Phycodnaviruses have a significant role in modulating the dynamics of phytoplankton, thereby influencing community structure and succession, nutrient cycles and potentially atmospheric composition because phytoplankton fix about half the carbon dioxide (CO₂) on the planet, and some algae release dimethylsulphoniopropionate when lysed by viruses. Despite their ecological importance and widespread distribution, relatively little is known about the evolutionary history, phylogenetic relationships and phylodynamics of the Phycodnaviruses from freshwater environments. Herein we provide novel data on Phycodnaviruses from the largest river system on earth—the Amazon Basin—that were compared with samples from different aquatic systems from several places around the world. Based on phylogenetic inference using DNA polymerase (pol) sequences we show the presence of distinct populations of Phycodnaviridae. Preliminary coarse-grained phylodynamics and phylogeographic inferences revealed a complex dynamics characterized by long-term fluctuations in viral population sizes, with a remarkable worldwide reduction of the effective population around 400 thousand years before the present (KYBP), followed by a recovery near to the present time. Moreover, we present evidence for significant viral gene flow between freshwater environments, but crucially almost none between freshwater and marine environments.

**Introduction**

The family Phycodnaviridae comprises six genera that include several large and diverse icosahedral double-stranded DNA viruses, with genomes ranging from 160 kbp to 560 kbp (Van Etten et al., 2002; Dunigan et al., 2006; Wilson et al., 2008), which infect a wide spectrum of eukaryotic algae. Phycodnaviruses have a significant role in modulating the dynamics of phytoplankton, thereby influencing the community structure and succession (Castberg et al., 2001; Brussaard, 2004), nutrient cycles (Gobler et al., 1997; Wilhelm and Suttle, 1999; Rusch et al., 2007) and potentially atmospheric composition. This is because phytoplankton fix about half of the CO₂ on the planet (Sabine et al., 2004; Deman et al., 2007; Houghton, 2007), and some algae release dimethylsulphoniopropionate when lysed by viruses (Liss et al., 1997; Hill et al., 1998; Malin et al., 1998). Despite their ecological importance and widespread distribution (Cottrell and Suttle 1991; Clasen and Suttle, 2009), relatively little is known about the evolutionary history and phylogenetic relationships among the Phycodnaviruses (Chen and Suttle, 1996; Short and Suttle, 2002; Wilson et al., 2006; Larsen et al., 2008) particularly as they pertain to freshwater environments (Short and Short, 2008; Clasen and Suttle, 2009). Moreover, it has been postulated that, based on molecular phylogenies, viruses and other microorganisms have experienced a limited number of marine–freshwater transitions during their evolution (Logares et al., 2009).

The Phycodnaviridae family was recognized based on a group of viruses that infect endosymbiotic Chlorella-like algae (Van Etten and Ghabrial, 1991), and is now known to be part of the large nuclear and cytoplasmic DNA virus group. These viruses encode several conserved proteins performing most key life-cycle processes. The choice for using the B-family DNA polymerase as a conserved marker for the evolutionary history of nuclear and cytoplasmic DNA virus group, and its use as a target in related studies, lay on the fact that this protein contains a polymerase domain that is highly conserved at the amino acid level, and less conserved at the nucleotide level, allowing for deeper phylogenetic inference (Zhang and Suttle, 1994).
The Amazon is one of the most ancient freshwater environments, raising the issue if it could function as a reservoir from which freshwater viruses could disperse to other environments. Despite the magnitude of the Amazonian river system little is known in general about the microbes that inhabit these waters (Benner et al., 1995; Hungria et al., 2005; Rejas et al., 2005; Hewson et al., 2006). Particularly, there is a lack of studies on Phycodnaviruses in tropical freshwater environments in general and in the Amazonian river system in particular. This study constitutes the first report on the composition of Phycodnaviruses of tropical waters in South America from the Amazon Basin, which is the source of 20% of the free-flowing freshwater on Earth, as well as being one of the world’s most important ecosystems in terms of biodiversity.

Materials and methods

Study sites

The two main rivers of the Amazon Basin are the Solimões and the Negro, which, after meeting near the city of Manaus, form the great Amazon River. Despite being geographically close, both rivers are physically, chemically and biologically distinct. The Solimões River is considered as a white-water river, having brownish turbid waters, rich in particulate suspension material and pH ranging from 6.2 to 7.2 (Sioli, 1975a, 1975b). The Negro River, on the contrary, is considered a black-water one. The dark brown color of the water results from the high concentration of humic substances in it, originating from the decomposition of high quantities of organic matter present in the river. Its acidic waters (pH from 3.8 to 4.9) have low concentrations of dissolved salts, reflecting the chemical poverty of the soils the waters pass over (Sioli, 1975a, 1975b). Both the Negro River, and its tributary, the Cuieiras River, share the same black-water-river characteristics.

