Cyanophage Diversity, Inferred from g20 Gene Analyses, in the Largest Natural Lake in France, Lake Bourget

Ursula Dorigo, Stéphan Jacquet, and Jean-François Humbert*
Equipe de Microbiologie Aquatique, Station INRA d’Hydrobiologie Lacustre, UMR CARRTEL, 74203 Thonon Cedex, France

Received 8 October 2003/Accepted 13 November 2003

The genetic diversity of the natural freshwater community of cyanophages and its variations over time have been investigated for the first time in the surface waters of the largest natural lake in France. This was done by random screening of clone libraries for the g20 gene and by denaturing gradient gel electrophoresis (DGGE). Nucleotide sequence analysis revealed 35 distinct cyanomyovirus g20 genotypes among the 47 sequences analyzed. Phylogenetic analyses showed that these sequences fell into seven genetically distinct operational taxonomic units (OTUs). The distances between these OTUs were comparable to those reported between marine clusters. Moreover, some of these freshwater cyanophage sequences were genetically more closely related to marine cyanophage sequences than to other freshwater sequences. Both approaches for the g20 gene (sequencing and DGGE analysis) showed that there was a clear seasonal pattern of variation in the composition of the cyanophage community that could reflect changes in its biological, chemical, and/or physical environment.

During the last 2 decades, viruses have been shown to be a key component of aquatic microbial communities because of their abundance, ubiquity, and potential impact on both biogeochemical and ecological cycles through the infection and lysis of bacterial and phytoplankton communities (9, 33). Several reports concerning the impact of viral lysis on the dynamics and clonal composition of bacterial and algal host communities have been published (see, for instance, references 2, 20, 22, and 28), but in contrast little is yet known about the diversity of viral communities (e.g., reference 37). However, during the last few years, significant advances have been made in assessing the diversity of natural viral communities in marine ecosystems (3, 13, 23, 34, 37), whereas we still know relatively little about the viral diversity of freshwater ecosystems (17, 36).

The development of molecular tools and genetic techniques have made it possible to reveal the extensive viral diversity. Until recently, this parameter had been greatly underestimated by culture-dependent methods and morphological identification (5, 32). We now have methods that can target either the entire genome, such as pulsed-field gel electrophoresis (12, 37), total community DNA-DNA hybridization (32), and restriction digestion (3, 17, 34), or single genes, such as cloning and sequencing methods (to create clone libraries) and denaturing gradient gel electrophoresis (DGGE) (19, 25). However, microbiologists attempting to study viral diversity from single genes still face a major problem: there is no “universal target,” such as the rRNA genes, that occurs in both bacteria and eukaryotic microorganisms (24).

In the general context of our attempts to assess the diversity and functioning of pelagic microbial communities in the three great subalpine lakes (lakes Annecy, Bourget, and Geneva), it is very important to be able to evaluate the diversity and dynamics of viruses, especially the cyanophages. Freshwater cyanobacteria play a key role within the phytoplanktonic community in these ecosystems, and we want to know what contribution viruses make to cyanobacterial structure and dynamics and their role in control and mortality during and following bloom episodes. Picocyanobacteria such as Synechococcus spp. dominate the phytoplankton community biomass in Lake Annecy (oligotrophic), whereas large, filamentous, toxin-producing cyanobacteria such as Planktothrix rubescens can be predominant during much of the year in the mesotrophic lakes Bourget and Geneva (16; S. Jacquet, unpublished data).

In this study, we tested the CSP1-CSP8 cyanophage primers targeting the g20 gene encoding the viral capsid structure, previously defined by Zhong et al. (39) for marine viruses, in surface water samples from Lake Bourget, and we subsequently characterized the cyanophage diversity by both cloning-sequencing and DGGE. The abundance of viruses, heterotrophic bacteria, picocyanobacteria, and Planktothrix rubescens was also assessed by flow cytometry or microscopic counting. All of these determinations were performed once or twice a month between September 2002 and January 2003.

