Virology 466-467 (2014) 146–157

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/yviro

Prasinovirus distribution in the Northwest Mediterranean Sea is affected by the environment and particularly by phosphate availability

Camille Clerissi [a,b], Nigel Grimsley [a,b,*], Lucie Subirana [a,b], Eric Maria [c], Louise Oriol [d,e], Hiroyuki Ogata [f], Hervé Moreau [a,b], Yves Desdevises [a,b]

[a] Sorbonne Universités, UPMC Univ Paris 06, UMR 7232, Biologie Intégrative des Organismes Marin, Observatoire Océanologique, Avenue du Fontaulé, F-66650 Banyuls-sur-Mer, France
[b] Sorbonne Universités, CNRS, UMR 7232, Observatoire Océanologique, Biologie Intégrative des Organismes Marins, Avenue du Fontaulé, F-66650 Banyuls-sur-Mer, France
[c] Sorbonne Universités, UPMC Univ Paris 06, UMS 2348, Observatoire Océanologique, Avenue du Fontaulé, F-66650 Banyuls-sur-Mer, France
[d] Sorbonne Universités, UPMC Univ Paris 06, UMR 7621, Laboratoire d'Océanographie Microbienne, Observatoire Océanologique, Avenue du Fontaulé, F-66650 Banyuls-sur-Mer, France
[e] Sorbonne Universités, CNRS, UMR 7621, Observatoire Océanologique, Laboratoire d'Océanographie Microbienne, Avenue du Fontaulé, F-66650 Banyuls-sur-Mer, France
[f] Institute for Chemical Research, Kyoto University, Uji, Kyoto 611-0011, Japan

ARTICLE INFO

Article history:
Received 11 April 2014
Returned to author for revisions
13 May 2014
Accepted 8 July 2014
Available online 8 August 2014

Keywords:
DNA polymerase
Virus
Diversity
Distribution
Amplicon
454-pyrosequencing

ABSTRACT

Numerous seawater lagoons punctuate the southern coastline of France. Exchanges of seawater between these lagoons and the open sea are limited by narrow channels connecting them. Lagoon salinities vary according to evaporation and to the volume of freshwater arriving from influent streams, whose nutrients also promote the growth of algae. We compared *Prasinovirus* communities, whose replication is supported by microscopic green algae, in four lagoons and at a coastal sampling site. Using high-throughput sequencing of DNA from a giant virus-specific marker gene, we show that the environmental conditions significantly affect the types of detectable viruses across samples. In spatial comparisons between 5 different sampling sites, higher levels of phosphates, nitrates, nitrites, ammonium and silicates tend to increase viral community richness independently of geographical distances between the sampling sites. Finally, comparisons of *Prasinovirus* communities at 2 sampling sites over a period of 10 months highlighted seasonal effects and the preponderant nature of phosphate concentrations in constraining viral distribution.

Introduction

Aquatic viruses have a strong influence on microbial community composition, since they can control the blooms of their hosts (Bratbak et al., 1993,1996; Jacquet et al., 2002; O'Kelly et al., 2003; Schroeder et al., 2003) and enhance their diversification through an antagonistic interaction (Van Valen, 1973; Paterson et al., 2010; Thomas et al., 2011). Moreover, they are the most abundant biological entities on Earth (Suttle, 2005), but little is known about how viral communities change in both spatial and temporal scales.

Generally bacteriophage abundance is high in eutrophic systems compared to oligotrophic systems (Boehme et al., 1993; Weinbauer et al., 1993; Wommack and Colwell, 2000). However, other studies did not find any direct link between viral abundance and trophic state, but rather with bacterial activity and abundance (Corinaldesi et al., 2003), and also with sampling depth of the water column (Taylor et al., 2003). Although changes in viral community composition are less well documented, available evidence suggests that (i) viral communities are composed of the same strains, but are structured according to environmental conditions (Short and Suttle, 2002; Breitbart and Rohwer, 2005; Bellec et al., 2010a), and (ii) that rapid variations can occur within a month (Schroeder et al., 2003; Snyder et al., 2007). Furthermore, few studies have focused on natural viral community compositions and the role of environmental factors, among which nutrients such as phosphates can influence virus-host interactions (Bratbak et al., 1993; Wilson et al., 1996).

* Corresponding author at: UPMC, CNRS UMR 7232, Observatoire Océanologique, Avenue du Fontaulé, 66650 Banyuls-sur-Mer, France. Tel.: +334 68887396, fax: +334 68887398.
E-mail address: nigel.grimsley@obs-banyuls.fr (N. Grimsley).

http://dx.doi.org/10.1016/j.virol.2014.07.016
0042-6822/

Some of the most studied eukaryote-infecting marine viruses are those in the genus *Prasinovirus* (Bellec et al., 2009), a member of the *Phycodnaviridae* family (Van Etten et al., 2002; Wilson et al., 2009), that infect green photosynthetic picoeukaryotes in the class Mamiellophyceae (Marin and Melkonian, 2010). In the marine environment, this class is dominated by three genera (*Bathycoccus*, *Micromonas* and *Ostreococcus*), which comprise different species: for instance, three species have been described so far in *Ostreococcus*, *Ostreococcus tauri* (Courties et al., 1994), *O. lucimarinus* (Palenik et al., 2007), and *O. mediterraneus* (Subirana et al., 2013). Three groups of *Prasinovirus* are well described and consist of *Bathycoccus* viruses (BpV), *Micromonas* viruses (MpV), and *Ostreococcus* viruses (OlV, OmV and OtV for *Ostreococcus lucimarinus* viruses, *Ostreococcus mediterraneus* viruses, and *Ostreococcus tauri* viruses, respectively) (Bellec et al., 2009; Moreau et al., 2010). They were collected worldwide (Weynberg et al., 2009; Bellec et al., 2010b; Manrique et al., 2012), and were found to be more abundant in coastal sites (Bellec et al., 2010a). Bellec et al. (2010a) tried to disentangle the influence of environmental factors over a two-year period in the Mediterranean coastal station SOLA and found some links between OtV abundance and both concentration of chlorophyll *a* and phosphates. However, these authors only considered viruses infecting a particular host strain, maintained in culture as the clonal cell line RCC745 (*O. tauri*, Courties et al., 1994), and did not consider viral genetic distances. Here we investigated both spatial and temporal variations of the whole communities using a *Prasinovirus*-specific primer set for the DNA polymerase (PolB) (Clerissi et al., 2014) and a 454-pyrosequencing approach. Five coastal locations in the Mediterranean Sea were sampled the same week in February 2011 for spatial comparisons, and two sites were sampled during 10 months from February to November 2011 to permit a temporal analysis. We measured several environmental variables at the different sites and we attempted to disentangle their influence on *Prasinovirus* communities. Our results suggest that although the five sites are geographically close, the environment selects *Prasinovirus* genotypes, and particularly highlight a correlation with phosphate availability.

