Research in Microbiology 166 (2015) 644–654

www.elsevier.com/locate/resmic

Original article

Structure and diversity of ssDNA *Microviridae* viruses in two peri-alpine lakes (Annecy and Bourget, France)

Xu Zhong[1], Baptiste Guidoni, Louis Jacas, Stéphan Jacquet*

INRA, UMR CARRTEL, 75 Avenue de Corzent, 74203 Thonon-les-Bains, France

Received 20 April 2015; accepted 10 July 2015
Available online 27 July 2015

Abstract

Microviridae is a subset of single-stranded DNA (ssDNA) viruses infecting bacteria. This group of phages has been previously observed to be very abundant (representing >90% of the total known viral metagenomic sequences) in Lake Bourget. However, this observation was made only during one period (in summer) and from a single sample collected at a single depth (near surface). This result suggests the importance of these viruses, poorly examined thus far, especially in fresh waters. In this study, performed on the two largest natural lakes in France (e.g. Lakes Annecy and Bourget), *Microviridae* structure was determined each month throughout the year (2011) using PCR-DGGE, with primers that target the major-capsid-protein-encoding gene *VP1;* cloning/sequencing was used to investigate their diversity. Our results confirm that *Microviridae* are diverse in peri-alpine lakes and are mainly represented by gokushoviruses. We also found for the first time ssDNA viruses belonging to *Alpavirinae*, another subfamily within *Microviridae* recently proposed by Krupovic and Forterre (2011), generally prophages infecting members of the Phylum Bacteroidetes. Our data also support highly variable community composition and dynamics of individual components whose patterns were different between lakes, suggesting distinct host communities and/or abiotic influences between the two ecosystems. We point out that most of the major observed ssDNA *Microviridae* viruses display boom-bust patterns (with a sharp increase/decline) in their dynamics, with high relative abundances, suggesting brutal control of hosts and rapid regulation of the host community structure.

Keywords: Lakes; Viruses; ssDNA virus; *Microviridae*; Structure; Diversity

1. Introduction

Viruses are the most abundant biological entities in the biosphere [46]. They are highly diverse, both genetically and morphologically. Our knowledge is largely based on dsDNA viruses, especially the tailed dsDNA viruses of the Caudovirales family [2,3]. However, over the last decade and, more specifically, during the past 5 years, with the emergence of viral metagenomics tools and protocols (for instance, use of multiple displacement amplification [MDA] that preferentially amplifies circular single-stranded DNA [18]), an increasing amount of research has revealed the importance of ssDNA viruses in a variety of ecosystems (reviewed in [36]). It is now recognized that the ssDNA viral community has been largely underestimated (because of technical constraints, such as insufficient staining in fluorescence-based methods such as flow cytometry and epifluorescence microscopy, or insufficient material when using PFGE). Its importance (both in terms of diversity and/or abundance) has been highlighted in lakes [23,38], desert ponds of the Mauritanian Sahara desert [10], reclaimed water [37], aquifers [45], stromatolithes [9], corals [55], ground and marine sediments [18,62], coastal estuaries [20–22,27] and the ocean in general [1,4,53].

Single-stranded DNA viruses are generally small (15–30 nm), non-tailed, with a circular genome between 1.4 and 8.5 kb and, on average, less than 10 genes [20,39]. According to the International Classification of Taxonomy of

* Corresponding author.
E-mail address: sjacquet@thonon.inra.fr (S. Jacquet).
[1] Current address: Department of Earth, Ocean and Atmospheric Sciences, University of British Columbia, Vancouver, BC, Canada.

http://dx.doi.org/10.1016/j.resmic.2015.07.003
0923-2508/

Viruses (ICTV 2012), 10 families have been recognized in this group. In aquatic ecosystems, the *Microviridae* has been reported to be the most frequently represented family, and, in particular, viruses from the *Gokushovirinae* subfamily [13,39,53]. Viruses of the *Microviridae* are constituted by five genes, three encoding for the protein capsid (with *VP1* coding for the major part of this viral capsid), REP encoding for the replication protein and ORF5 encoding for the encapsulation protein. Thus far, genes targeted to investigate *Microviridae* ssDNA diversity have been mainly *VP1* [13,21,22] and REP [53].

Over the past few years, we have developed an interest in the diversity of viruses in peri-alpine lakes (Bourget, Annecy and Geneva lakes). Our previous works [30, 57–61] focused only on dsDNA viruses, and we showed that this community is numerous, very diverse and plays a key role in the functioning of these lakes. However, an earlier metagenomic study [38] of a summer sample collected at a single depth in Lake Bourget suggested that the viral community was mainly composed of ssDNA viruses (>90% in relative abundance) and that these viruses belonged mainly to the *Microviridae* family. Despite the quantitative importance of these ssDNA viruses, their ecology is still largely unknown and needs to be assessed in these lakes. Current known *Microviridae* studies [13,21,53] focused on their geographical and vertical distribution.

In this study, we sought first to investigate the diversity and temporal changes in the ssDNA virus community using PCR-DGGE and cloning/sequencing approaches in two peri-alpine lakes characterized by distinct trophic states and host bacterial communities. Secondly, we attempted to reveal factors likely to affect the community structures of these viruses. Our questions were thus numerous: Can PCR-DGGE be used to detect ssDNA viruses in peri-alpine lakes? What are the structural patterns of ssDNA viruses in these lakes? Can we explain these patterns by environmental variables? Can new viruses be identified for peri-alpine lakes?

2. Materials and methods

2.1. Sampling

Water samples were collected once every month between January and November 2011 at reference stations of Lakes Annecy (Lat N 45.8727, Long E 6.1645333) and Bourget (Lat N 45.7469, Long E 5.86015), corresponding to the deepest part of each lake. We obtained 14 samples for Lake Annecy and 18 for Lake Bourget. Twenty-one liters, integrating the water column from surface to 20 m depth, was collected using an electric pump and tubing, and water was stored in a polycarbonate flask placed in the dark at 4 °C before filtration steps. A few hours following sampling, 20 l samples were first filtered through a 60 μm mesh and then filtered through 1 μm pore-size filters (Millipore, Bedford, MA, USA). The filtrate (i.e. <1 μm fraction) was concentrated to a volume of 200–250 ml using a 30,000 Da-molecular-weight-cutoff spiral-wound millipore ultrafiltration cartridge (regenerated cellulose, PLTK Prep/scale TFF, 1 ft2; Millipore). To ensure that all remaining small free-living bacteria were removed, we twice filtered the <1 μm concentrated fraction through 0.45 μm pore-size filters (Millipore). The absence of cellular contamination was verified using flow cytometry (not shown).

2.2. Physico-chemical variables

A variety of classical limnological variables were obtained using multiparameter probes such as the Sea-Bird SEB 19 SEACAT Profiler and the chlorophyll fluorescence Fluoroprobe (BBE Moldenke, Germany), that enabled obtaining vertical profiles for water temperature, pH, conductivity, dissolved oxygen and chlorophyll *a* concentrations. Nutrient concentrations (total nitrogen, $N-NH_4$, $N-NO_3$, SiO_2, $P-PO_4$ and total phosphorus) were measured at each station following standard AFNOR protocols (details are available at http://www.afnor.org).