MATERIALS AND METHODS

Site description, sampling strategy, and cyanophage isolation. A complete description of Lake Bourget, which is the largest natural lake in France, has been provided elsewhere (S. Jacquet, J.-F. Briand, C. Leboulanger, G. Paolini, L. Oberhaus, B. Tassin, B. Vinçon-Leite, J.-C. Dutart, O. Anneville, and J.-F. Humbert, submitted for publication). Lake water was collected at a depth of approximately 5 m at the reference station located in the middle and deepest part of the lake once a month from September 2002 to January 2003. Twenty liters of lake water was sampled by using an electric pump on a boat deck and kept in a plastic flask while being transported to the laboratory. The water samples were kept at 4°C in the dark for no more than 1 week before being processed. The water was filtered through GF/F filters in order to remove both phyto- and zooplankton. The resulting filtrate was then concentrated at least 100-fold by tangential flow filtration using a mini-ultrasette with a 100-kDa cutoff membrane (Vivathow; Vivasciences). The concentrate was stored at 4°C until PCR amplification, since phage communities can be stored in these conditions for days to months without significant loss of titer (S. Jacquet, unpublished data, and K. Rodda and C. A. Suttle, unpublished data).

* Corresponding author. Mailing address: UMR CARRTEL, Station INRA d’Hydrobiologie Lacustre, Equipe de Microbiologie Aquatique, BP 511, 74203 Thonon Cedex, France. Phone: 33-4-50-26-78-09. Fax: 33-4-50-26-78-06. E-mail: humbert@thonon.inra.fr.
Flow cytometry analyses. Samples were analyzed using a FACSCalibur (Becton Dickinson) flow cytometer equipped with a blue laser beam fixed at 488 nm and with the original filter set up. Picoycyanobacteria were analyzed without any fixative or dye, and the community was identified on the basis of its chlorophyll and autofluorescence and the right-angle light scatter. To count the heterotrophic bacteria and viruses, samples underwent preliminary fixing with glutaraldehyde (0.25% final concentration) for 30 min in dim-light conditions and were then filtered through 0.2-μm-pore-size filters. For the analysis of heterotrophic bacteria, samples were diluted 50-fold with water from the lake sampled the same day and subjected to filtration with 0.2-μm-pore-size filters. For the analysis of viruses, the samples were diluted 100-fold in Tris-EDTA buffer (filtered through 0.02-μm-pore-size filters; pH 7.8) and heated to 75°C for 10 min. Samples of these two communities were stained by using the nucleic acid dye SYBR Green I (see reference 18 for more details). Cellular parameters were determined relative to the values found for 1-μm beads (Molecular Probes). Data were collected in listmode files and then analyzed using CY-TOWIN software (30) (available at http://www.sb-rosequ.fr/Phyto cytoto.html).

Counting of *P. rubescens*. At each sampling date, 300 ml of water was preserved with Lugol’s iodine solution for subsequent microscopic counts. Two-hundred-micrometer units of *P. rubescens* filaments were counted using the Utermöhl inverted microscope technique after sedimenting 25 to 50 ml of water. The number of cells was estimated by assuming that the mean length of a *P. rubescens* cell (estimated from 100 measurements) was 5 μm.

PCR amplification, cloning, and sequencing. All PCRs were performed using a T-Personal DNA thermal cycler (Biometra). Three pairs of oligonucleotides (CPS1-CPS2, CPS3-CPS4, and CPS1-CPS8, as the gene of cyanophages belonging to the family Myoviridae) were used. Oligonucleotides were manually corrected using GeneDoc. A phylogenetic tree was constructed for the whole data set by neighbor joining on the Jukes-Cantor distances by using the PHYLIP software package (7). The bootstrap option was used to run 1,000 replicates. Several operational taxonomic units (OTUs) were defined on the basis of their bootstrap proportions. The Chao-1 and abundance-based coverage estimators of species richness (15) were calculated using EstimateS software (http://viceroy.eeb.uconn.edu/estimates), and a rarefaction curve was obtained using PAST software (http://folk.uio.no/ohammer/past).

DGGE analysis. For the DGGE analysis, the CPS1 primer was altered by adding a 40-nucleotide GC-rich sequence (GC clamp) to the 5’ end and hence renamed CPS1GC. The sequence was as follows: 5’GC CGG CGG CGG CCC GGG CCC GTC CGG CCC CCC CCC G-GTA-G/1A/ATTTTTCACTAT TGAGC/CTTGTGGG 3’. The 50-μl reaction mix for PCR-DGGE contained 10× Taq reaction buffer (Eurobiol), 2.5 mM MgCl₂, 200 μM concentrations of each deoxynucleoside triphosphate, 1 μM concentrations of each primer (CPS1GC and CPS8), 1.25 U of Tag DNA polymerase (Eurobiolacq; Eurobiol), and 10 μl of the viral concentrate. For each set of reactions, a negative control sample was included that contained all of the reagents but without the viral DNA. PCRs were carried out as described by Zhong et al. (39), with the exception of the annealing temperature, which was higher (46°C) than that used previously. PCR products (13 μl) were subjected to electrophoresis on a 1.4% (wt/vol) agarose gel.