Results

Sampling sites

The five sampled stations represent four lagoons and one coastal site, SOLA (Fig. 1). Because the maximum depth at SOLA was higher relative to lagoon stations (Table S1), we considered this as an additional variable for the different PCA (Principal Components Analysis). The first and second components (hereafter named C1 and C2) of PCA of the 19 samples (Fig. 2) explained ~83% of the total variation, and divides them according to the salinity and their nutrient content (Table 1), with lagoons being less salty and richer in nutrients than SOLA. Moreover, the addition of the depth variable shows its positive link with salinity (Fig. 2). In the PCA of the five locations sampled in February 2011 (Fig. S2), C1 explained 66.38% of the total variation, and is mostly driven by salinity (which also seems to be depth-dependent) and silicates (Table S2). Finally, PCA on temporal analyses at SOLA and Leucate were very similar to the global analysis and suggest that the variation is mainly explained by nitrates (C1) and phosphates (C2) for Leucate (Fig. S3, Table S2), and by nutrients (C1) and salinity (C2) for SOLA (Fig. S4, Table S2). For both locations, winter and summer samples contain higher values for nutrients and salinity, respectively.

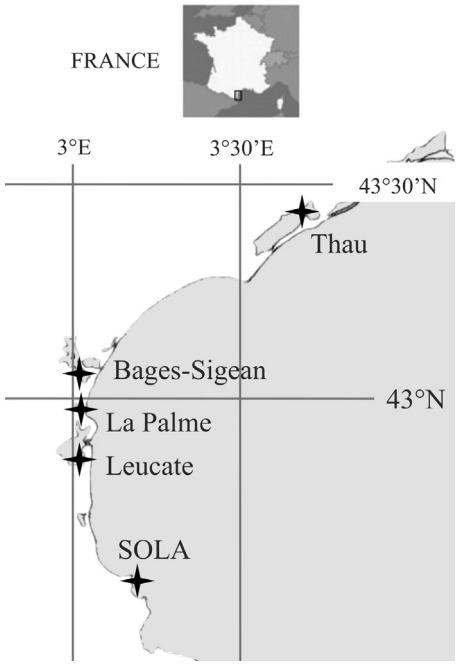

Fig. 1. Locations of sampling sites. Four locations are lagoons (Leucate: 42°48′24″N, 03°01′27″E; La Palme: 42°57′18.04″N, 3° 0′3.56″E; Bages-Sigean: 43°03′14.76″N, 2°59′51.63″E and Thau: 43°26′01.40″N, 03°39′54.49″E) and one is a coastal station (SOLA: 42°29′18″N, 3°8′42″E) included in the French marine monitoring network SOMLIT.

Links between Prasinovirus communities, geography and environments

There were 5 and 16 *Prasinovirus* samples for the spatial and the temporal analyses, respectively. The temporal analysis is divided in 10 samples for Leucate and 6 for SOLA, due to the small amount of PCR products for the coastal station samples, particularly from spring and autumn (samples from March, April, August and October are missing). Since PCR and sequencing steps may bias the abundance of environmental sequences, we used dereplicated sequence data to estimate genotype richness. Furthermore, in order to standardize the whole dataset for subsequent statistical analyses, we defined a OTU (Operational Taxonomic Unit) cutoff of 90% for *Prasinovirus* using rarefaction curves with different thresholds (Fig. S1). This percentage was chosen according to an appropriate asymptotic curve, suggesting that each sample was representative of the population at this OTU cutoff. An average of 84 OTU per sample and a total of 486 different OTU were found for the 19 samples (Table 2).

Prasinovirus sequence annotation suggested that most genotypes had no cultured representatives (54.66%), and that *Micromonas* viruses were the most abundant (33.72%), followed by *Ostreococcus* viruses (8.22%), *Bathycoccus* viruses (3.37%), and *Chlorovirus* (0.02%). The phylogenetic tree highlights a high degree of diversity (Fig. 3). One clade contains the reference *Prasinovirus* genotypes together with many environmental sequences, and many unknown sequences have an intermediate position between known chloroviruses and *Prasinovirus*.

The distribution of *Prasinovirus* in the 19 samples was first described using a clustering approach (Fig. 4). This analysis based on the composition of *Prasinovirus* communities tends to cluster lagoons according to the sampling time, and shows that viral communities in SOLA exhibit different structures (proportions of the different OTU for a cutoff of 90%, Bray-Curtis dissimilarity) and content (only presence/absence of the different OTU, Jaccard index) compared to lagoons, particularly to Leucate, during

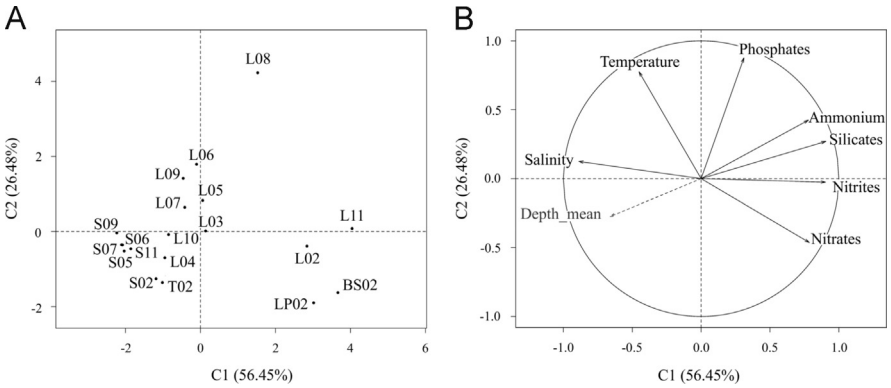

Fig. 2. Principal component analysis of the 19 sampling stations according to the environmental factors. A. Distances between samples. B. Correlations between variables. Stations are named from L02 to L11 (Leucate-February to Leucate-November); from S02 to S11 (SOLA-February to SOLA-November); T02 (Thau-February); LP02 (La Palme-February); BS02 (Bages-Sigean-February). Salinity (g/L); Temperature (°C); Nitrates (μM); Nitrites (μM); Ammonium (μM); Phosphates (μM); Silicates (μM); Depth_mean: mean depth of the bottom (m). Depth_mean was used as an additional variable, and did not contribute to build the different components.

Table 1
Correlations between environmental variables and first and second components of the PCA. Only the significant links are shown.

Variable	Component	Correlation	*p*-value
Silicates	1	0.906	9.591e−08
Nitrites	1	0.903	1.194e−07
Nitrates	1	0.786	6.575e−05
Ammonium	1	0.779	8.411e−05
Depth_mean	1	−0.660	2.084e−03
Salinity	1	−0.889	3.689e−07
Phosphates	2	0.877	8.502e−07
Temperature	2	0.775	9.650e−05
Nitrates	2	−0.463	4.575e−02

Table 2
Prasinovirus sequencing data.