2.3. Transmission electron microscopy (TEM) analysis

TEM was used to visualize and count the different morphotypes of viruses. Briefly, 8 ml formalin-fixed samples kept at 4 °C were harvested by ultracentrifugation onto 400 mesh NI electron microscope grids with a carbon-coated Formvar film, using a Centrikon TST 41.14 Swing-Out-Rotor run at 70,000 × g for 20 min at 4 °C [43,56]. Each grid was then stained for 30 s with uranyl acetate (2% wt/wt) and examined using a JEOL 1200EX TEM operated at 80 kV and 40,000× magnification. The negatives were scanned with Adobe Photoshop and phage dimensions were determined using IMAGEJ software.

2.4. Bacterial and virus quantification

Bacteria and VLPs (virus-like particles with stained nucleic acids, likely virions) were quantified using a FACSCalibur flow cytometer (Becton Dickinson) equipped with an air cooled laser providing 15 mW at 488 nm. Cells in the raw-water sample (untreated) were fixed with glutaraldehyde (0.5% final concentration, grade I, Merck) for 30 min in the dark, then diluted in 0.02 μm filtered lake water (for the bacteria) and 0.02 μm filtered Tris EDTA (for the viruses). Samples were incubated with SYBR Green I (at a final 10^{-4} dilution of the commercial stock solution; Molecular Probes), for 15 min at ambient temperature for the bacteria and for 10 min with heating for the VLPs [15,31]. For VLPs, it is noteworthy that FCM enables observing mainly dsDNA viruses (but not ssDNA viruses; [12]). Listmode files obtained were then transferred and analyzed on a PC using the custom-designed freeware CYTOWIN [54]. More details about FCM analysis and data treatment can be obtained elsewhere [15].

2.5. PCR and cloning sequencing

Prior to PCR, viral DNAs were purified for each sample using the QIAamp MinElute Virus Spin Kit (Qiagen), and then

genomic DNAs were pre-amplified by enzyme phi29 using the Genomiphi V2 DNA Amplification Kit (GE Healthcare) following the manufacturer's instructions. The latter step is crucial and enables enrichment of ssDNA, conversion from ssDNA to dsDNA and linearization required for PCR reactions [22]. It is noteworthy that, whatever the conditions tested, we did not succeed by PCR in directly amplifying *VP1* from purified ssDNA without pre-amplification by MDA. Pre-amplified genomic DNAs were used as a template for PCR. PCRs were carried out in a DNA Thermal Cycler T-Professional (Biometra) to amplify a 800 bp major capsid protein encoding gene *VP1* using primer set microVP1-F1/R2 [21]. The mix was made with 25 μL of 1X PCR buffer, 4 mM $MgCl_2$, 25 μg of bovine serum albumin (Invitrogen), 200 mM of dNTP, 1 μM of each primers, 1 U of Platinum® Taq DNA polymerase (Invitrogen) and 1 μL of DNA template. The PCR program started with initial denaturation of 5 min at 95 °C, followed by 30 cycles of denaturation at 95 °C for 30 s, annealing at 54 °C for 30 s and elongation at 72 °C for 50 s, and a final elongation for 5 min at 72 °C. It is noteworthy that some samples (July and October for Lake Bourget) could not be amplified despite MDA protocol.

Amplicons were purified using the QIAquick PCR purification kit (Qiagen), and then cloned to pCR®4-TOPO® vectors using the TOPO TA Cloning kit (Invitrogen). On an average, 24 selected clones for each monthly sample were sent to GATC Biotech (Constance, Germany) for sequencing. After trimming and correcting sequences using BioEdit 7.0.5.3, we obtained 99 distinct non-redundant nucleotide sequences, with nucleotide similarity varying between 41.8% and 99.9%, following pairwise alignments using SDT [28]. Sequences were deposited in GenBank under reference accession numbers KR092804 to KR092902.

2.6. Phylogeny

The obtained nucleotide sequences were translated into amino acids using the online translate tool (http://www.hiv.lanl.gov/content/sequence/TRANSLATE/translate.html). These deduced amino acid sequences were aligned using MAFFT version 7 [17], with representatives of the cultured *Microviridae* and environmental sequences of different defined clusters from other studies [13,19–21,39,53,62]. Multiple alignments were then curated using Gblocks [7] employing a less stringent option that allowed for gaps inside the final blocks. We constructed phylogenies using both the Bayesian inference and maximum likelihood methods. Bayesian inference was conducted using MrBayes 3.2.1 [35] with two runs, four chains, three millions generations, sampling every 100 generations, a burn-in value of 25% and mixed models of amino acid substitution. The maximum likelihood phylogeny was constructed using PhyML 3.0 [11] with 100 bootstrap replicates and with the best model of acid amino substitution and rate heterogeneity. The best model for the aligned sequence dataset was determined using MEGA 6 [47] and was the LG model and gamma-distributed substitution rates.

2.7. PCR-DGGE

DGGE was used to reveal the community structure of *Microviridae* in both lakes. Prior to DGGE, a separate PCR was conducted for each sample as described above but using the GC primers (the same primers, but with the addition of a GC clamp of 40 nucleotides on the reverse primer), to add a GC tail on the amplicons. These amplicons were then purified using the QIAquick PCR purification kit (Qiagen).

DGGE was conducted in 6% polyacrylamide gels with optimized linear denaturing gradient (100% denaturant is defined as 7 M urea and 40% deionized formamide). The linear denaturing gradient was 20–60% for *VP1* amplicons. Approximately 200 ng DNA (quantified using a Nanodrop 1000 Spectrophotometer (Thermo Scientific)) were loaded into wells with 5 μl of 5 X loading buffer (12.5% Ficoll, 25 mM tris, 5 mM EDTA [pH 8.0], 0.5% SDS, 0.1% [wt/vol] xylene cyanol and 0.1% [wt/vol] bromophenol blue). Electrophoresis was carried out for 16 h in 1 X TAE buffer [pH 7.4] (40 mM Tris-base, 20 mM sodium acetate, 1 mM EDTA) at 120 V and a constant temperature of 60 °C using the CBS-DGGE 2000 system (C.B.S.Scientific Co., Inc.). Gels were stained in a 2 X SYBR Green I (Molecular Probles, Invitrogen) solution for 45 min and were visualized on a UV transilluminator (Tex-35M, Bioblock Scientific) and photographed with GelDoc (BioRad).

The DGGE banding patterns were analyzed using the GelCompare II software package (Applied Maths, Kortrijk, Belgium) after digitalization of the DGGE gels. We selected the bands manually, one by one, and not by using the automatic function of the software, since we noticed that automatic selection of bands could lead to underestimation. Two operators repeated this operation twice. Similarities between samples were estimated by Jaccard's similarity index based on the absence or presence of bands for which positions were discriminated unambiguously with 1% tolerance. The UPGMA method was used to construct the dendogram from a binary matrix of similarity values.

2.8. Statistics

To investigate the relationships between virus community structure and measured environmental variables, Canonical Correlation Analysis (CCA) was performed using software package XLSTAT-ADA. CCA generates an ordination plot that shows the main patterns of variation in community structure as accounted for measured environmental variables. Different variables were submitted to the forward selection procedure, in which the statistical significance of the term was tested by the unrestricted Monte Carlo permutation test (999 permutations). Explanatory variables with p-values greater than 0.05 were excluded from further analyses.

To evaluate whether the *VP1 Microviridae* assemblage (based on sequences) of Lake Annecy differed from that of Lake Bourget, we carried out statistical analyses using the UniFrac distance metric statistical tools available at http://unifrac.colorado.edu/ [24,25]. We used the unweighted

UniFrac option in order to compare community composition based on the presence/absence (i.e. on qualitative data) from the Bayesian phylogenetic tree. Finally, we conducted a P-test based on the UniFrac distance matrix generated for *VP1 Microviridae* assemblages of Lake Annecy and Lake Bourget.