However, the Cuieiras River flows over an area with almost no human impact.

Sample collection and viral concentration

Freshwater samples were collected from the Solimões (n = 3), Negro (n = 4) and Cuieiras (n = 1) Rivers (Table 1), during the flood of 2007 (July), using a submersible pump. The water was transported to the laboratory in 20-l polypropylene carboys, protected from light and heat and processed within 2–3 h. A volume of 10–100l of water were pressure filtered through 142 mm diameter cellulose fiber paper filter (Millipore, Barueri, Brazil), 142 mm diameter 1.2 μm pore size glass fiber (Millipore) and 142 mm diameter 0.45 μm pore size PVDF (Millipore) membrane filters connected in series. The remaining particulate material was concentrated ~1000-fold by ultrafiltration through a 30-kDa-cutoff tangential flow cartridge (Millipore) (Suttle et al., 1991) followed by ultracentrifugation at 100 000g for 2 h in a swinging bucket rotor (Mehnert and Stiewien, 1993). Pellets were resuspended in supernatant and stored at 4 °C in the dark until the DNA was extracted.

Phycodnaviruses DNA amplification

After treatment with Vertrel XF (DuPont, Barueri, Brazil) to remove lipids and proteins (Queiroz et al., 2001), DNA from each concentrate was extracted using the PowerSoil DNA kit (MoBio, Cotia, Brazil) and stored at −20 °C until use. Phycodnavirus DNA polymerase (pol) gene fragments were amplified in two PCR rounds with the degenerate primers, AVS1/AVS2 and AVS1/POL (Chen and Suttle, 1995), using a Hybaid PCR Express thermocycler (Ashford, UK). In the first round 2 μl of template DNA were added to a mixture of 1 U of Platinum Taq DNA polymerase (Invitrogen, Burlington, Canada), 0.2 mM of each deoxyribonucleotide triphosphate, 10 pmol of AVS1, 60 pmol of AVS2, 1.5 mM MgCl₂,
manufacturer-provided PCR reaction buffer and water, to a final volume of 50 μL. In the negative and positive control reactions, template DNA was substituted for water and *Micromonas pusilla* virus SP-1 (MpV-SP1) DNA, respectively. PCR conditions were 90-s initial denaturation at 95 °C, followed by 40 cycles of 45-s denaturation at 95 °C, 45-s annealing at 54 °C and 45-s extension at 72 °C, and plus a final 10-min extension.

A pool of four PCR reactions from each sample (the total volume of 200 μL) was electrophoresed with 6 x loading buffer on a 1.5% agarose gel, at 90 V for 60 min, in 0.5 x Tris-Borate-EDTA buffer. The ethidium bromide-stained gel was visualized with ultraviolet for <5 s using an AlphaImager 3400 (Alphalnnotec, Toronto, Canada). The ~700 bp bands were excised, and then purified using a MinElute PCR Purification kit (Qiagen, Toronto, Canada) before elution in 50 μl of water.

A nested second-round PCR was performed, to confirm that the correct target was amplified. This was done by adding 2 μL of purified DNA to a mixture of 0.625 U of Platinum Taq DNA polymerase, 0.16 mM of each deoxyribonucleotide triphosphate, 30 pmol of both AVS1 and POL primers, 25 μL 1.5 mM MgCl2 and manufacturer-provided PCR buffer, 0.16 mM of each deoxyribonucleotide triphosphate, 30 pmol of both AVS1 and POL primers, 25 μL 1.5 mM MgCl2 and manufacturer-provided PCR buffer.

Cloning and sequencing
Purified 500 bp pol fragments were cloned into pCR4-TOPO vectors (Invitrogen), *Escherichia coli* strain TOP-10 chemocompetent cells were transformed, and plated on Luria-Bertani/agar medium selective for ampicillin (100 μg mL−1), according to the manufacturer’s protocol. Bacterial colonies containing pol inserts were randomly picked and inoculated in 80 μl of Luria-Bertani with ampicillin (100 μg mL−1) in 96-well plates. After a 12-h incubation period at 37 °C, 40 μl of sterile 60% glycerol were added to each well. Plates were incubated for an additional hour at 37 °C and stored at −80 °C until sent to the Genome Québec Innovation Centre at McGill University for sequencing on a 3730xl DNA Analyzer (Applied Biosystems, Carlsbad, CA, USA).