Station abbreviation[a]	PolB sequences (raw data)	*Prasinovirus* genotypes	*Prasinovirus* OTU (90%)	*Prasinovirus* richness (90%)
S02	31,566	147	46	55.4
L02	40,676	275	106	131.3
LP02	27,772	419	118	140.8
BS02	30,467	358	106	130.7
T02	46,977	281	100	122.8
L03	14,619	351	97	116.6
L04	30,900	226	94	116.1
L05	12,354	403	96	113.5
L06	23,041	216	44	52.2
L07	22,753	110	56	70.2
L08	8050	177	53	65.2
L09	34,132	271	98	122.3
L10	14,956	273	70	81.1
L11	23,758	342	90	106.1
S05	19,698	273	82	95.8
S06	10,212	145	46	55.7
S07	25,205	430	115	137.3
S09	11,954	235	96	117.7
S11	31,799	243	77	93.4
Mean	24,257	272	83.7	101.3

[a] Station abbreviations correspond to station names (S for SOLA; L for Leucate; LP for La Palme; BS for Bages-Sigean and T for Thau) and sampling month (from 02 for February to 11 for November).

one year. In particular, SOLA samples were more similar to February-sampled lagoons than to other samples from Leucate with regard to their *Prasinovirus* OTU composition. Next, we tried to find out whether *Prasinovirus* community compositions (based on the proportion of OTU for a cutoff of 90%, Bray–Curtis dissimilarity) were linked to geography and/or to the environmental variables for both the spatial and the temporal analyses, and for the whole dataset (i.e. the 19 samples). On one hand, the whole dataset and the spatial analysis revealed that *Prasinovirus* community compositions were correlated to environmental conditions (Mantel test $r=0.21$, p-value$=0.045$ and $r=0.54$, p-value$=0.031$, respectively), but not to geography ($r=0.01$, p-value$=0.484$ and $r=-0.20$, p-value$=0.595$, respectively). On the other hand, the temporal analysis showed a correlation between *Prasinovirus* community and geography ($r=0.38$, p-value$=0.001$), but not with the environment ($r=0.11$, p-value$=0.206$), neither for Leucate ($r=0.04$, p-value$=0.392$) nor SOLA ($r=0.30$, p-value$=0.099$). Nevertheless, after selection of the variables that best explain the variation of *Prasinovirus* using a correspondence analysis and environmental vector fitting, significant correlations were found between *Prasinovirus* assemblages and environment, (i) when only phosphates, temperature and salinity were retained for the whole temporal analysis dataset ($r=0.31$, p-value$=0.004$), (ii) when only phosphates and temperature were retained for Leucate ($r=0.33$, p-value$=0.029$), (iii) when only phosphates, temperature and salinity were retained for SOLA ($r=0.55$, p-value$=0.024$). Furthermore, an indirect link with the environment was found for the temporal analysis of Leucate, since the month of sampling explained a high variation of *Prasinovirus* communities ($r=0.43$, p-value$=0.007$), but not for SOLA ($r=0.36$, p-value$=0.099$). To summarize, *Prasinovirus* communities are correlated mainly to the environmental conditions of the 19 samples, and the significant link obtained with the geography for the temporal analysis is related to environmental differences between the two contrasting sampling sites, particularly for phosphates, temperature and salinity.

Prasinovirus richness and the variables explaining their distribution

An approximated unimodal distribution of OTU richness was found in SOLA during a year with a maximum value in July (Fig. 5). In contrast, OTU richness in Leucate was lower in summer. *Prasinovirus* communities in lagoons seem to be more diverse than in SOLA for the month of February, since we found an average richness of 131.4 in lagoons and only 55.4 in the coastal station. However, *Prasinovirus* richness is not constant along the year in the coastal station and can be higher than lagoons' value (Fig. 5, Table 2). The CCA coupled with a forward-selection procedure highlighted that phosphate concentration is the only variable that explains the distribution of *Prasinovirus* for the whole dataset (p-value$=0.005$), and separately for both spatial (p-value$=0.008$; Fig. 6) and the temporal analyses (p-value$=0.026$; Fig. 7). Although nitrates showed a similar trend to phosphates for the

Fig. 3. Phylogenetic tree of environmental OTU and 23 reference sequences of *Prasinovirus* and *Chlorovirus*, reconstructed using Bayesian inference. OTU are defined for a nucleotide identity of 75%. The tree was rooted using the chloroviruses. Numbers are posterior probabilities (%) reflecting clade support. The 23 reference sequences are indicated in bold.

spatial analysis (Fig. 6), this variable did not explain a significant level of the residual variation in *Prasinovirus* composition. For the temporal analysis, phosphates were selected when only Leucate samples (p-value = 0.001) were considered (Fig. 7), but only the salinity was significant in SOLA (p-value = 0.010; Fig. 8), and this variable showed a trend similar to the temperature (Fig. 8). Finally, phosphate availability in Leucate is much more variable than in SOLA (Fig. 9).

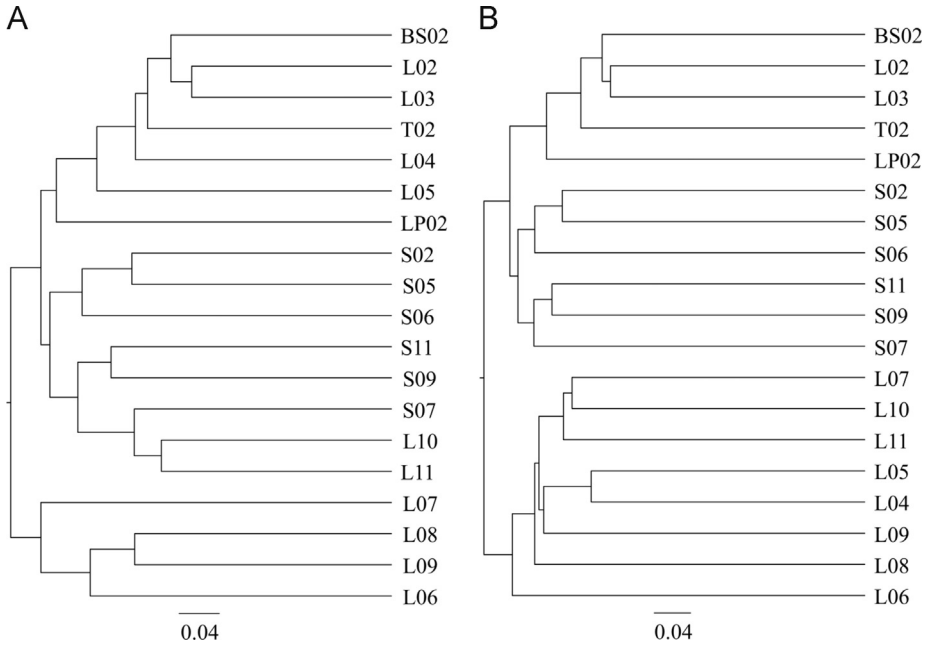

Fig. 4. Clustering of the 19 samples based on *Prasinovirus* assemblages (UPGMA). A. Bray–Curtis dissimilarities. B. Jaccard index. OTU are defined for a nucleotide identity of 90%.