Positive PCR products were ligated to the pGEM-T System II vector (Promega) and then transformed into JM109-competent cells (Promega) according to the manufacturer’s instructions. Among 15 positive clones (white colonies) from each clone library were randomly selected, and plasmid DNA was isolated by PCR, using the commercial primers SP6 and T7, and then sequenced (6).

Phylogenetic analyses. The sequences were aligned using the Pileup module of the GCG package (Genetics Computer Group, Inc., Madison, Wis.), and alignments were manually corrected using GeneDoc. A phylogenetic tree was constructed for the whole data set by neighbor joining on the Jukes-Cantor distances by using the PHYLIP software package (7). The bootstrap option was used to run 1,000 replicates. Several operational taxonomic units (OTUs) were defined on the basis of their bootstrap proportions. The Chao-1 and abundance-based coverage estimators of species richness (15) were calculated using EstimateS software (http://viceroy.eeb.uconn.edu/estimates), and a rarefaction curve was obtained using PAST software (http://folk.uio.no/ohammer/past).

DGGE analysis. For the DGGE analysis, the CPS1 primer was altered by adding a 40-nucleotide GC-rich sequence (GC clamp) to the 5’ end and hence renamed CPS1GC. The sequence was as follows: 5’GC CGG CGG CGG CCC GGG CCC GTC CGG CCC CCC CCC G-GTA-G/1A/ATTTTTCACTAT TGAGC/CTTGTGGG 3’. The 50-μl reaction mix for PCR-DGGE contained 10× Taq reaction buffer (Eurobiol), 2.5 mM MgCl₂, 200 μM concentrations of each deoxynucleoside triphosphate, 1 μM concentrations of each primer (CPS1GC and CPS8), 1.25 U of Tag DNA polymerase (Eurobiolacq; Eurobiol), and 20 μl of viral lysate. The same PCR and electrophoresis conditions as described above were used.

DGGE analysis was performed using the CBS-DGGE 2000 system (CBS. B.S. Scientific Co., Inc.). PCR products (40 μl) were loaded onto a 1-mm-thick 6% polyacrylamide gel in 1× TAE (40 mM Tris acetate [pH 7.4], 20 mM sodium acetate, 1 mM Na₂-EDTA) which contained a 30 to 70% linear denaturing gradient (100% is defined as 7 M urea plus 40% deionized formamide), as previously established for perpendicular DGGE (data not shown). Electrophoresis was performed at a constant voltage of 100 V and a temperature of 60°C for the optimal duration of 16 h (data not shown). Separated PCR products were stained for 45 min in the dark with SybrGold (Molecular Probes), visualized on a UV transilluminator (Tex-35 M; Bioblock Scientific), and photographed with a Kodak DC290 camera.

Nucleotide sequence accession numbers. The g20 nucleotide sequences have been deposited in the GenBank-EMBL database under accession numbers AY426128 to AY426174.

RESULTS

Dynamics of microbial communities. The dynamics of viruses, heterotrophic bacteria, and cyanobacteria (both picocyanobacteria and *P. rubescens*) in Lake Bourget were monitored in the autumn and early winter of 2002-2003. During this period, we observed a net decrease in the cell (or particle) abundance of picocyanobacteria, heterotrophic bacteria, and viruses, whereas the cell density of *P. rubescens* rose (Fig. 1). There was a clear decline in the density of the viral community throughout the period of interest (from 26 × 10⁸ to 5.8 × 10⁷ parts/ml), apart from an isolated peak in October (Fig. 1A). The abundance of the heterotrophic bacteria was halved during the period studied, but the dynamics of this community was characterized by a succession of alternating decreasing and increasing phases until January 2003 (Fig. 1B). For picocyanobacteria, there was an initial decrease in cell density from 2.1 × 10⁷ to 3.0 × 10⁶ cells ml⁻¹ in September, followed by a stationary phase until the end of October and then by a second decrease in November (Fig. 1C). Finally, the abundance of *P. rubescens* rose almost 10-fold from September to November and then decreased slowly (Fig. 1D). Viral abundance was positively correlated with that of heterotrophic bacteria (*r* = 0.76; *P* < 0.05) and that of picocyanobacteria (*r* = 0.77; *P* < 0.01). In contrast, there was no correlation between the abundance of viruses and that of *P. rubescens* (*r* = −0.50; not significant).