**Phylogenetics and movement of Phycodnaviruses**

**Phylogenetic analysis.** At first, a global maximum likelihood (ML) tree of pol sequences was constructed from an alignment of 244 inferred amino acids from 638 taxa, including environmental sequences (Supplementary information 1) and representatives from all Phycodnaviridae genera: *Paramecium bursaria* Chlorella virus—PBCV (Chlorovirus), *Chrysochromulina brevisilum* virus—CbV and *Phaeocystis globosa* virus—PgL (Pyrenesiviavirus), *Ectocarpus siliculosus* virus—EsV (Phaeovirus), *Emiliania huxleyi* virus—EhV (Coccolithovirus), *Heterosigma akashiwo* virus—HaV1 (Raphidovirus), *M. pusilla* virus—MpV (Prasinovirus) and *Ostreococcus virus*—OsV and *Bathyctecus virus*—BpV (unclassified) (Supplementary information 2). To characterize and better define deep associations among Phycodnaviruses, a diverse set of taxa were used as outgroups (HaV1, EhV, EsV, PgV, CbV and PBCV) during protein-based phylogenetic inferences.

Protein trees were recovered with PhyML v2.4.4 (Guindon and Gascuel, 2003) using the JTT model of aminoacid substitution, with the following parameters optimized from the data: (i) proportion of invariable sites, (ii) tree topology and (iii) shape parameter (γ) of the γ-distributed substitution rates. Additionally, in a second analysis using DNA sequences, the outgroups were removed to better estimate the associations among closely related groups by reducing saturation at degenerate sites introduced by highly divergent sequences. DNA-based phylogenies for Phycodnaviruses alone were obtained from the analysis of complete data sets using BEAST v.1.4.8 (Drummond and Rambaut, 2007; see details below). Maximum clade credibility trees summarizing the posterior probability of nodes on the phylogeny were obtained from trees sampled at stationary during the Markov Chain Monte Carlo (MCMC) runs.

**Phyldynamic analysis.** Based on the coalescent theory of Kingman (1982), intra-species gene genealogies have been extensively used to infer various demographic parameters for a diverse set of organisms (Page and Holmes, 1998), allowing the time and mode of evolution of exogenous viruses (Zanotto et al., 1996, Grenfell et al., 2004, Pybus and Rambaut, 2009), endogenous viruses (Romano et al., 2007) and even complex metazoa (Campos et al., 2010) to be understood. To infer gene genealogies closely related pol sequences were subdivided into demes. We considered demes as discrete subpopulation of viruses with definable genetic relationships. As not much is known about any Phycodnavirus population structure and nor how comprehensive our samples would be of any of the viral ‘species’ we sampled from, we partitioned our sequences into demes (that is, subpopulations) respecting the proper coding frame with Clustal X (Thompson et al., 1997) and Codon Align (http://www.sinauer.com/hall/2e/).
following three criteria: (i) they had to belong to a monophyletic group (that is, they had to share a common relatively recent ancestral node), (ii) have 17 or more samples, which implies a probability of 0.0074 of adding any new sequence that would alter the time to the most recent common ancestor (that is, adding 136 new sequences per deme, Felsenstein, 1993) and (iii) not show genetic divergence within the monophyletic group >2%. The average pairwise genetic distance among taxa from apical monophyletic groups found in the global tree was calculated from the ML estimates of the number of substitutions per site obtained with the DNADIST program (Felsenstein, 1993), version 3.6, available in the PHYLIP package (Felsenstein, 1993).

Demes were analyzed using a Bayesian MCMC approach implemented in BEAST. Bayesian priors for the mutation rate per site per year for Phycodnaviruses were approximated using the formula: log(rate) = 0.95 log (genome size) + 2.67 (r² = 0.97), obtained by interpolating the genome size ranges for Phycodnaviruses (160–560 kbp) into the per-site mutation rate versus genome size curve for organisms ranging from viroids to higher eukaryotes (Gago et al., 2009, data available from authors upon request). A relaxed (uncorrelated lognormal) molecular clock, with a normal prior distribution of mean = 1.4E–8 s/s/y (substitutions per site per year) and s.d. = 2.34E–7 s/s/y was used, as dates of sampling would be uninformative for this time scale. This rate obtained by interpolation agrees with rates of substitution per site per year for double-stranded DNA viruses (Duffy et al., 2008), such as Herpes virus type I, with a rate of 1.8E–8 s/s/y (Drake and Hwang, 2005) and a genome size within the same order of magnitude of the Phycodnavirus. Therefore, we believe that the prior distribution we used would include realistic values of the rate of change per site per year. The behavior of the Phycodnavirus population size through time was inferred using virus sequences and a prior distribution for the mutation rate. This is because, based on phylogenies of Phycodnaviruses sampled during the MCMC, BEAST infers the population size parameter based on the number of cladogenetic events (that is, nodes or bifurcation events along the tree from the tips back to the root) to estimate the joint parameter (Ne × g) that is, effective population size (Ne) times the generation time in years (g). Nevertheless, from our data or prior knowledge of our system, we do not have ways to disentangle the actual values of effective population size from generation time, but merely report their joint behavior in time, as shown in Bayesian skyline plots.