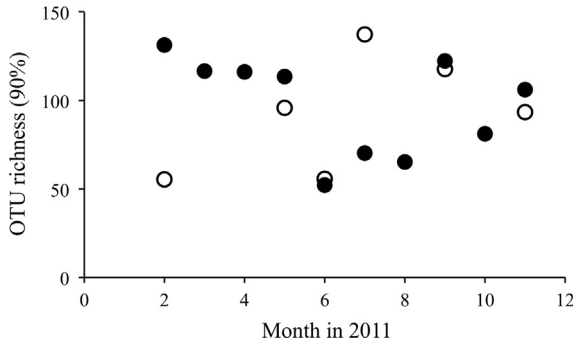

Fig. 5. *Prasinovirus* OTU richness for the temporal analysis. •: Leucate; ○: SOLA. OTU are defined for a nucleotide identity of 90% and richness was estimated using the bootstrap calculator.

Discussion

Our results suggest that although *Prasinovirus* were not abundant every month in SOLA (in accordance with the difficulty for obtaining enough PCR product and our previous observations (Bellec et al., 2010a)), the richness peak observed in July was higher than for some lagoon samples. Moreover, while SOLA showed a high environmental inertia, variations in Leucate were of greater magnitude. For example, changes in phosphate concentrations were up to 3 times higher in Leucate than in SOLA during 10 months. Finally, we found that spatial and temporal distributions of *Prasinovirus* were directly correlated with phosphates, except in SOLA where the overall seasonality pattern could explain differences between communities.

Sampling sites

Prasinovirus of the five locations are well described, particularly those inhabiting lagoons, where they are very abundant (Bellec et al., 2009; Bellec et al., 2010a; Bellec et al., 2010b; Thomas et al., 2011; Clerissi et al., 2012, 2013). The geomorphology and especially the depth of the water column are the most important differences between the five stations (Figs. 2 and S1). Indeed, Mediterranean lagoons are shallow rather confined basins with some influent freshwater tributaries and that are connected with the sea by a few narrow channels (Arnaud and Raimbault, 1969; Llas et al., 1990; Ifremer, 2008; Derolez et al., 2012). As a consequence, their salinities were lower, and their nutrient contents were higher than in SOLA (Fig. 2). In addition, the Leucate lagoon showed high variations of physico-chemical variables along the year, since water flows in from the land, in and out from the sea, and also the evaporation of water from the lagoon changes according to seasons. This seasonal pattern was well described in this analysis, since (i) in summer, the salinity was higher than in winter for SOLA and Leucate samples (these differences were in part due to evaporation, since salinity showed a similar trend to temperature), (ii) in winter, the nutrients were mostly more concentrated, probably because of increased influx from agricultural runoff from freshwater inlets (precipitation varies from an average of 38 mm April–September to 57 mm October–March over the period 1981–2010 [source: Météo-France]), and because their uptake by algae was less important. Indeed, most of the primary production occurs in spring with the phytoplankton bloom in the Mediterranean Sea (Vidussi et al., 2000; Bec et al., 2005).

Links between Prasinovirus communities, geography and environment

Prasinovirus communities were described using new molecular markers (Clerissi et al., 2014), but such approaches might produce biased estimates of the viral diversity. In this study, *Prasinovirus* communities were dominated by unknown genotypes, and also by *Micromonas* viruses. While *Prasinovirus* representatives were clustered into a clade with environmental sequences, other unknown genotypes form multiple groups of taxa, suggesting that they infect a diverse set of hosts. However, the cophylogenetic pattern inferred between *Prasinovirus* and their hosts (Bellec et al., 2014) (i.e. related viruses tend to infect related hosts), suggests that most of the unknown host species also belong to the class Mamiellophyceae. In agreement with this hypothesis, Viprey et al. (2008), using degenerate primers to amplify a chlorophyte plastid marker