3. Results

3.1. Diversity of ssDNA viruses

We obtained 99 unique sequences, among which 36 were exclusively found in Lake Bourget, 47 were detected only in Lake Annecy and 16 were common to both ecosystems. After phylogenetic analysis, all *VP1* sequences obtained from these two lakes were grouped into the *Microviridae* family, of which 92% clustered into *Gokushovirinae* subfamily and the other 8% were grouped with Bacteroidetes prophages of the *Alpavirinae* subfamily (bootstrap support value = 58; Fig. 1). We were able to discriminate three major clusters (based on both Bayesian inference posterior probabilities and maximum-likelihood bootstrap values at nodes superior to 20) in the gokushoviruses for a large proportion of sequences, referred to as LAB-Sub1, LAB-Sub2 and LAB-Sub3 that included 26%, 36% and 21% of total *VP1* sequences, respectively. Clusters LAB-Sub2 and LAB-Sub3 contained *VP1* sequences of gokushovirus genomes obtained from Lake Bourget by Roux et al. (2011b) which were previously assigned to clade "Lake-1″ as described in Labonté et al. (2015). Yet, LAB-Sub1 could be a new evolutionary cluster, as it does not include any *VP1* sequences of known gokushoviruses genomes. When examining the habitat/origin of these *VP1* sequences, we could not discriminate them based on ecosystem, season or month. However, when all *VP1* sequences were pooled, Unifrac analysis revealed significant differences between Lake Annecy and Lake Bourget ($p = 0.01$), suggesting distinct community compositions between the two lakes.

3.2. Structure of the ssDNA virus community

The number of *VP1* DGGE bands varied between 1 (February) and 13 (May) in Lake Annecy and between 5 (March and May) and 13 (June) in Lake Bourget (Fig. 2). We obtained 36 and 19 different bands for Lake Annecy and Lake Bourget, respectively. For both lakes, a total of 43 unique DGGE bands were obtained, among which 12 were common to both ecosystems, and the band patterns varied through time between the lakes. Only one band could be detected in all samples from the two lakes. DGGE banding pattern analysis (Fig. 2) revealed the existence of four clusters for *Microviridae* assemblages in Lake Annecy, discriminating the end of summer and autumn (September, October, November) from the summer months (July, August), with a part of the spring period (symbolized by March and May) and the other months corresponding to the end of winter and spring (February, April, June). However, such seasonality was relatively weak, since a significant difference was consistently recorded between each month, also suggesting a very dynamic structure over a shorter time scale (see Discussion). For Lake Bourget, samples were grouped into two main clusters, from late winter to mid-spring and from late-spring to autumn, with 25% similarity between the two. There was no clear seasonal pattern.

3.3. Dynamic patterns of individual ssDNA viruses

We examined the dynamics of individual constituents of *Microviridae* (corresponding to DGGE bands) for each sample obtained throughout the year. We detected variability in intensity in nearly all bands. The relative abundance of a band could increase or decrease up to 29-fold between two consecutive sampling dates. As an example, the relative abundance of band b12 increased from 2.6 to 74.1% between April and May in Lake Bourget (Fig. 3, Fig. S1). Some other bands also switched from undetectable to dominant and vice versa. For instance, band b10 'bloomed' in April in Lake Bourget, but disappeared just after (Fig. 3, Fig. S1). Overall, we observed three different dynamic patterns for the different *VP1* bands: (i) a "bloom-bust" pattern (a dramatic increase or decrease with a sharp peak, either episodic or very occasional); (ii) relatively stable fluctuations (within relatively narrow limits); (iii) only gradual progression through time (including seasonal variation, monotonic increase or decrease). Often, different patterns in the two lakes could be observed for a selected band (e.g. b3, b6, b10, b12, etc.; Fig. S1). However, similar dynamics could also be detected in both lakes, but the amplitude and timing of the increase/decline was different (e.g. b1, b7, b13, etc.; Fig. S1); b13, for instance, boomed from undetectable in June in both lakes, while it started to re-bloom in Lake Annecy in October, but not in Lake Bourget; b1 increased from undetectable during September–October in both lakes, but the amplitude of the change was different (relative abundance increase to 24.2% in Lake Bourget, but only up to 11% in Lake Annecy); b7 began to boom earlier in Lake Bourget (August) than in Lake Annecy (October).

The examination of each monthly sample revealed that *Microviridae* virus assemblage was dominated by only a few bands (Fig. S2). The most abundant bands accounted for up to 89.5% of total *Microviridae* virus abundance in Lake Annecy (e.g. b109 in August) and up to 74.1% in Lake Bourget (e.g. b12 in May). We found that the identity of the dominant bands changed with time and differed between lakes, as did band dominance. When examining the dynamics of 13 of these dominant bands obtained from either Lake Annecy or Bourget across the year (Fig. S2), we found that: (i) they were detected only at certain times of the year (e.g. b7 in autumn in both lakes, etc.) or even just once (e.g. b10 in Lake Bourget in April; b115 in Lake Annecy in June), while none of them consisted of persistent bands (detectable throughout the year); (ii) most of them dominated only once in our samples, (iii) a few others could reoccur as dominants (e.g. b104 in Lake Annecy); (v) a majority (4 out of 5) of the dominant bands of Lake Bourget (b7, b10, b12 and b15) could be detected in Lake Annecy, while none of the predominant bands of Lake Annecy (b104, b105, b106, b109, b112, b115, b118 and b112) were detected in Lake Bourget.