To be as conservative as possible an unconstrained Bayesian skyline coalescent prior was used under the best-fit model of nucleotide substitution (HKY + Γ + I, Shapiro et al., 2006). To improve the MCMC search operator values were set to a function of the number of taxa in each group. After optimizing the values of the MCMC operators during preliminary runs, up to 10 additional MCMC runs individually, consisting of 20 million generations, were done to obtain convergence of parameter estimates, indicated by values of effective sampling sizes above 200. In all cases the convergence of parameters during the MCMC was inspected with Tracer v.1.4 (Drummond and Rambaut, 2007), with uncertainties depicted as 95% high probability density (HPD) intervals. To provide an independent assessment of the robustness of our analysis of evolutionary dynamics, the time to the most recent common ancestor of each data set was also estimated using the Path-O-Gen software (available at http://tree.bio.ed.ac.uk/software/pathogen/), which uses a linear regression of root-to-tip genetic distances against sampling time, based on the ML phylogenies described above.

Phycodnaviruses movement among aquatic environments. The BayesTraits software (Pagel and Meade, 2004) was used to derive posterior distributions and ML estimates of values for rates of change among traits, which were assumed as posterior probability estimates of both the rates of movement (that is, migration) among the three environments (riverine, lacustrine and marine) and the posterior probability of the state at the root of the trees (that is, origin of the viruses). Analyses were performed using 10000 trees sampled after the MCMC reached stationarity (usually after 20 million states), and by coding each taxa as discrete multistate traits representing their environmental sources (‘river’ = R, ‘lake’ = L and ‘sea’ = S). Several freshwater samples could not be defined as either R or L; therefore the ambiguity character coding RL (that is, either ‘river’ or ‘lake’) was used instead (Pagel and Meade, 2004). Missing traits were coded as hyphens. To improve the convergence of estimates, at least 100 attempts in finding the likelihood for each tree with Bayes-Trait were used.

Results

Sequences obtained from the eight water samples collected in this study yielded 65 and 39 Phycodnavirus pol sequences from the Solimões and Cuiabá Rivers, respectively, providing the first data on Phycodnaviruses from tropical rivers. These sequence data have been submitted to the GenBank database under accession numbers HQ424349–HQ424430. This data set allowed us to do the first comparison between riverine Phycodnaviruses clones from both hemispheres and from temperate and tropical environments.

Molecular characterization of Amazonian Phycodnaviruses

No Phycodnavirus amplicons were obtained from the Negro River samples. The lack of
Phycodnaviruses in the Negro River could be because of naturally occurring conditions because we did observe phages by electron microscopy but not Phycodnaviruses (data not shown). It is likely that these viruses were either missing at the spot or below the level of detection of our method. Nevertheless, DNA was isolated and sequences from other viruses could be amplified from these samples (data not shown). The physical and chemical conditions of these waters were very similar to those found in the Cuieiras River, from where Phycodnavirus amplicons were obtained. We argue that, possibly, the unique conditions of the river were responsible for the lack of viral detection, assuming that there was no unanticipated methodological problem specific to Phycodnaviruses at those particular sampling localities. All 104 Amazonian sequences were aligned to 550 pol sequences from environmental samples, and to sequences from cultured isolates, to create a 638-taxa data set. Phylogenetic analysis with both DNA and inferred amino acid sequences yielded trees that were similar to those inferred in previous studies (Larsen et al., 2008; Short and Short, 2008; Clasen and Suttle, 2009, Figure 1). Our choice of outgroups for the phylogenetic analysis (Figure 1) was justified as these taxa were more divergent than all Phycodnaviruses we sequences or used as references, in line with previous studies (Short and Suttle, 2003; Larsen et al., 2008; Short and Short, 2008; Clasen and Suttle, 2009). Furthermore, to minimize the impact of saturation at the degenerate coding sites at the DNA level and to properly depict the divergence among the more recently diverged ingroup we excluded the divergent outgroups (HaV1, EhV, EsV, PgV, CbV and PBCV). Nevertheless, our phylogenies revealed highly divergent monophyletic groups showing intense cladogenesis near the tips of the tree branches that are consistent with demographic expansions of closely related populations (Figure 2). These monophyletic groups were identified as demes (shown as yellow sectors in the tree) for the purpose of subsequent demographic studies. The branching pattern on the tree also supports a distant relationship between Phycodnavirus clone demes from freshwater and marine lineages (Figure 2).