Analysis of nucleotide g20 sequences. Among the 47 sequences obtained by random sequencing in the different clones libraries, 12, 13, 12, 6, and 4 were obtained from clones taken on September 26, October 28, November 27, December 23, and January 28, respectively. On the two last dates, only a few sequences were obtained due to a low cloning efficiency (low number of white colonies); this was linked to the low PCR efficiency (data not shown). Consequently, the sequences obtained on these last two dates were pooled in the subsequent analyses. The 47 sequences included 35 different haplotypes. The total nucleotide diversity was not significantly different for the four sampling periods: 0.29 (±0.15) in September, 0.32 (±0.16) in October, 0.22 (±0.12) in November, and 0.34 (±0.18) in December and January.

Phylogenetic analyses of g20 sequences. Phylogenetic analysis revealed the existence of six clearly distinguished clusters containing 11, 14, 4, 3, 12, and 2 sequences, respectively (Fig. 2). In each of these clusters, the percentage of nucleotide sequence similarity between cyanophage sequences was always >96%. On the other hand, the similarity ranged between 54 and 60% when sequences belonging to different clusters were compared. All of these clusters were highly supported by their bootstrap proportion (1,000 resamplings). One sequence (“Seq. 1” in Fig. 2) stood apart from all of these clusters. Thus, a total of seven OTUs were identified in this study. Both the Chao-1 and abundance-based coverage richness estimators were equal to 7 ± 1 (standard deviation), which means that we obtained a good estimation of the OTU richness in the cyano-
nophage community of Lake Bourget. Similarly, the asymptotic rarefaction curve (data not shown) confirmed the representative nature of our sequence sample.

The phylogenetic distances between these freshwater cyanophage sequence clusters were of the same order as those found when they were compared to marine cyanophage sequences (Fig. 2). Interestingly, the sequences in some of our freshwater cyanophage clusters (clusters 2 and 4) were genetically more similar to some marine cyanophage sequences than to the other clusters defined in this study. For example, there was 73 to 74% nucleotide sequence similarity between cluster 4 sequences and marine cyanophage sequences, ay152732, ay152738, and ay152741, but less than 60% similarity with other freshwater cyanophage sequences.

Temporal variation in the cyanophage community. The distribution of the sequences belonging to the different clusters determined above (Fig. 3) revealed first that sequences belonging to cluster 1 were only obtained in September and October. Moreover, it appears that these sampling months were characterized by very similar patterns for the sequence distributions, distinguished by the dominance of cluster 1 sequences, but also by the presence of cluster 2, 3, and 5 sequences. November was characterized by a high dominance of sequences belonging to cluster 2 (Fig. 3). For December and January, sequences belonging to five of the six clusters were found in quite similar proportions (Fig. 3). Only two clusters (clusters 2 and 5) were found throughout the whole season. The first (cluster 2) showed high variations in regard to the sampling months, while for the second (cluster 5), no variation was observed.

DGGE analysis. DGGE analysis was performed on the same samples as those used for the sequencing approach. On the one hand, the study of the DGGE migration profile (Fig. 4) revealed that the first two sampling months (September and October) were characterized by similar band patterns. On the other hand, considerable differences in the band intensities were observed in the migration profile obtained for the November sample (Fig. 4). During the last two sampling months (December and January), there was a low intensity in the profiles due to a low PCR amplification efficiency. It seems, however, that these patterns were more similar to that obtained for the November sample than to those for the September and October samples (Fig. 4).

DISCUSSION

Our findings show that the phylogenetic diversity of natural cyanophages in the mesotrophic Lake Bourget seems to be very great. Seven OTUs were identified among the 47 sequences obtained over a limited period of time (6 months in 2002-2003). This relatively high diversity of the cyanophage community has already been reported, but as far as we are
aware, it has only been shown in marine ecosystems (19, 39). Zhong et al. (39) and Sullivan et al. (26) proposed that this could be due to important genetic exchanges between phage and host or between coinfecting phages. Another complementary hypothesis is that the high diversity within the cyanophage community may be linked to high diversity in the host community. In their study of picocyanobacteria from oligo- and mesotrophic lakes, Becker et al. (1) found that different lineages of the picoplankton clade sensu Urbach et al. (29) were present in Lake Constance (Germany), an ecosystem which has physical, chemical, and biological characteristics very similar to those of Lake Bourget. Following the same approach, Crosbie et al. (4) recently showed that there are at least seven clusters within nonmarine picocyanobacteria. Despite its partial nature, our study of the molecular characteristics of freshwater cyanophages suggests once again that genetic diversity may be high within this community.