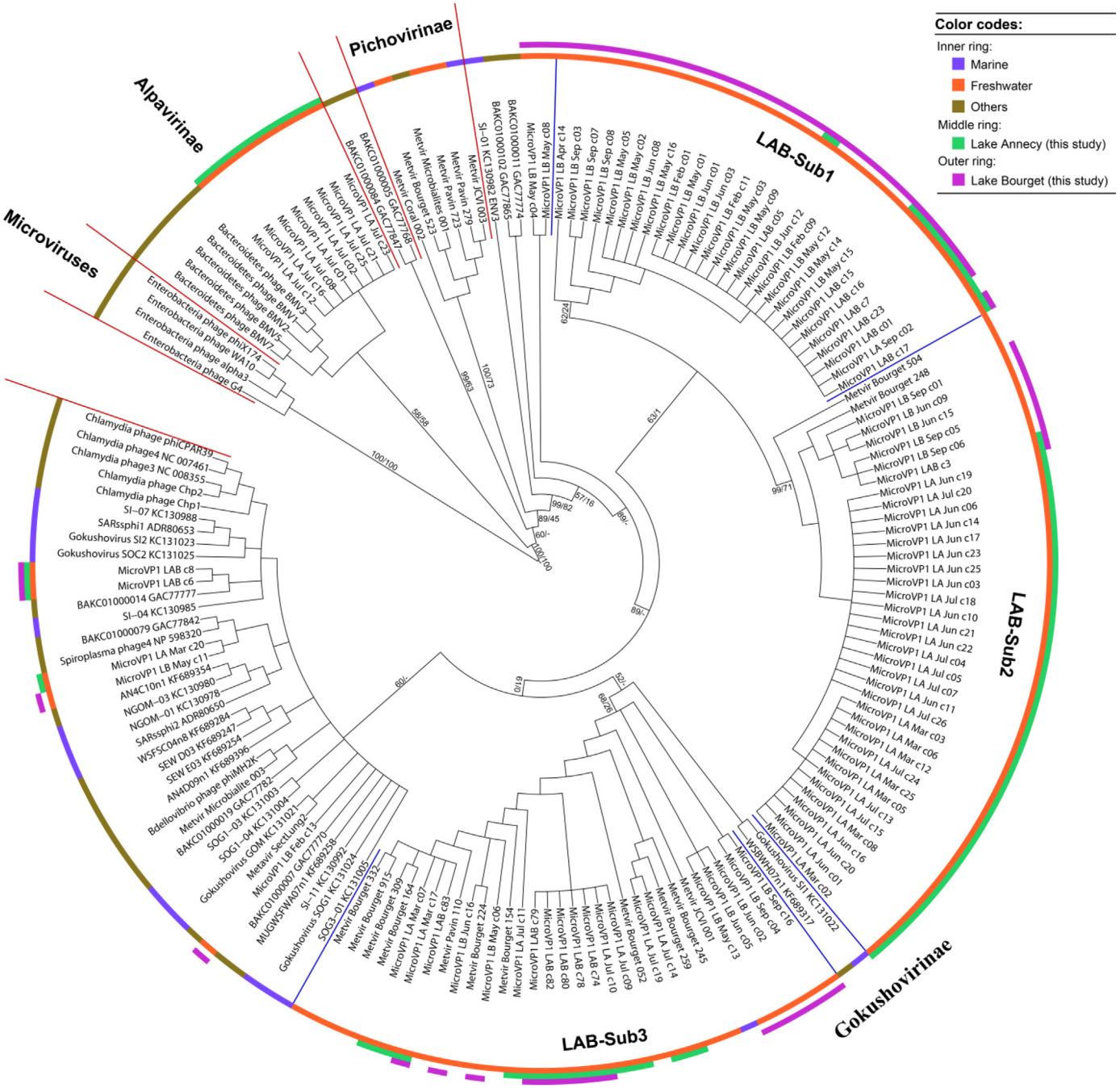

Fig. 1. Bayesian phylogenetic tree based on alignment of 194 homologous amino acid positions of the major capsid protein encoding gene *VP1* from 167 *VP1* sequences. Sequences are from lakes Annecy and Bourget, *Microviridae* isolates and other environments (e.g. representative sequences of different defined clusters from other studies). Values at nodes are the Bayesian inference (BI) posterior probabilities and maximum likelihood (ML) bootstrap values (BI/ML). The inner ring represents *VP1* sequences of the marine field (violet), freshwater (orange) and the others (antique bronze). The middle ring with color strips in green corresponds to sequences from Lake Annecy. In the outer ring, sequences labeled with color strips (purple) are from Lake Bourget. "Others" in the color key panel indicates *VP1* sequences obtained from marine and freshwater sediments, microbialites, sewage water and human gut, mouth, feces or tissue.

3.4. Viral assemblage structure in relation to biotic and abiotic factors

We examined the relationships between *VP1* assemblage and environmental and biological variables using canonical correspondence analysis (CCA), and then tested the significance of the relationships using Monte Carlo permutations (999 iterations). Six descriptors explained 53.5% of the variance in *VP1* ssDNA virus assemblage in Lake Bourget. These factors were the temperature and the abundance of VLP1, heterotrophic bacteria, *Synechococcus*, PO_4 and NH_4. Comparatively, for Lake Annecy, the number of variables was

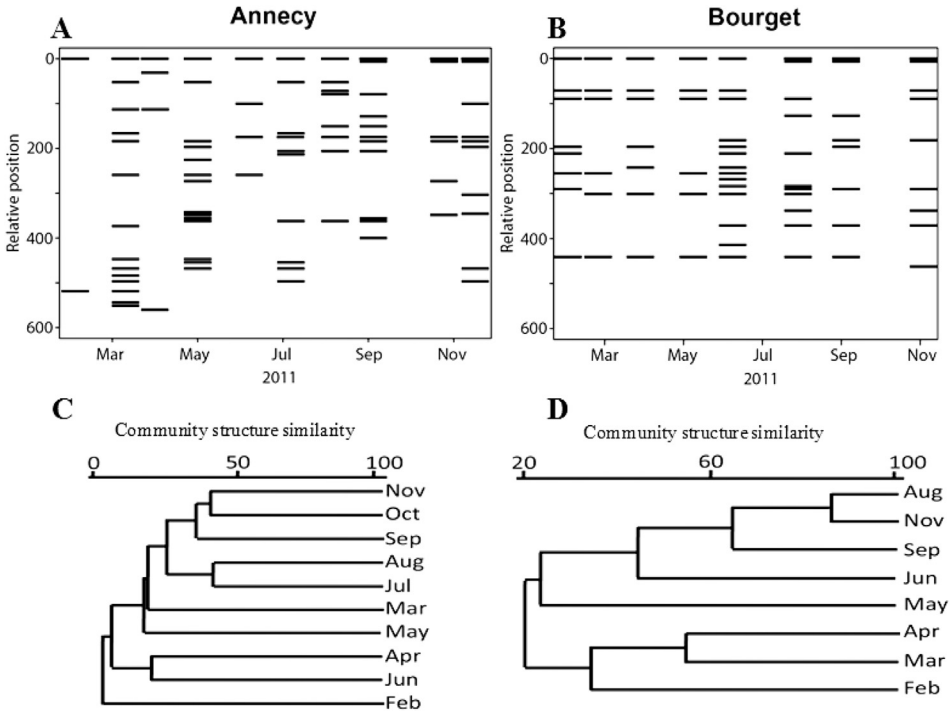

Fig. 2. Schematic outline of the presence/absence of the *VP1* DGGE bands in the two lakes (A, B) during 2011. Cluster analyses of *VP1* DGGE fingerprinting patterns (C, D), obtained using UPGMA clustering of Jaccard's similarity matrix based on the presence/absence of bands.

seven (plus NO_3) and explained only 40.7% of the variance of *VP1* ssDNA virus assemblage. Each variable could play a (direct or indirect) role at different periods of the year for each lake (Fig. 4, see Discussion).

4. Discussion

During the last decade, we have explored viral ecology (abundance, dynamics, distribution, lysis vs. lysogeny) [16,31,44,49,58] and diversity [31,57,59–61] in peri-alpine lakes. Our previous works focused only on dsDNA viruses, and we showed that this community is abundant, highly diverse and likely to play an important role in the functioning of these lakes (e.g. Lakes Annecy, Bourget and Geneva). However, in 2012, Roux and colleagues, for one of these peri-alpine lakes (Lake Bourget), showed that ssDNA viruses were largely underestimated and might form the bulk of viral diversity, with more than 90% sequence representatives. In a parallel study, Holmfeldt et al. (2012) also reported that ssDNA viruses were largely underestimated using conventional techniques such as flow cytometry or microscopy, or molecular tools more suited to targeting dsDNA viruses. Our study is one of the rare studies indicating that ssDNA viruses are likely an important component of the phage community in lakes, the role of which remains to be determined.

Our results suggest that microviruses are diverse in (peri-alpine) lakes. The number of *VP1* genotypes obtained throughout a complete year, as revealed by both cloning/sequencing and PCR-DGGE, was higher in Lake Annecy than in Lake Bourget. This may be related higher richness of potential hosts of *Microviridae* in oligotrophic Lake Annecy compared to mesotrophic Lake Bourget. This observation is in contrast to most of the examined dsDNA viruses (cyanophages, T4-like myophages and phycodnaviruses; [59]) for which we had previously reported an inverse result for the same survey.