Population dynamics of Phycodnaviruses

Population size estimates shown in the Bayesian skyline plot for all Phycodnaviruses (Figure 3) were calculated as the effective population times the generation time (Ne × g). There was a significant reduction in Ne × g and the 95% HPD reached a nadir between 500 and 300 thousand years before present (KYBP), suggesting a complex

![Figure 1](image-url) ML tree of Phycodnaviruses based on amino-acid sequences inferred from DNA pol sequences. Nodes marked in orange contain marine sequences, nodes marked in light blue contain sequences from lakes and rivers, nodes marked in green contain sequences from lakes and nodes marked in dark blue contain sequences from rivers.
fluctuating dynamics of Phycodnaviruses over a large time scale. The data also provided evidence of a considerable increase in the number of lineages after a bottleneck around 300 KYBP, followed by a plateau around 100 KYBP (Figure 3). This agrees with the signature observed near the present (Figure 4). Moreover, there were no obvious differences in population sizes of demes from the Cuieiras and Solimões Rivers compared with other localities. Deme 2 (from the Chadfield reservoir and the South Platte River) and Deme 4 (from the Solimões River) did not show the population reduction about 300 000 years ago observed elsewhere. Rather, the Solimões River kept a relatively constant size as indicated by a stable Ne × g. Moreover, we noticed that Deme 2 was the only population to increase throughout that period.

Movement of Phycodnaviruses among water environments

To evaluate the apparent restriction of gene flow among Phycodnaviruses, the likelihood method in BayesTraits was used to estimate the lineage exchange among aquatic environments. Results indicated significant genetic exchange among rivers and lakes, but exchange rates between marine and fresh waters were very low, ranging from 100 to 10 000-fold less (Figure 5). Significant differences among rates were indicated by a lack of overlap among the 95% HPD estimates, rejecting the hypothesis of random exchange (panmixis) of Phycodnaviruses among marine and fresh waters, and providing evidence for restricted gene flow. Because exchange values between rivers and lakes were the highest, but within the same order of magnitude, we evaluated how the ratio varied.

Figure 2  ML tree of Phycodnavirus DNA pol inferred amino acid sequences. Nodes are identified by their origin as marine (orange), riverine (dark blue), lacustrine (green) and fresh water (light blue), with sequences from both rivers and lakes. Demes (yellow) are identified by numbers from 1 to 9. Demes 1 and 2—samples from the Chatfield Reservoir (CO, USA) and the South Platte River (CO, USA); Deme 3—samples from the Chatfield Reservoir, the South Platte River and Lake Ontario (ON, Canada); Deme 4—samples from the Solimões River (AM, Brazil); Deme 5—samples from the South Platte River; Deme 6—samples from the Cuieiras River (AM, Brazil); Demes 7 and 9—samples from Lake Ontario and Deme 8—samples from Lake Ontario and Crawford Lake (ON, Canada).
depending on the tree used for the estimate. Joint estimates of rates obtained with BayesTraits from a sample of trees visited by the MCMC method indicated that the ratio was robust to values of gene flux from river to lake, mostly larger than from lake to river (Figure 6). That happened irrespective of the rates co-estimated for each tree that was summarized by the line with a high correlation coefficient ($r^2 = 0.99$) rising above the diagonal. Consequently, this result supports the idea of further asymmetries in Phycodnavirus gene flow among freshwater systems, possibly with greater influx of viruses from ‘rivers’ into ‘lakes’ than the other way around. An attempt to infer the root position on MCMC trees
showed higher probabilities for a marine origin (S) of Phycodnaviruses, although large overlapping 95% HPD values for the alternative environments (R, L) indicate that the root habitat is unresolved (data not shown).

