains of the root nodule actinobacterium *Frankia*. BMC Genomics 10:468.
3. Bourgoin, F., G. Guédon, M. Pébay, Y. Roussel, C. Panis, and B. Decaris. 1996. Characterization of a mosaic IS*S1* element and evidence for the recent horizontal transfer of two different types of IS*S1* between *Streptococcus thermophilus* and *Lactococcus lactis*. Gene 178:15–23.
4. Canchaya, C., G. Fournous, S. Chibani-Chennoufi, M.-L. Dillmann, and H. Brüssow. 2003. Phage as agents of lateral gene transfer. Curr. Opin. Microbiol. 6:417–424.
5. Cassier-Chauvat, C., M. Poncelet, and F. Chauvat. 1997. Three insertion sequences from the cyanobacterium *Synechocystis* PCC6803 support the occurrence of horizontal DNA transfer among bacteria. Gene 195:257–266.
6. Codd, G.A., J. Lindsay, F.M. Young, L.F. Morrison, and J.S. Metcalf. 2005. Harmful cyanobacteria, p. 1–23. *In* J. Huisman, H.C.P. Matthijs, and P.M. Visser, (ed.), Harmful Cyanobacteria. Springer, Dordrecht.
7. Coleman, M.L., M.B. Sullivan, A.C. Martiny, C. Steglich, K. Barry, E.F. DeLong, and S.W. Chisholm. 2006. Genomic islands and the ecology and evolution of *Prochlorococcus*. Science 311:1768–1770.
8. Filée, J., P. Forterre, and J. Laurent. 2003. The role played by viruses in the evolution of their hosts: a view based on informational protein phylogenies. Res. Microbiol. 154:237–243.
9. Forterre, P. 1999. Displacement of cellular proteins by functional analogues from plasmids or viruses could explain puzzling phylogenies of many DNA informational proteins. Mol. Microbiol. 33:457–465.
10. Frangeul, L., P. Quillardet, A.M. Castets, *et al.* 2008. Highly plastic genome of *Microcystis aeruginosa* PCC 7806, a ubiquitous toxic freshwater cyanobacterium. BMC genomics 9:274.
11. Fuhrman, J.A. 1999. Marine viruses and their biogeochemical and ecological effects. Nature 399:541–548.
12. Hartl, D.L., and S.A. Sawyer. 1988. Why do unrelated insertion sequences occur together in the genome of *Escherichia coli*? Genetics 118:537–541.
13. Kalia, A., A.K. Mukhopadhyay, G. Dailide, Y. Ito, T. Azuma, B.C.Y. Wong, and D.E. Berg. 2004. Evolutionary dynamics of insertion sequences in *Helicobacter pylori*. J. Bacteriol. 186:7508–7520.
14. Kamikawa, R., H. Nishimura, and Y. Sako. 2009. Analysis of the mitochondrial genome, transcripts, and electron transport activity in the dinoflagellate *Alexandrium catenella* (Gonyaulacales, Dinophyceae). Phycol. Res. 57:1–11.
15. Kaneko, T., N. Nakajima, S. Okamoto, *et al.* 2007. Complete genomic structure of the bloom-forming toxic cyanobacterium *Microcystis aeruginosa* NIES-843. DNA Res. 14:247–256.
16. Kasai, F., M. Kawachi, M. Erata, and M.M. Watanabe. 2004. NIES-Collection List of Strains, 7th ed. The National Institute for Environmental Studies, Tsukuba, Japan.
17. Kato, K., K. Ohtsuki, H. Mitsuda, T. Yomo, S. Negoro, and I. Urabe. 1994. Insertion sequence IS*6100* on plasmid pOAD2, which degrades nylon oligomers. J. Bacteriol. 176:1197–1200.
18. Kersulyte, D., A.K. Mukhopadhyay, M. Shirai, T. Nakazawa, and D.E. Berg. 2000. Functional organization and insertion specificity of IS*607*, a chimeric element of *Helicobacter pylori*. J. Bacteriol. 182:5300–5308.
19. Kondo, R., M. Komura, S. Hiroishi, and Y. Hata. 1998. Detection and 16S rDNA sequence analysis of a bloom-forming cyanobacterial genus *Microcystis*. Fish. Sci. 64:840–841.
20. Lawrence, J.G., H. Ochman, and D.L. Hartl. 1992. The evolution of insertion sequences within enteric bacteria. Genetics 131:9–20.
21. Leavis, H.L., R.J.L. Willems, W.J.B. van Wamel, F.H. Schuren, M.P.M. Caspers, and M.J.M. Bonten. 2007. Insertion sequence-driven diversification creates a globally dispersed emerging multiresistant subspecies of *E. faecium*. PLoS Pathog. 3:e7.
22. Lindell, D., M.B. Sullivan, Z.I. Johnson, A.C. Tolonen, F. Rohwer, and S.W. Chisholm. 2004. Transfer of photosynthesis genes to and from *Prochlorococcus* viruses. Proc. Natl. Acad. Sci. USA 101:11013–11018.