When checking the phylogenetic affiliations of the *VP1* sequences, 92% grouped together with gokushoviruses, for which the primers had been initially designated [21]. However, 8% of *VP1* sequences were closely related to Bacteroidetes phages of the *Alpavirinae*, a subfamily of *Microviridae* recently proposed [19]. It is noteworthy that these Bacteroidetes ssDNA viruses (BMV1 to BMV7) have not been isolated thus far, and they have been referred to prophages infecting members of the phylum Bacteroidetes (i.e. *Bacteroides eggerthii*, *Bacteroides plebeius*, *Prevotella bergensis* and *Prevotella buccalis*) obtained from the human gut, mouth, feces and tissues [19]. Our finding is thus the first detection of such ssDNA *Alpavirinae* viruses (either lysogenic or lytic) in aquatic environments using a PCR-based approach. The potential hosts of these ssDNA *Alpavirinae* viruses could be members of the Bacteroidetes. This is not impossible, since a previous metagenomic study [8,14] suggested that Bacteroidetes were indeed diverse, the third most abundant bacterial phylum in Lakes Annecy and Bourget. More interestingly, these *Alpavirinae* were only recorded once in Lake Annecy in July, suggesting specific geographic and temporal distribution for these viruses. This could be explained by the fact that *Alpavirinae* are lysogenic and were induced in July in Lake Annecy in response to favorable conditions.

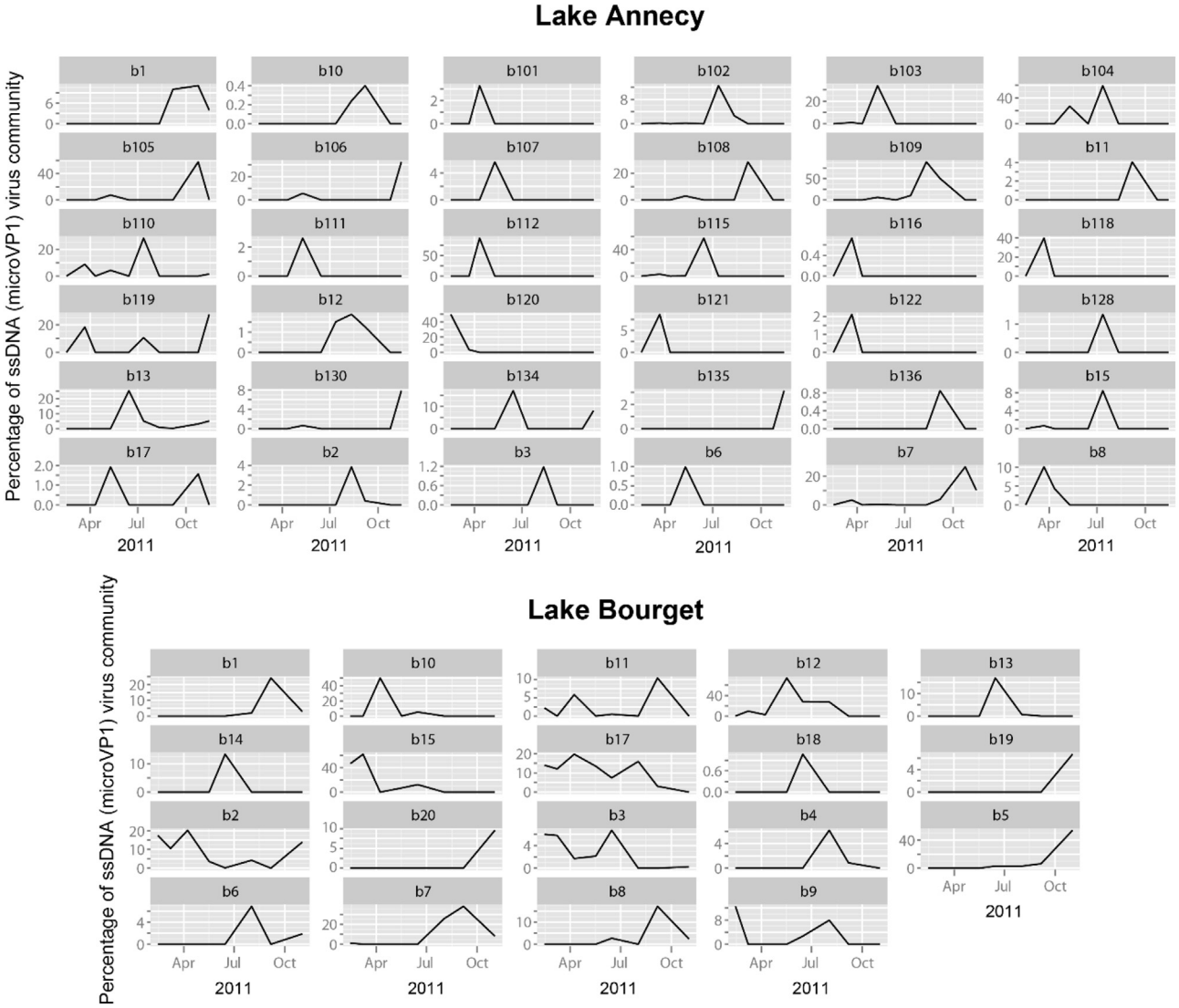

Fig. 3. Dynamics of all *Microviridae VP1* DGGE bands in 2011 for the two lakes. The Y-axis indicates the relative abundance for each band, i.e. the normalized percent contribution of each band to the total ssDNA virus assemblage in the sample.

In this study, we did not obtain sequences related to *Pichovirinae*, another new group of *Microviridae* proposed by Roux et al. (2011b), which is characterized by a distinct genomic organization within core genes compared to the three other subfamilies (*Alpavirinae*, *Microvirus* and *Gokushovirinae*). This does not mean that *Pichovirinae* are absent in perialpine lakes. Indeed, the metagenomic study of Roux et al. (2012a) for a single sample obtained from Lake Bourget in July 2008 revealed significant numbers of sequences related to *Pichovirinae,* and those authors succeeded in assembling them in an intact genome named "Bourget_523 or MET7_523" [39]. After analysis of the sequence, we noticed that the MicroVP1 primers do not match VP1 of *Pichovirinae* Bourget_523. Thus, the design of other primers could be very helpful for more specifically targeting these viruses.

We observed highly dynamic structures for *Microviridae* communities in Lakes Annecy and Bourget. We measured changes in the identity of predominant hosts and modifications in the relative abundance of individual constituents and rapid shifts in the community structure/composition across the year which also differed between lakes. Indeed, we observed: (i) different community compositions and identities for the dominant bands; (ii) distinct dynamic patterns for individual community constituents; (iii) seasonality in Lake Annecy, but not in Lake Bourget. This could reflect the different trophic status of the two lakes where distinct virus and potential host communities exist [59]. A high proportion of bands (88.9% of the bands in Lake Annecy and 68.4% of the bands in Lake Bourget; Fig. 3) displayed a boom-bust pattern (i.e. a sharp increase/decrease, potentially indicative of the *killing the winner* strategy; [48]). Most of these bands were marked by a single sharp peak (particularly in Lake Annecy; e.g. b10, b12, b14, b109, b112, b115, etc.; Fig. 3). It seems likely that these *Microviridae* were opportunistic viruses (either lytic or lysogenic, able to emerge as dominant from rare/undetectable when conditions were favorable), suggesting they play a key