Discussion

Evolutionary dynamics of Phycodnaviruses

The phylogenetic trees of DNA pol sequences indicated that all samples studied, including those from the Amazon, belong to Group C Phycodnaviruses (Larsen et al., 2008), which are found in a wide range of marine and freshwater environments. The phylodynamics of DNA pol sequences was characterized by a complex signature of fluctuating dynamics. It depicted a remarkable reduction of the effective population size \( \text{Ne} \times g \) around 400 KYBP, suggestive of a genetic diversity bottleneck with small 95% HPD (Figure 3), indicating a reduction in cladogenetic events before the recent burst of lineages on most demes. We argue that sampling biases alone would not explain this effect. Crucially, increased sampling would not only reduce even further the likelihood of finding a new older root for any particular deme (Felsenstein, 2004), but also magnify the recent growth signal. Moreover, a similar near-present expansion was observed independently in all fresh- and salt-water systems. Therefore, assuming that the pruning of lineages causing the \( \text{Ne} \times g \) reduction we observed is not due to sampling, it could be explained, to some extent, by a boom and bust process (Ferguson et al., 1999), which causes older lineages to die out within demes (Zanotto et al., 1996). Notably, fluctuations in viral effective sizes may have an important role in controlling the succession dynamics of algal populations, causing abrupt bloom termination (Bratbak et al., 1993; Jacquet et al., 2002). This is part of the well-known boom and bust dynamics of host algal population (that is, ‘kill the winner’ hypothesis—Thingstad, 2000). In addition, other non-biotic factors could also have had an impact on the overall dynamics we observe.

For example, the gradual reduction of viral effective population size may also be compatible with the known paleontological record. According to our temporal reconstruction of events based on a rate of change of 1.4E-8 s/s/y, the demographic reduction in the number of lineages (Figure 3) appears to have taken place after the Gelasian age, during the Pleistocene, between 1.5 million and 300 thousand years ago. This is a period known for key changes such as the extinction of the calcareous nanofossils Discostaer surculus and D. pentaradiatus and the planktonic foraminifer Globigerinoides extremus (Rio et al., 1998; Gradstein et al., 2005). Interestingly, however, after considerable growth in \( \text{Ne} \times g \) coinciding with the warmer Eemian interglacial period, a decline was observed leading to the present (Figure 3).

Is the environmental movement of Phycodnaviruses dictated by physico-chemistry?

Because we had information related to the place of collection of a diverse set of samples, we did use phylogenetic associations to study the movement of Phycodnaviruses between aquatic environments. Our findings supported the notion that Phycodnaviruses may not be completely mixed (that is, panmictic) with respect to water environments (riversine, lacustrine and marine), and that this gene flow restriction among these environments could have a significant role in their evolution and dispersal. Furthermore, we observed freshwater and marine samples clustering separately, in accordance with previous results obtained from smaller data set (Clasen and Suttle, 2009) in particular and with what has been observed for several other microorganisms in general (Logares et al., 2009). Data on North American Phycodnaviruses did not present good evidence for the deep phylogenetic relationship between marine and freshwater isolates of group I and II as defined by Short and Short (2008). However, our results agree with the finding that freshwater Phycodnaviruses clones are more closely related to some marine isolates (MpV, OsV and BpV) than to freshwater isolate PBCV (Short and Short, 2008). The fact that PBCV radiated deep in the tree, adjacent to highly divergent seawater samples, also hindered a precise placement of the root-tree position, even though the inclusion of sequences from cultivated viruses had placed it at ‘sea’ in the maximum clade credibility tree.

Our multi-state character change analyses with BayesTraits suggested significant restrictions in Phycodnaviruses movement between salt and freshwater, which could possibly explain further restrictions on the movement of these viruses among aquatic environments, or the colonization of specific freshwater systems like the Negro River. On a similar vein, a drastic reduction in freshwater viral counts upon addition of seawater was observed by Cissoko et al. (2008), suggestive of salinity restriction for virus movement between freshwater and marine environments. However, the same effect was not observed upon the addition of freshwater to marine samples, indicating that marine viruses can better tolerate reductions in salinity, as do marine bacteria. Similarly, marine phytoplankton are known to cope well with changes in salinity, probably because of their necessity to deal with tidal variation, precipitation and other environmental phenomena (Kirst, 1989), whereas freshwater phytoplankton are usually totally replaced by marine species in a range of 0.5–10% salinity (Lionard et al., 2005). Moreover, bacteriophages isolated from lake water and soil were able to replicate in marine microbes (Sano et al., 2004; Logares et al., 2009), which, given that bacteria tolerate greater salinity variations than algae, suggests that the restriction of Phycodnaviruses movement between fresh and salt water environments...
References

Benner R, Opsahl S, Chín-Leo G, Richey JE, Forsberg BR. (1995). Bacterial carbon metabolism in the Amazon river system. Limnol Oceanog 40: 1262–1270.