The genetic distances we found between the different clus-

![FIG. 2. Phylogenetic relationships among the partial g20 gene sequences from the cyanophage community of Lake Bourget. The unrooted tree was constructed from pairwise Jukes and Cantor distances by using the neighbor-joining method. Marine cyanophage sequences, identified by their GenBank accession numbers, were also added. Only bootstrap values of >80% are indicated at the nodes of the tree.](http://aem.asm.org/)

![FIG. 3. Temporal changes in the abundance of the six clusters defined by the phylogenetic analysis of the g20 gene sequences from the cyanophage community of Lake Bourget.](http://aem.asm.org/)
ters were in the same range as those found by Zhong et al. (39) for isolates and natural marine cyanophages. Very interestingly, we found that some of our freshwater cyanophage sequences were more similar to marine cyanophage sequences than to other freshwater cyanophage sequences (Fig. 2). This finding suggests that some marine and freshwater cyanophages may have shared a common ancestor. This finding is also consistent with the fact that a close phylogenetic relationship can be found between several marine and freshwater strains of *Synechococcus* spp. (14).

Seasonal changes recorded in the (cyano)phage community sampled in surface waters of Lake Bourget affected both the abundance and the diversity of these microorganisms. As mentioned above, we observed a clear reduction in the total viral abundance from September to January and, at the same time, a decrease in PCR efficiency for the g20 gene. This probably reflected a decrease in the target number and thus in cyanomyophage abundance. Many authors have reported a winter decrease of this sort in the viral abundance in aquatic ecosystems, and this has been attributed to a positive correlation between host and virioplankton abundance (see the review by Wommack and Colwell [38]). In our study, virioplankton abundance was indeed positively correlated with that of the picobacteria (both auto- and heterotrophs). In contrast, there was no clear relationship between the viral community and the filamentous cyanobacterium *P. rubescens*, even though the concentration of the latter was relatively high. This finding suggests that our sequences are likely to belong only to picocyanobacterial phages and/or that important resistance mechanisms exist for filamentous cyanobacteria, such as *P. rubescens*, of which major blooms without extinction episodes have been recorded (S. Jacquet et al., submitted). Such host resistance has often been reported for other cyanobacterial species (27, 31).

With regard to the seasonal pattern of g20 phylogenetic diversity, sequencing and DGGE approaches revealed similar values in September and October, which at first sight validates both methodological approaches. Such seasonal changes in the relative cyanophage g20 genotypes were also reported by Marston and Sallee (19) in coastal cyanophage communities and by Wommack et al. (37) in the composition of natural virioplankton communities in Chesapeake Bay. It would have been very interesting to evaluate changes in the composition of the host communities as well in order to determine the relationships between host and virus compositions. This is one of the prospects for future research arising from this study.

From a practical point of view, this study confirmed that the DGGE fingerprinting technique is a very efficient tool for monitoring changes in the composition of natural cyanophage communities. This had previously been reported in marine cyanophage communities through the use of another set of primers by Frederickson et al. (8) and by Wilson et al. (35). Combined with quantitative PCR, it may make it possible to assess the relative abundance of the various strains in these communities.

**FIG. 4.** Temporal changes in the DGGE band patterns obtained for PCR products resulting from the amplification of a 592-bp fragment of the g20 gene in the cyanophage community of Lake Bourget. Arrows indicate the principal differences between the September-October and November band patterns. Lanes: A and G, mix of September-October samples; B, September sample; C, October sample; D, November sample; E, December sample; F, January sample.
communities. Indeed, this approach is now presently used in clinical virology (e.g., see references 11 and 21), and its application to aquatic viral ecology could be extremely promising.

ACKNOWLEDGMENTS

We are very grateful to F. Chen (University of Maryland Biotechnology Institute, Baltimore) for his valuable advice about the PCR conditions for g20 gene amplification. Monika Ghosh is acknowledged for improving the English version of the manuscript.

The DGEI system was funded by la Fondation pour la Recherche Médicale. The flow cytometer was funded by the Institut National de la Recherche Agronomique and l’Université de Savoie.

REFERENCES