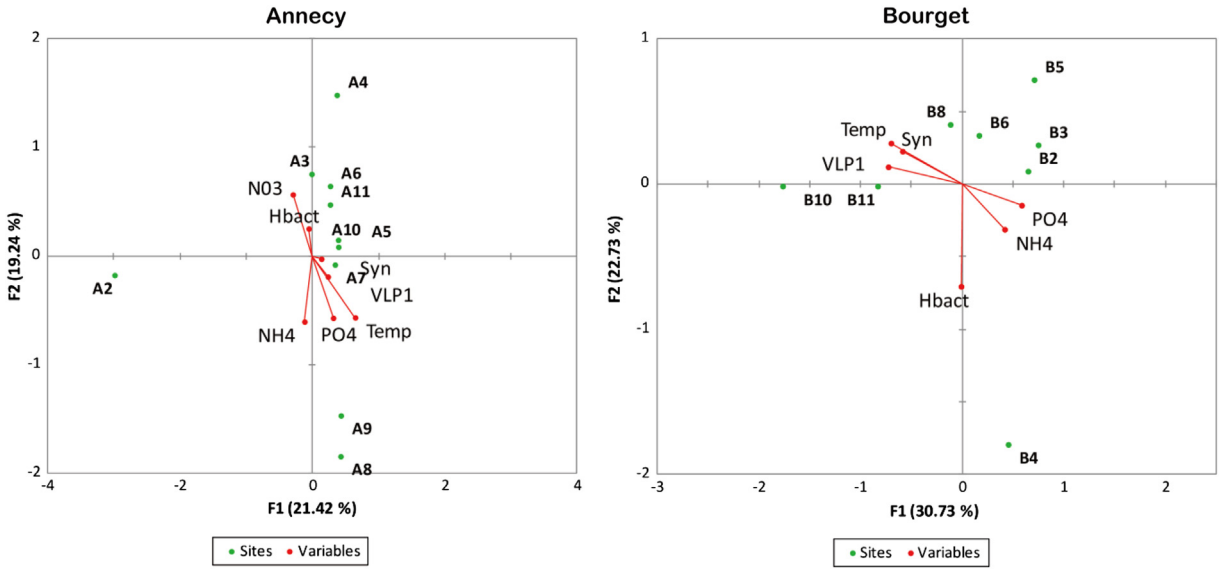

Fig. 4. CCA using both environmental variables and biological counts as constrained variables (Temp = water temperature; PO_4; NO_3; NH_4; VLP1 = virus-like particles group 1; Syn = PE-rich picocyanobacteria belonging to *Synechococcus* spp; Bact = heterotrophic bacteria) to explain temporal changes in ssDNA virus assemblage structure (based on presence/absence of DGGE band and relative abundance). The sample codes give the lake (A = Annecy, B = Bourget) and the month of sampling (2–11). A6 is the sample taken in June from Lake Annecy and B6 is thus the sample taken in June from Lake Bourget.

role in host mortality and community structuring. Interestingly, these viruses could display similar patterns in both lakes (e.g b13 in June; Fig. S1), while, in most cases, they differed (e.g. b10, b12, b15, b17, b2). This suggests that ssDNA viruses act differently toward their potential hosts in each lake where biotic and abiotic influence differ.

In comparison to dsDNA viruses that also infect some bacterial hosts (e.g. T4-like myoviruses; [57]), we found in *Microviridae* a higher relative abundance when a band boomed (up to 89.5%; Fig. S2, Fig. 3). This may be explained by the fact that ssDNA viruses are smaller in size and genome (up to 8.5 kb), and consequently with a short latent period, thus resulting in high burst size [20,21]. However, we may not rule out bias associated with MDA which may unevenly amplify some single-stranded DNAs [22] and also the DGGE where one band may contain multiple *VP1* genotypes/sequences [59], so that the high relative abundance recorded for some DGGE bands might be an artifact of amplification or the contribution of several different ssDNA viruses. These biases could also be true for analysis of community structure (Fig. 2), the dynamics of individual constituents (Fig. 3) and CCA analysis (Fig. 4). However, because we used temporal resolution to investigate ssDNA virus community dynamics, assuming the bias related to MDA is random, our results are likely to reflect overall ssDNA virus community structure and dynamics.

CCAs revealed that biotic environmental variables explained between 40.7% and 53.5% of the variance of *Microviridae* community structures. Potential bacterial hosts (i.e. the heterotrophic bacteria *Synechococcus* spp.), viral group VLP1 and nutrients (PO_4 and NH_4 in both lakes, NO_3 in Lake Annecy) were retained as significant variables (i.e. $P < 0.05$) mediating the relationships between virus assemblage, host abundance and environmental factors. In general, the number and relative contribution of variables shaping the ssDNA virus assemblage structure were not dramatically different between the two lakes. Rather, these variables acted differently on each viral group at different times of the year in each lake. This may explain a large portion of the temporal variability in viral community structure in each lake. The fact that we found some relationships between the abundances of VLP1 with the ssDNA virus assemblage is surprising, since ssDNA viruses are too small to be counted in FCM. This suggests that some ssDNA viruses were counted with VLP1 (i.e. both ssDNA and dsDNA viruses were associated with small-genome-size bacteriophages, e.g. <80 kb, in this FCM signature). Another possibility that would explain possible relationships between VLPs and ssDNA viruses is that associations exist between these communities that remain to be determined. This hypothesis makes sense in light of results recently obtained by [40], who detected, for SUP05 bacteria (i.e. chemoautotrophic bacteria that drive sulfur cycling and energy flow in oxygen minimal zones of the ocean), co-infection by *Microviridae-Caudovirales*, suggesting that some hosts could be "shared" by different viral types (e.g. ssDNA viruses and dsDNA viruses). Finally, we should bear in mind that this result could simply be an artifact due to technical bias associated with DGGE and/or MDA.

The correlations found between the abundance of heterotrophic prokaryotes and ssDNA virus assemblage were expected, as bacteria are well-known potential hosts for these viruses. Currently known *Microviridae* (either the lytic isolates or the prophage genome obtained from lysogenized host genomes) have been detected in the bacterial phyla of the Proteobacteria, Tenericutes, Chlamydiae and Bacteroidetes [5,12,19,22,40], and these groups are indeed important in

these lakes [8,15]. It is noteworthy that a recent study conducted in coastal seawaters revealed diverse ssDNA viruses infecting SUP05 (a clade of sulfur oxider within gamma-Proteobacteria), and suggested that the host range of gokushoviruses is much wider. Interestingly, we also found *Synechococcus* abundance to be correlated with changes in ssDNA virus assemblage, also suggesting OTUs of this community to be potential hosts. This is not impossible, since one *Microviridae* prophage has already been detected for *Synechococcus* spp. by [26]. However, the biological relationship we found was relatively weak and likely to point to missing variables. For instance, ssDNA viruses have been shown to infect diatoms [29,50–52] and Archaea [32–34,41].