Brathbuk G, Egge JK, Heldal M. (1993). Viral mortality of the marine alga Emiliania huxleyi (Haptophyceae) and termination of algal blooms. Mar Ecol Prog Ser 93: 39–48.

Brussaard CDP. (2004). Viral control of phytoplankton populations—a review. I Eukaryot Microbiol 51: 125–138.

Campos PF, Willerslev E, Sher A, Orlando L, Axelson E, Tikhonov A et al. (2010). Ancient DNA analyses exclude humans as the driving force behind late Pleistocene musk ox (Ovibos moschatus) population dynamics. PNAS 107: 5675–5680.

Castberg T, Larsen A, Sandaa RA, Brussaard CDP, Egge JK, Heldal M et al. (2001). Microbial population dynamics and diversity during a bloom of the marine coccolithophorid Emiliania huxleyi (Haptophyta). Mar Ecol Prog Ser 221: 39–46.

Chen F, Suttle CA. (1995). Nested PCR with 3 highly degenerate primers for amplification and identification of DNA from related organisms. Biotechniques 18: 609–612.

Chen F, Suttle CA. (1996). Evolutionary relationships among large double-stranded DNA viruses that infect microalgae and other organisms as inferred from DNA polymerase genes. Virology 219: 170–178.

Cissoko MM, Desmues A, Bouvy M, Sime-ngando TL, Verling E, Bettarel Y. (2008). Effects of freshwater and seawater mixing on viro and bacterioplankton in a tropical estuary. Freshwater Biol 53: 1154–1162.

Clasen JL, Suttle CA. (2009). Identification of freshwater Phycodnaviridae and their potential phytoplankton hosts, using DNA pol sequence fragments and a genetic-distance analysis. Appl Environ Microbiol 75: 991–997.

Cottrell MT, Suttle CA. (1991). Wide-spread occurrence and clonal variation in viruses which cause lysis of a cosmopolitan, eukaryotic marine phytoplankter, Micromonas pusilla. Mar Ecol Prog Ser 78: 1–9.

Donnan KL, Brasseur G, Chidhasiong A, Giais P, Cox PM, Dickinson RE et al. (2007). Couplings between changes in the climate system and biogeochemistry. In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Avert KB, Tignor M, Miller HL (eds). Climate Change 2007—The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press: Cambridge, UK; New York, NY, pp 499–587.

Drake JW, Hwang CBC. (2005). On the mutation rate of Herpes Simplex Virus Type 1. Genetics 170: 969–970.

Drummond AJ, Rambaut A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol 7: 214.

Duffy S, Shackelton LA, Holmes EC. (2008). Rates of evolutionary change in viruses: patterns and determinants. Nature Rev Genet 9: 267–276.

Dunigan DD, Fitzgerald LA, Van Etten JL. (2006). Phycodnaviruses: a peek at genetic diversity. Virus Res 117: 119–132.

Felsenstein J. (1993). PHYLIP (Phylogeny Inference Package) version 3.5c. Distributed by the author. Department of Genetics, University of Washington: Seattle.

Acknowledgements

We thank Alex Krusche (CENA, USP) for the logistical support, Telma A Monezi (ICB-USP) for technical support, Jonismar Souza, Alice Plaskievetz and Anthony Lopes (INPA), and Patricia Garrafa and Bernardo Gimenes (ICB—USP) for the valuable help during the sample collection and processing. We also thank Jérôme Payet, Caroline Chênard, Matthias Fischer, Jessica Labonté and Amy Chan (EOS-UBC) for their priceless assistance with specific protocols and techniques involved in the project. This project was financed by FAPESP (Grant no. 05/55164-1; scholarship no. 04/11287-0), supported by CAPES and in part by a Discovery Grant from the Natural Engineering and Research Council of Canada to CAS. PMAZ holds a CNPq PQ scholarship. We thank two anonymous referees for their meticulous reading and relevant contributions to the manuscript.
Felsenstein J (ed.) (2004). *Inferring Phylogenies*. Sinauer: Sunderland.

Ferguson N, Anderson R, Gupta S. (1999). The effect of antibody-dependent enhancement on the transmission dynamics and persistence of multiple-strain pathogens. *Proc Natl Acad Sci* 96: 790–794.

Gago S, Elena SF, Flores R, Sanjuán R. (2009). Extremely high mutation rate of a hammerhead viroid. *Science* 323: 1308.