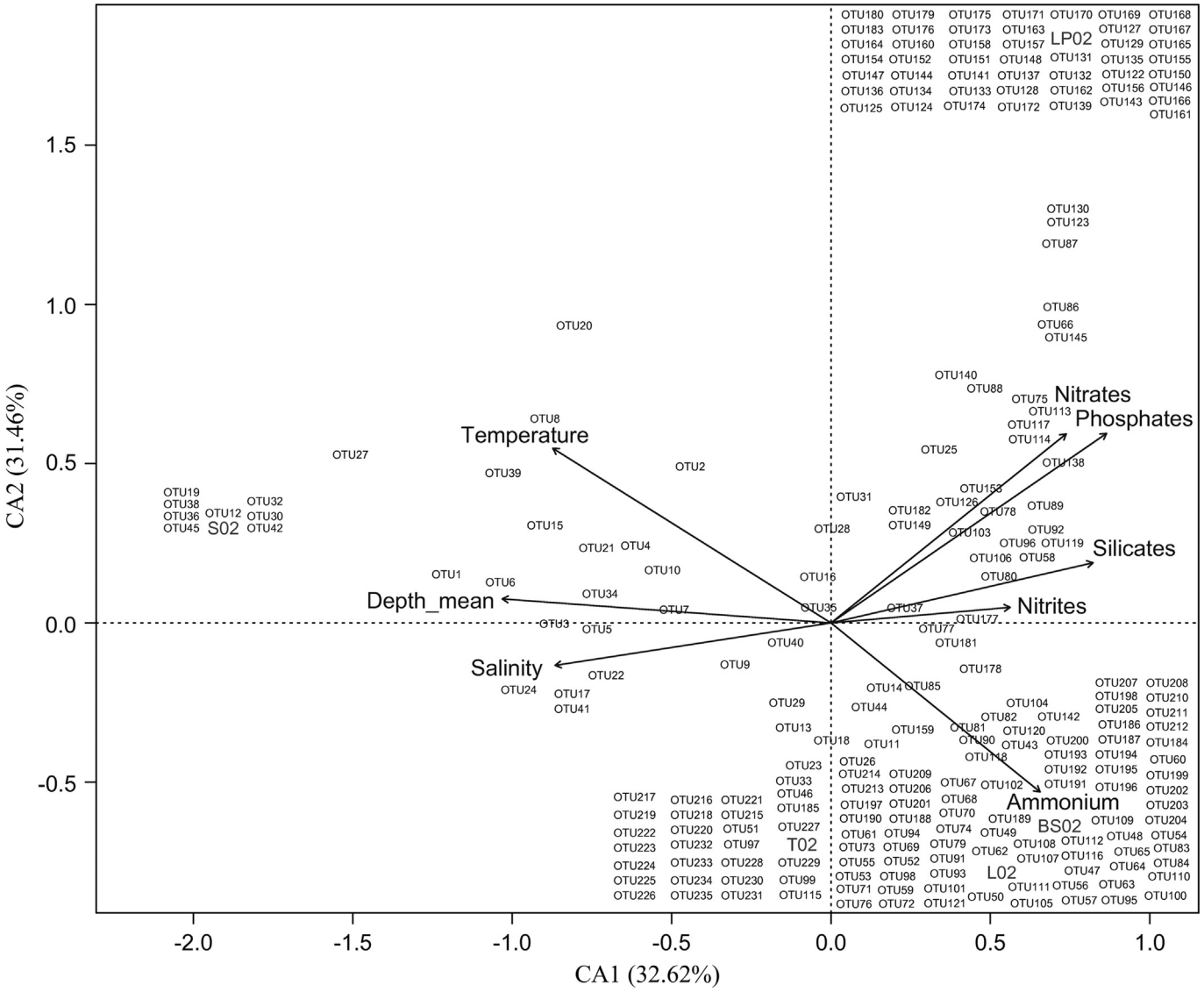

Fig. 6. Correspondence analysis of the five locations on *Prasinovirus* assemblages and environmental vector fitting. S02, L02, LP02, BS02, and T02 correspond to SOLA-February, Leucate-February, La Palme-February, Bages-Sigean-February, and Thau-February, respectively. OTU are defined for a nucleotide identity of 90%. Salinity (g/L); Temperature (°C); Nitrates (μM); Nitrites (μM); Ammonium (μM); Phosphates (μM); Silicates (μM); Depth_mean: mean depth of the bottom (m). Only the phosphates were retained using the forward-selection implemented in the CCA analysis.

gene, showed that among 3890 environmental sequences from the Mediterranean Sea, the majority (99%) were prasinophytes. *Micromonas* and *Ostreococcus* were present mainly from mesotrophic regions, whereas *Mamiella*, *Crustomastix* and *Dolichomastix* were more common in oligotrophic surface waters. Overall, the present study shows the limitation of culture-dependent approaches, arising because in practice relatively few strains can be screened in laboratory cultures (Derelle et al., 2008; Bellec et al., 2010a; Clerissi et al., 2012).

Using Mantel tests, we found that *Prasinovirus* communities among the 19 samples tested were affected much more by their environment than by the geographical distances between sampling sites. Indeed, although no link was found between *Prasinovirus* communities and geography for the five sites, a significant correlation was obtained with the environmental conditions for the whole dataset and the spatial analysis. The lack of correlation between viral compositions and geographic distances for these sampled sites is in concordance with previous observations (Clerissi et al., 2012). These results suggest that the five locations are connected and that the environment might select viral assemblages, consequently highlighting ecological differences between these stations. Moreover, our temporal analysis of Leucate and SOLA samples also suggests that *Prasinovirus* assemblages are constrained by the environment, and particularly by seasons. Indeed, we show that viral communities are significantly linked to sampling times for Leucate (see the clustering of samples in Fig. 7), and a significant correlation was obtained when the environmental factors were reduced to phosphates and temperature. In contrast, no link was found for the sampling time in SOLA, but a significant correlation came out when phosphates, temperature and salinity were considered as environmental variables. Because the salinity was positively related to temperature and in summer almost no freshwater arrives from the Baillaury river at Banyuls-sur-Mer (the main influent for the bay containing SOLA), the increase in salinity could be related to increased evaporation and/or lack of freshwater input. Hence, our data strongly suggest that viral assemblages for both Leucate and SOLA are linked to seasonal constraints. Accordingly, Lami et al. (2009) also suggested that bacterial communities and activities were linked to seasons in SOLA.