Among abiotic factors, water temperature, PO_4 and NH_4 seemed to be related to *Microviridae* assemblage composition in both lakes, but they acted differently over the months. Temperature is an important factor, probably acting as a primary factor driving host growth, temporal changes in host availability and viral decay, thereby acting directly and indirectly on viral community structure [6,59]. Phosphorus is a limiting nutrient for both autotrophs and heterotrophs in oligotrophic lakes such as Lake Annecy, and is seasonally depleted even in Lake Bourget. Our results suggested that PO_4 could explain part of the variation in the assemblage of *Microviridae* in mid-summer to early autumn in Lake Annecy and in the spring in Lake Bourget. This was likely to be related to the rapid growth of heterotrophic bacteria following proliferation of phytoplankton and nutrient depletion. CCA analysis also suggested NH_4 to be related to the ssDNA viral structure. This result echoes the study of [42], who showed that bacterial lysis can result in ammonium production through the liberation of dissolved organic N re-mineralized by uninfected bacteria that, in turn, can fuel primary production. Note that we also reported this possible effect for dsDNA viruses in a previous study [59]. Lastly, the NO_3 concentration was also associated with the ssDNA virus structure, but only in Lake Annecy. From the same samples, we did not detect such a relationship for *Microviridae* in Lake Bourget or for previously examined dsDNA viruses (e.g phycodnaviruses, cyanophages, T4-like myoviruses) in the two lakes. The mechanism is a puzzle, but such a result could suggest bottoms-up control by this resource only in oligotrophic ecosystems. Once again, it is noteworthy that both DGGE (see details in [59]) and MDA (uneven amplification of initial template either among distinct taxa from a sample or the same taxon from different samples; see [22]) are biased methods. We are aware that our interpretation concerning community structure, the dynamics of individual constituents and CCA analysis (that links environmental factors to community structure) should thus be taken with caution.

In conclusion, we reveal a diverse and variable ssDNA virus community whose composition and temporal dynamic patterns for individual components could be very different between lakes. The fact that ssDNA viruses replicate/multiply quickly may explain our observation of the significant variations in community composition and the relative abundance of individual constituents, the high proportion of bands with boom-bust patterns and the dramatic and sharp increases in their relative abundance when booming. Unique sequences were also discovered. Futures studies should: (i) monitor specific *Microviridae* and their hosts using qPCR, for instance, to reveal their importance in mortality and community structuring; (ii) examine how important lysogeny can be in ssDNA viruses; (iii) study other groups of ssDNA viruses and their potential hosts from which a limited number of isolates have been obtained in aquatic environments.

Conflict of interest

The Authors declare that there is no conflict of interest.

Acknowledgments

We thank Drs. Yves Desdevises and Simon Roux for stimulating discussion and improvement of a former version of this article. TEM data were provided by J. Colombet. We particularly wish to thank two anonymous reviewers who helped us to construct this article.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.resmic.2015.07.003.

References

[1] Angly F, Felts B, Breitbart M, Salamon P, Edwards R, Carlson C, et al. The marine viromes of four oceanic regions. PLoS Biol 2006;4:2121–31.
[2] Ackermann HW. 5500 Phages examined in the electron microscope. Arch Virol 2007;152:227–43.
[3] Ackermann HW, Prangishvili D. Prokaryote viruses studied by electron microscopy. Arch Virol 2012;157:1843–9.
[4] Breitbart M, Salamon P, Andresen B, Mahaffy JM, Segall AM, Mead D, et al. Genomic analysis of uncultured marine viral communities. Proc Natl Acad Sci U S A 2002;99:14250–5.
[5] Brentlinger K, Hafenstein S, Novak C, Fane B, Borgon R, McKenna R, et al. *Microviridae*, a family divided: isolation, characterization, and genome sequence of fMH2K, a bacteriophage of the obligate intracellular parasitic bacterium *Bdellovibrio bacteriovorus*. J Bacteriol 2002;184:1089–94.
[6] Callieri C, Caravati E, Corno G, Bertoni R. Picocyanobacterial community structure and space-time dynamics in the subalpine Lake Maggiore (N. Italy). J Limnol 2012;71:95–103.
[7] Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol 2000;17:540–52.
[8] Debroas D, Humbert J-F, Enault F, Bronner G, Faubladier M, Cornillot E. Metagenomic approach studying the taxonomic and functional diversity of the bacterial community in a mesotrophic lake (Lac du Bourget-France). Environ Microbiol 2009;11:2412–24.
[9] Desnues C, Rodriguez-Brito B, Rayhawk S, Kelley S, Tran T, Haynes M, et al. Biodiversity and biogeography of phages in modern stromatolites and thrombolites. Nature 2008;452:340–3.
[10] Fancello L, Trape S, Robert C, Boyer M, Popgeorgiev N, Raoult D, et al. Viruses in the desert: a metagenomic survey of viral communities in four perennial ponds of the Mauritanian Sahara. ISME J 2013;7:359–69.
[11] Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 2010;59:307–21.