23. Lindell, D., J.D. Jaffe, M.L. Coleman, *et al.* 2007. Genome-wide expression dynamics of a marine virus and host reveal features of co-evolution. Nature 449:83–86.
24. Liu, Y.G., N. Mitsukawa, T. Oosumi, and R.F. Whittier. 1995. Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. Plant J. 8:457–463.
25. Liu, Y.G., and R.F. Whittier. 1995. Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. Genomics 25:674–681.
26. Mahillon, J., and M. Chandler. 1998. Insertion sequences. Microbiol. Mol. Biol. Rev. 62:725–774.
27. Mann, N.H., A. Cook, A. Millard, S. Bailey, and M. Clokie. 2003. Marine ecosystems: bacterial photosynthesis genes in a virus. Nature 424:741.
28. Monier, A., J.-M. Claverie, and H. Ogata. 2007. Horizontal gene transfer and nucleotide compositional anomaly in large DNA viruses. BMC Genomics 8:456.
29. Nevers, P., and H. Saedler. 1977. Transposable genetic elements as agents of gene instability and chromosomal rearrangements. Nature 268:109–115.
30. Proctor, L.M., and J.A. Fuhrman. 1990. Viral mortality of marine bacteria and cyanobacteria. Nature 343:60–62.
31. Ronquist, F., and J.P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574.
32. Sakaguchi, Y., T. Hayashi, K. Kurokawa, *et al.* 2005. The genome sequence of *Clostridium botulinum* type C neurotoxin-converting phage and the molecular mechanisms of unstable lysogeny. Proc. Natl. Acad. Sci. USA 102:17472–17477.
33. Sawyer, S.A., D.E. Dykhuizen, R.F. DuBose, L. Green, T. Mutangadura-Mhlanga, D.F. Wolczyk, and D.L. Hartl. 1987. Distribution and abundance of insertion sequences among natural isolates of *Escherichia coli*. Genetics 115:51–63.
34. Sharon, I., A. Alperovitch, F. Rohwer, *et al.* 2009. Photosystem I gene cassettes are present in marine virus genomes. Nature 461:258–262.
35. Smith, M.C.M., and H.M. Thorpe. 2002. Diversity in the serine recombinases. Mol. Microbiol. 44:299–307.
36. Stamatakis, A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22:2688–2690.
37. Suttle, C.A. 2007. Marine viruses—major players in the global ecosystem. Nat. Rev. Microbiol. 5:801–812.

38. Takashima, Y., T. Yoshida, M. Yoshida, Y. Shirai, Y. Tomaru, Y. Takao, S. Hiroishi, and K. Nagasaki. 2007. Development and application of quantitative detection of cyanophages phylogenetically related to cyanophage Ma-LMM01 infecting *Microcystis aeruginosa* in fresh water. Microbes Environ. 22:207–213.
39. Tanabe, Y., F. Kasai, and M.M. Watanabe. 2007. Multilocus sequence typing (MLST) reveals high genetic diversity and clonal population structure of the toxic cyanobacterium *Microcystis aeruginosa*. Microbiology 153:3695–3703.
40. Tomioka, N., T. Nagai, T. Kawasaki, A. Imai, K. Matsushige, and K. Kohata. 2008. Quantification of *Microcystis* in a eutrophic lake by simple DNA extraction and SYBR Green real-time PCR. Microbes Environ. 23:306–312.
41. Yoshida, M., T. Yoshida, Y. Takashima, R. Kondo, and S. Hiroishi. 2005. Genetic diversity of the toxic cyanobacterium *Microcystis* in Lake Mikata. Environ. Toxicol. 20:229–234.
42. Yoshida, M., T. Yoshida, M. Satomi, Y. Takashima, N. Hosoda, and S. Hiroishi. 2008. Intra-specific phenotypic and genotypic variation in toxic cyanobacterial *Microcystis* strains. J. Appl. Microbiol. 105:407–415.
43. Yoshida, M., T. Yoshida, Y. Yoshida-Takashima, A. Kashima, and S. Hiroishi. 2010. Real-time PCR detection of host-mediated cyanophage gene transcripts during infection of a natural *Microcystis aeruginosa* population. Microbes Environ. 25:211–215.
44. Yoshida, T., Y. Takashima, Y. Tomaru, Y. Shirai, Y. Takao, S. Hiroishi, and K. Nagasaki. 2006. Isolation and characterization of a cyanophage infecting the toxic cyanobacterium *Microcystis aeruginosa*. Appl. Environ. Microbiol. 72:1239–1274.
45. Yoshida, T., K. Nagasaki, Y. Takashima, Y. Shirai, Y. Tomaru, Y. Takao, S. Sakamoto, S. Hiroishi, and H. Ogata. 2008. Ma-LMM01 infecting toxic *Microcystis aeruginosa* illuminates diverse cyanophage genome strategies. J. Bacteriol. 190:1762–1772.