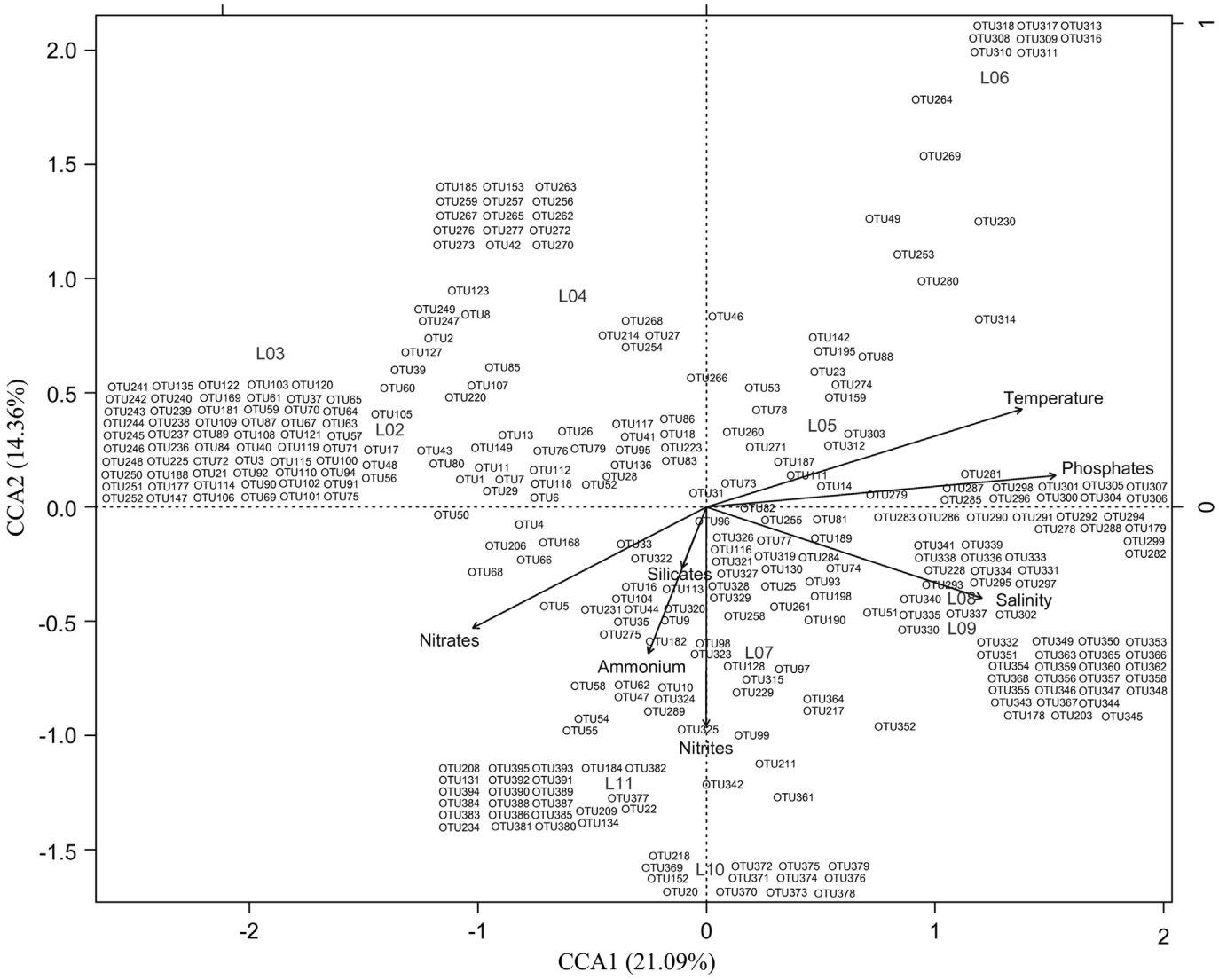

Fig. 7. Canonical correspondence analysis of the Leucate samples on *Prasinovirus* assemblages constrained by environmental data. Stations are named from L02 to L11 for Leucate-February to Leucate-November. OTU are defined for a nucleotide identity of 90%. Salinity (g/L); Temperature (°C); Nitrates (μM); Nitrites (μM); Ammonium (μM); Phosphates (μM); Silicates (μM). Only the phosphates were retained using the forward selection implemented in the CCA analysis.

Prasinovirus richness and variables explaining their distribution

The highest *Prasinovirus* richness was found in La Palme. In addition, this lagoon contained the highest numbers of viral particles as observed in Clerissi et al. (2012). The high richness of *Prasinovirus* in La Palme lagoon probably arises because it is rather confined. Indeed, water exchange with the sea may be very limited at many times of the year, since its single passage can be naturally closed (Ifremer, 2008). This might explain the eutrophic nature of this lagoon, which certainly promotes the growth of algae and their associated *Prasinovirus*, without dilution in the sea.

The variation of richness in Leucate was regular, except from June to August 2011, when the lowest values were found. This decrease could be related to lower abundances and diversities of host cells in the water column, since grazing rates on picophytoplankton can be higher in summer in lagoons (Bec et al., 2005). However, these values could also be biased by the flocculation method used to extract viruses. Indeed, the iron flocculation of viruses occurs certainly by adsorption of negatively charged particles onto the positively charged iron (Zhu et al., 2005; John et al., 2010). Hence, iron did not flocculate viruses specifically, but all negatively charged particles. Since the samples from Leucate were difficult to filter in summer, the low richness could be the result of less efficient flocculation of *Prasinovirus*. In contrast, the richness of *Prasinovirus* during a year in SOLA is close to a unimodal distribution, displaying its highest value in July 2011. Interestingly, this value is higher than all of the lagoon richness values for February 2011 (except in La Palme), which suggests that their diversity is not less important in this coastal station at the community scale. Accordingly, Bellec et al. (2010a) found that Leucate contained more OtV, but exhibited a similar genotype diversity to that seen in SOLA. Did this increase of diversity in SOLA have an autochthonous or allochthonous origin? On one hand, this increase of diversity in summer could be directly influenced by the composition of host community in SOLA: the increase in richness could be the result of bottom-up factors favouring host growth, whereas the decrease observed in September 2011 could be related to top-down factors, such as ciliate and flagellate grazers (Raven et al., 2005). However, it has been observed that picoeukaryotes were more abundant in winter over a one-year cycle (Charles et al., 2005). On the other hand, Bellec et al. (2010a) found that viral abundance in SOLA was positively correlated with chlorophyll *a*, but negatively correlated with phosphates, and we found that phosphate concentrations are up

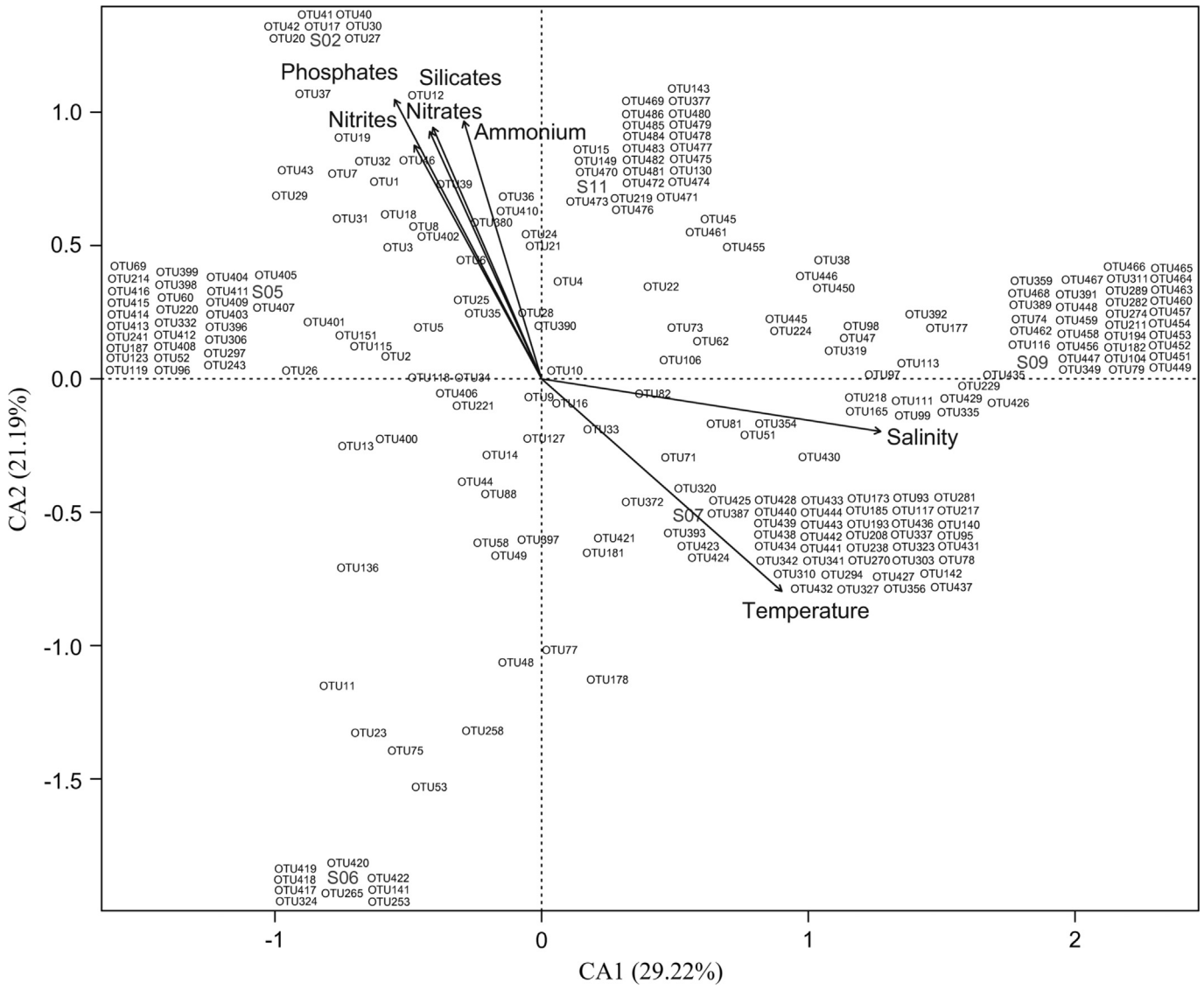

Fig. 8. Correspondence analysis of the SOLA samples on *Prasinovirus* assemblages and environmental vector fitting. Stations are named from S02 to S11 for SOLA-February to SOLA-November. OTU are defined for a nucleotide identity of 90%. Salinity (g/L); Temperature (°C); Nitrates (μM); Nitrites (μM); Ammonium (μM); Phosphates (μM); Silicates (μM). Only the salinity was retained using the forward-selection implemented in the CCA analysis.