[12] Holmfeldt K, Odic D, Sullivan MB, Middelboe M, Riemann L. Cultivated single-stranded DNA phages that infect marine Bacteroidetes prove difficult to detect with DNA-binding stains. Appl Environ Microbiol 2012;78:892–4.
[13] Hopkins M, Kailasan S, Cohen A, Roux S, Tucker KP, Shevenell A, et al. Diversity of environmental single-stranded DNA phages revealed by PCR amplification of the partial major capsid protein. ISME J 2014;8:2093–103.
[14] Humbert J-F, Dorigo U, Cecchi P, Le Berre B, Debroas D, Bouvy M. Comparison of the structure and composition of bacterial communities from temperate and tropical freshwater ecosystems. Environ Microbiol 2009;11:2339–50.
[15] Jacquet S, Dorigo U, Personnic S. A few tests prior to flow cytometry and epifluorescence analyses of freshwater bacterio- and virioplankton communities [Chapter 1]. In: Papandreou Stefanos, editor. Flow cytometry: principles, methodology and applications. New York: Nova Science Publishers; 2013. p. 1–30.
[16] Jacquet S, Domaizon I, Personnic S, Duhamel S, Pradeep-Ram AS, Heldal M, et al. Estimates of protozoan- and viral-mediated mortality of bacterioplankton in Lake Bourget. (France). Fresh Biol 2005;50:627–45.
[17] Katoh K, Misawa K, Kuma K, Miyata T. MAFFT, a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nuc Acids Res 2002;30:3059–66.
[18] Kim K, Chang H, Nam Y, Roh SW, Kim M, Sung Y, et al. Amplification of uncultured single-stranded DNA viruses from rice paddy soil. Appl Environ Microbiol 2008;74:5975–85.
[19] Krupovic M, Forterre P. Microviridae goes temperate: microvirus-related proviruses reside in the genomes of Bacteroidetes. PLoS One 2011;6:e19893.
[20] Labonté JM, Suttle CA. Previously unknown and highly divergent ssDNA viruses populate the oceans. ISME J 2013a;7:2169–77.
[21] Labonté JM, Suttle CA. Metagenomic and whole-genome analysis reveals new lineages of gokushoviruses and biogeographic separation in the sea. Front Microbiol 2013b;4:404.
[22] Labonté JM, Suttle Curtis A. Corrigendum: metagenomic and whole-genome analysis reveals new lineages of Gokushoviruses and biogeographic separation in the sea. Front Microbiol 2015;6:114.
[23] López-Bueno A, Tamames J, Velázquez D, Moya A, Quesada A, Alcamí A. High diversity of the viral community from an Antarctic Lake. Science 2009;326:858–61.
[24] Lozupone C, Knight R. UniFrac, a new phylogenetic method for comparing microbial communities. Appl Environ Microbiol 2005;71:8228–35.
[25] Lozupone CA, Hamady M, Kelley ST, Knight R. Quantitative and qualitative beta diversity measures lead to different insights into factors that structure microbial communities. Appl Environ Microbiol 2007;73:1576–85.
[26] McDaniel LD, delaRosa M, Paul JH. Temperate and lytic cyanophages from the Gulf of Mexico. J Mar Biol Assoc UK 2006;86:517–27.
[27] McDaniel L, Rosario K, Breitbart M, Paul J. Comparative metagenomics: natural populations of induced prophages demonstrate highly unique, lower diversity viral sequences. Environ Microbiol 2014;16:570–85.
[28] Muhire BM, Varsani A, Martin DP. SDT: a virus classification tool based on pairwise sequence alignment and identity calculation. PLoS One 2014;9:e108277. http://dx.doi.org/10.1371/journal.pone.0108277.
[29] Nagasaki K, Tomaru Y, Takao Y, Nishida K, Shirai Y, Suzuki H, et al. Previously unknown virus infects marine diatom. Appl Environ Microbiol 2005;71:3528–35.
[30] Parvathi A, Zhong X, Jacquet S. Dynamics of various viral groups infecting autotrophic plankton in Lake Geneva. Adv Limnol Oceanog 2012;3:171–91.
[31] Personnic S, Domaizon I, Dorigo U, Berdjeb L, Jacquet S. Seasonal and spatial variability of virio-, bacterio-, and picophytoplanktonic abundances in three peri-alpine lakes. Hydrobiol 2009;627:99–116.
[32] Pietilä MK, Demina TA, Atanasova NS, Oksanen HM, Bamford DH. Archaeal viruses and bacteriophages: comparisons and contrasts. Trends Microbiol 2014;22:334–44.
[33] Roine E, Kukkaro P, Paulin L, Laurinavicius S, Domanska A, Somerharju P, et al. New, closely related haloarchaeal viral elements with different nucleic acid types. J Virol 2010;84:3682–9.
[34] Pietilä MK, Roine E, Paulin L, Kalkkinen N, Bamford DH. An ssDNA virus infecting archaea: a new lineage of viruses with a membrane envelope. Mol Microbiol 2009;72:307–19.
[35] Ronquist F, Teslenko M, Van der Mark P, Ayres DL, Darling A, Höhna S, et al. MrBayes 3.2, efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol 2012;61:539–42.
[36] Rosario K, Duffy S, et al. Diverse circovirus-like genome architectures revealed by environmental metagenomics. J Gen Virol 2011;90:2418–24.
[37] Rosario K, Nilsson C, Lim YW, Ruan Y, Breitbart M. Metagenomic analysis of viruses in reclaimed water. Environ Microbiol 2009;11:2806–20.
[38] Roux S, Enault F, Robin A, Ravet V, Personnic S, Theil S, et al. Assessing the diversity and specificity of two freshwater viral communities through metagenomics. PLoS One 2012a;7:e33641.
[39] Roux S, Krupovic M, Poulet A, Debroas D, Enault F. Evolution and diversity of the *Microviridae* viral family through a collection of 81 new complete genomes assembled from virome reads. PloS One 2012b;7:e40418.
[40] Roux S, Hawley AK, Torres Beltran M, Scofield M, Schwientek P, Stepanauskas R, et al. Ecology and evolution of viruses infecting uncultivated SUP05 bacteria as revealed by single-cell- and meta- genomics. eLife 2014. http://dx.doi.org/10.7554/eLife.03125.
[41] Senčilo A, Jacobs-Sera D, Russell DA, Ko C-C, Bowman CA, Atanasova NS, et al. Snapshot of haloarchaeaal tailed virus genomes. RNA Biol 2013;10:803–16.
[42] Shelford EJ, Middelboe M, Møller EF, Suttle CA. Virus-driven nitrogen cycling enhances phytoplankton growth. Aqua Microb Ecol 2012;66:41–6.
[43] Sime-Ngando T, Mignot JP, Amblard C, Bourdier G, Devilettes C, Quilblier-Lloberas C. Characterization of planktonic virus-like particles in a French mountain lake: methodological aspects and preliminary results. Ann Limnol Int J Lim 1996;32:1–5.
[44] Sime-Ngando T, Colombet J, Personnic S, Domaizon I, Dorigo U, Perney P, et al. Short-term variations in abundances and potential activities of viruses, bacteria and nanoprotists in Lake Bourget (France). Ecol Res 2008;23:851–61.
[45] Smith AM, Adler FR, Ribeiro RM, Gutenkunst RN, McAuley JL, McCullers JA, et al. Kinetics of coinfection with influenza A virus and Streptococcus pneumoniae. PLoS Pathog 2013;9.
[46] Suttle CA. Marine viruses-major players in the global ecosystem. Nat Rev Microbiol 2007;5:801–12.
[47] Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5, molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum Parsimony methods. Mol Bio Evol 2011;28:2731–9.
[48] Thingstad TF. Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. Limnol Oceanogr 2000;45:1320–8.
[49] Thomas R, Berdjeb L, Sime-Ngando T, Jacquet S. Viral abundance, production, decay rates and life strategies (lysogeny versu lysis) in Lake Bourget (France). Environ Microbiol 2011;13:616–30.
[50] Tomaru Y, Shirai Y, Suzuk IH, Nagumo T, Nagasaki K. Isolation and characterization of a new single-stranded DNA virus infecting the cosmopolitan marine diatom *Chaetoceros dehilis*. Aqua Microb Ecol 2008;50:103–12.
[51] Tomaru Y, Takao Y, Suzuki H, Nagumo T, Koike K, Nagasaki K. Isolation and characterization of a single-stranded DNA virus infecting *Chaetoceros lorenzianus* Grunow. Appl Environ Microbiol 2011;77:5285–93.
[52] Tomaru Y, Takao Y, Suzuki H, Nagumo T, Nagasaki K. Isolation and characterization of a single-stranded RNA virus infecting the bloom-forming diatom *Chaetoceros socialis*. Appl Environ Microbiol 2009;75:2375–81.
[53] Tucker KP, Parsons R, Symond EM, Breitbart M. Diversity and distribution of single-stranded DNA phages in the North Atlantic Ocean. ISME J 2011;5:822–30.
[54] Vaulot D. CYTOPC: processing software for flow cytometric data. Signal Noise 1989;2:8.

[55] Wegley L, Edwards R, Rodriguez-Brito B, Liu H, Rohwer F. Metagenomic analysis of the microbial community associated with the coral *Porites astreoides*. Environ Microbiol 2007;9(11):2707–19.

[56] Weinbauer MG, Peduzzi P. Frequency, size and distribution of bacteriophages in different marine bacterial morphotypes. Mar Ecol Prog Ser 1994;108:11–20.

[57] Zhong X, Jacquet S. Differing assemblage composition and dynamics in T4-like myophages of two neighbouring sub-alpine lakes. Fresh Biol 2014;59:1577–95.

[58] Zhong X, Ram ASP, Colombet J, Jacquet S. Variations in abundance, genome size, morphology, and functional role of the virioplankton in lakes Annecy and Bourget over 1-year period. Microb Ecol 2014;67:66–82.

[59] Zhong X, Rimet F, Jacquet S. Seasonal variations in PCR-DGGE fingerprinted viruses infecting phytoplankton in large and deep peri-alpine lakes. Ecol Res 2014;29:271–87.

[60] Zhong X, Jacquet S. Prevalence of viral photosynthetic and capsid protein genes in two large and deep peri-alpine lakes: a new insight on the diversity of freshwater cyanophages. Appl Environ Microbiol 2013;79:7169–78.

[61] Zhong X, Jacquet S. Contrasting diversity of phycodnavirus signature genes in two large and deep western European lakes. Environ Microbiol 2014b;16:759–73.

[62] Yoshida M, Takaki Y, Eitoku M, Nunoura T, Takai K. Metagenomic analysis of viral communities in (Hado)pelagic sediments. PLoS One 2013;8:e57271.