Gobler CJ, Hutchins DA, Fisher NS, Cosper EM, Sanudo-Wilhelmy AS. (1997). Release and bioavailability of C, N, P, Se and Fe following viral lysis of a marine chrysophyte. *Limnol Oceanogr* 42: 1492–1504.

Gradstein FM, Ogg JG, Smith AG (eds) (2005).

Ferguson N, Anderson R, Gupta S. (1999). The effect of antibody-dependent enhancement on the transmission dynamics and persistence of multiple-strain pathogens. *Proc Natl Acad Sci* 96: 790–794.

Gago S, Elena SF, Flores R, Sanjuán R. (2009). Extremely high mutation rate of a hammerhead viroid. *Science* 323: 1308.

Gobler CJ, Hutchins DA, Fisher NS, Cosper EM, Sanudo-Wilhelmy AS. (1997). Release and bioavailability of C, N, P, Se and Fe following viral lysis of a marine chrysophyte. *Limnol Oceanogr* 42: 1492–1504.

Gradstein FM, Ogg JG, Smith AG (eds) (2005).

Ferguson N, Anderson R, Gupta S. (1999). The effect of antibody-dependent enhancement on the transmission dynamics and persistence of multiple-strain pathogens. *Proc Natl Acad Sci* 96: 790–794.

Gago S, Elena SF, Flores R, Sanjuán R. (2009). Extremely high mutation rate of a hammerhead viroid. *Science* 323: 1308.

Gobler CJ, Hutchins DA, Fisher NS, Cosper EM, Sanudo-Wilhelmy AS. (1997). Release and bioavailability of C, N, P, Se and Fe following viral lysis of a marine chrysophyte. *Limnol Oceanogr* 42: 1492–1504.

Gradstein FM, Ogg JG, Smith AG (eds) (2005).

Ferguson N, Anderson R, Gupta S. (1999). The effect of antibody-dependent enhancement on the transmission dynamics and persistence of multiple-strain pathogens. *Proc Natl Acad Sci* 96: 790–794.

Gago S, Elena SF, Flores R, Sanjuán R. (2009). Extremely high mutation rate of a hammerhead viroid. *Science* 323: 1308.

Gobler CJ, Hutchins DA, Fisher NS, Cosper EM, Sanudo-Wilhelmy AS. (1997). Release and bioavailability of C, N, P, Se and Fe following viral lysis of a marine chrysophyte. *Limnol Oceanogr* 42: 1492–1504.

Gradstein FM, Ogg JG, Smith AG (eds) (2005).

Ferguson N, Anderson R, Gupta S. (1999). The effect of antibody-dependent enhancement on the transmission dynamics and persistence of multiple-strain pathogens. *Proc Natl Acad Sci* 96: 790–794.
are pathogens of marine phytoplankton. *Appl Environ Microbiol* **57**: 721–726.

Thingstad TF. (2000). Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. *Limnol Oceanogr* **45**: 1320–1328.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **25**: 4876–4882.

Van Etten JL, Ghabrial SA. (1991). *Phycodnaviridae*. In: Francki RIB, Fauguet CM, Knudson DL, Brown F (eds). Classification and Nomenclature of Viruses, fifth report. Springer-Verlag: Vienna, pp 137–139.

Van Etten JL, Graves MV, Müller DG, Boland W, Delaroque N. (2002). *Phycodnaviridae*—large DNA viruses. Springer-Verlag: Vienna, pp 1479–1516.

Wilhelm SW, Suttle CA. (1999). Viruses and nutrient cycles in the sea. *Bioscience* **49**: 781–788.

Wilson WH, Schroeder DC, Ho J, Canty M. (2006). Phylogenetic analysis of PgV-102P, a new virus from the English Channel that infects *Phaeocystis globosa*. *J Mar Biol Ass UK* **86**: 485–490.

Wilson WH, Van Etten JL, Allen MJ. (2008). The Phycodnaviridae: the story of how tiny giants rule the world. In: Van Etten J (ed.) *Lesser Known Large dsDNA Viruses*. Springer: Berlin, pp 1–42.

Zanotto PM, Gould EA, Gao GF, Harvey PH, Holmes EC. (1996). Population dynamics revealed by flaviviruses by molecular phylogenies. *Proc Natl Acad Sci* **93**: 548–553.

Zhang YP, Suttle CA. (1994). Design and use of PCR primers for B-family DNA polymerase genes to detect and identify viruses and microbes. In: *ASLO & PSA Joint Meeting*. American Society for Limnology and Oceanography: Lawrence, Kans, p a–85.

Supplementary Information accompanies the paper on The ISME Journal website (http://www.nature.com/ismej)