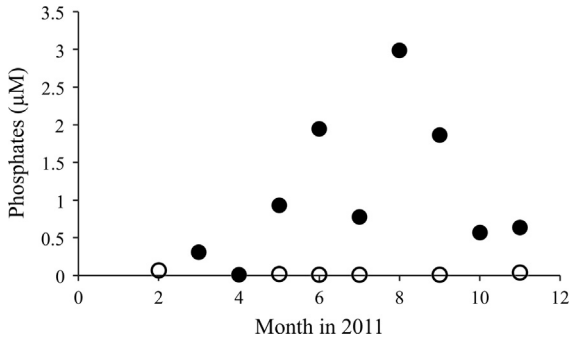

Fig. 9. Phosphate concentrations for the temporal analysis. •: Leucate; ○: SOLA.

to three times greater in Leucate during summer. Since freshwater influx is reduced in the summer because of lower precipitation (see above), this increased level of phosphates might be of anthropological origin, because the local population density increases several fold in the summer (Khan and Ansari, 2005; García-Pintado et al., 2007). If this increase of phosphate concentrations also occurred in the other lagoons and promoted higher primary production of photosynthetic picoeukaryotes, such as in Leucate (Bec et al., 2005), lagoons could thus act as natural chemostats that favor *Prasinovirus* richness in SOLA, by leakage of their richer contents to coastal waters. However, although connection was highlighted between lagoons and SOLA using the Mantel test for the whole dataset and the spatial analysis (where no geographic segregation was observed), the clustering analysis suggests that the environment selects the composition of viral communities in SOLA. As a consequence, the richness in SOLA could be enhanced by lagoons, but this contribution is not random.

Among the environmental variables, CCA suggests that salinity constrained the viral distribution in SOLA during the year. This variable is certainly a signal of seasonality, since, as previously discussed, the increase of the salinity was likely affected by a temperature increase and a decrease of freshwater input from the Baillaury river. Finally, among all the variables tested, phosphate was retained as the most significant variable constraining *Prasinovirus* assemblages, not only for the spatial and temporal analysis, but also for the whole dataset. We hypothesize that this constraint could be direct and/or indirect, since (i) nutrients could

influence the composition of picoplankton (Fuller et al., 2006; Mackey et al., 2009; Kirkham et al., 2011), which then permits growth of *Prasinovirus* strains, because they are able to infect available host cells, and (ii) synthesis of viral genomes requires a good supply of phosphates, highlighted by the observation that many *Prasinovirus* genomes encode phosphate transporters, certainly acquired from hosts by lateral gene transfers (Weynberg et al., 2011; Monier et al., 2012). Like the *psb*A gene found in *Prochlorococcus* and *Synechococcus* phage genomes (Mann et al., 2003; Lindell et al., 2004) that maintains host photosynthesis throughout the infection cycle (Lindell et al., 2005), phosphate transporters may increase *Prasinovirus* fitness in a similar way to the PstS gene in cyanobacteria (Kelly et al., 2013). Indeed, phosphate limitation was shown to affect viral lysis (Bratbak et al., 1993; Maat et al., 2014) and phage life cycle (Wilson et al., 1996), certainly because viruses have a high nucleic acid to protein ratio (Hershey and Chase, 1952). Hence, future studies should test whether transporters (or transporter combinations) have various affinities to phosphates, and if they play a crucial role in *Prasinovirus* fitness at different taxonomic resolutions.

To conclude, the five locations studied are environmentally different enough to harbor particular *Prasinovirus* communities. Lagoons contained higher specific richness than the coastal station SOLA, and particularly the La Palme lagoon, where *Prasinovirus* were the most diverse in this study and also the most abundant in a previous study by Clerissi et al. (2012). This spatial pattern was not the result of geographical barriers, but the signature of environmental selection of *Prasinovirus* genotypes. In addition, for the spatial comparison done in February 2011, nutrient-rich conditions probably increased the specific richness of *Prasinovirus*. The comparison of *Prasinovirus* assemblages at two sampling sites over a period of 10 months suggested (i) that in contrast to OtV abundances (Bellec et al., 2010a), specific richness of *Prasinovirus* in SOLA could be higher than that found in lagoons, (ii) that *Prasinovirus* assemblages were different throughout the year between SOLA and Leucate, (iii) that lagoons might contribute to the diversity of SOLA in summer, and (iv) that the environment and particularly seasonality shaped *Prasinovirus* communities in both locations. Finally, the preponderant link with phosphates and the presence of phosphate-uptake genes in *Prasinovirus* genomes should be considered as driving factors in future analyses that aim to understand the dynamics of *Prasinovirus* in natural environments.

Materials and methods

Sampling sites and environmental variables

The sample sites are located in NW Mediterranean Sea (Fig. 1) and include four lagoons (Leucate, La Palme, Bages-Sigean and Thau) and one coastal station (SOLA). Water samples from the five sites were collected at surface (between 0 and 5 m) by filtration between February 7th and 11th, 2011 and individual samples were processed the same day on return to the lab. Two sites (Leucate and SOLA) were sampled once a month during 10 months between 7 February 2011 and 17 November 2011 (Leucate sample representing the month of March was actually taken on February 25, 2011). The following variables were measured in order to describe each sample: mean depth (m), in situ temperature (°C), salinity (g/L), GPS coordinates (WGS84), ammonium (μM), phosphates (μM), silicates (μM), nitrites (μM), and nitrates (μM). Salinity was measured with a refractometer to make global comparisons, but CTD-estimated salinities were available for SOLA samples (national monitoring SOMLIT network).

Prasinovirus sampling and DNA extraction

Twenty litres of seawater were first passed through 20 μm mesh filters to remove larger plankton, then in series through 3, 0.45, and 0.22 μm filters (142 mm, 3 μm PC Membrane, Millipore, Ref. TSTP14250; 142 mm, 0.45 μm Express PLUS Membrane, Millipore, Ref. HPWP14250; 142 mm, 0.22 μm Express PLUS Membrane, Millipore, Ref. GPWP14250) using a peristaltic pump (Watson-Marlow, 624 S). Viruses were flocculated with ferric chloride after the filtration on 0.22 μm, and filtered on 1 μm filters (142 mm, 1 μm PC Membrane, Whatman, Ref. 112110) as described in John et al. (2010). Viruses were resuspended by adding 10 mL of resuspension buffer (2 M EDTA, 2.5 M Tris–HCl, pH 6) for 24 h at 4 °C, with occasional shaking by hand. Then, the virus-containing buffer was concentrated in 2 mL using sucrose gradient (35 g/L NaCl; 40% (2 mL), 50% (1 mL), 65% (1 mL) and 80% (1 mL) fractions) and ultracentrifugation (CO-LE80K Ultracentrifuge, Beckman-Coulter; Polyallomer centrifuge tubes (16 × 102 mm), No. 337986, Beckman, California, USA) for 6 h at 28,000 rpm (~107,000 g). Viruses were collected in the 65–80% fractions. DNA extractions were done on the virus-containing sucrose fractions, following a modified CTAB protocol (Winnepenninckx et al., 1993): (i) the 2 mL of sucrose fractions were incubated at 60 °C for one hour in 8 mL of a CTAB buffer (2.5% CTAB (hexadecyltrimethylammonium bromide); 125 mM TrisHCl (pH 8); 25 mM EDTA; 1.75 M NaCl; 0.25% β-mercaptoethanol; 0.125 mg/mL proteinase K; 12.5 mM DTT (dithiothreitol), (ii) DNA was purified using an equal volume of chloroform/isoamylalcohol (24:1) and a one-hour-long-RNase digestion step, and (iii) DNA was precipitated with a 2/3 volume of isopropanol and washed with 1 mL of a solution containing 76% v/v EtOH and 10 mM ammonium acetate solution. Finally, the extracted DNA samples were dissolved in 50 μL of laboratory grade deionized water and stored at −20 °C until the sequencing steps. An approximate yield of 220 ng/μL was obtained on average for each sample.

Primer, PCR, sequencing and data processing

This part was previously described in detail in Clerissi et al. (2014). Briefly, the PolB of *Prasinovirus* was amplified using the primer sets VpolAS4 (5′–GAR GGI GCI ACI GTI YTN GA-3′) - VpolAAS1 (5′–CCI GTR AAI CCR TAI ACI SWR TTC AT-3′). In order to multiplex different samples for sequencing (454-pyrosequencing), adapters, key and "MID" tags were incorporated at the 5′ ends during primer synthesis. PCR reactions were done in duplicate and set up as follows: 2 μL of environmental DNA (100 ng) was added to a 48 μL reaction mixture which contained 0.2 mM of each deoxyribonucleoside triphosphate, 30 pmol of each primer, 1/10 of Advantage 2 PCR buffer and 1 U of Advantage 2 Polymerase Mix (Cat. no. 639201, Clontech). The PCR was conducted in a Mastercycler (Eppendorf) with an initial step of 95 °C (3 min) followed by 38 rounds at 95 °C (30 s), 50 °C (60 s), 72 °C (90 s) and a final extension at 72 °C (4 min). The amount of environmental DNA (100 ng) was optimized according to experimental assays done using a range of 10–500 ng DNA from a test sample. The PCR protocol, particularly the number of cycles, was optimized using the genomic DNA of OtV5 (Derelle et al., 2008), OlV1, MpV1, BpV1, and BpV2 (Moreau et al., 2010). PCR products were electrophoresed in 0.5% TAE buffer in a 0.8% agarose gel and the expected band was excised by UV visualization after ethidium bromide staining. PCR bands were purified directly using a gel extraction kit (QIAquick, Cat. no. 28704, Qiagen). The DNA concentrations were measured using a Nanodrop 2000 Spectrophotometer (Thermo Scientific) and amplicons were pooled stoichiometrically before sequencing on a GS-FLX Titanium plate. Pyrosequencing and PCR noise were removed using AmpliconNoise v1.25 (Quince et al., 2011) implemented in

QIIME (Quantitative Insights Into Microbial Ecology) v.1.3.0 (Caporaso et al., 2010). Reads lacking a correct primer and/or having less than 360 successful pyrosequencing flows were removed. Then, after the removal of primer sequences, the following steps and the default parameters were used: PyroDist; PyroNoise (−s 60.0 −i 0.01); SeqDist; SeqNoise (−s 25.0 −i 0.08); Perseus. After theoretical translation in the appropriate reading frame, stop codon-containing sequences were removed and the denoised sequences of PolB were trimmed to ~300-bp after the protein motifs GKQLAYK. Finally, blastp searches (Altschul et al., 1990; using the PolB of OtV5 as a query with a 35.4% identity threshold) and phylogenetic analyses were performed to retain only potential *Prasinovirus* sequences, i.e. those positioned phylogenetically within known *Prasinovirus* or between *Chlorovirus* and *Prasinovirus*.

Sequence analysis

Sequence analyses were performed using Mothur (Schloss et al., 2009). Because PolB was amplified with degenerate primers, we dereplicated the sequence data to avoid proportional biases generated by PCR or sequencing. The dereplicated sequence data contained 4365 unique sequences, which are hereafter referred to as viral genotypes. As a consequence, this diversity analysis corresponds to viral specific richness. Sequences were aligned according to the codons, using Muscle v3.8.31 (Edgar, 2004) implemented in Seaview v4.0 (Galtier et al., 1996; Gouy et al., 2010), and clustered at different percentages of sequence identities, from 70 to 100%. Rarefaction curves were produced at different thresholds and allowed us to define the Operational Taxonomic Units (hereafter referred to as OTU) when the curves began to level off, i.e. at 90% for *Prasinovirus* (Fig. S1). This step provided an objective cutoff and permitted us to standardize the different samples. The samples were clustered according to their OTU composition and content using Bray–Curtis dissimilarities and Jaccard indices, respectively, and the function *tree.shared*. Finally, the estimation of the *Prasinovirus* richness was assessed using the bootstrap calculator for an OTU cutoff of 90%.

Prasinovirus annotation

In order to describe the environmental *Prasinovirus* diversity, a PolB alignment of 475 *Prasinovirus* and *Chlorovirus* isolates was reduced to 23 reference sequences, using a OTU cutoff of 90% (see below; Genbank accession numbers are indicated in brackets). These are 1 *Bathycoccus* virus (BpV): BpV87 (FJ267515); 14 *Micromonas* viruses (MpV; viral names indicate which host strain or host clade was used for isolation. For example, Mi1109V14 is the virus #14 isolated using the RCC1109 strain, and MicCV32 is the virus #32 isolated with a clade C-*Micromonas* strain): Mi1109V14 (KF378564); MicCV32 (KF378579); Mi497V14 (KF378565); MicAV11 (KF378568); Mi829V1 (KF378567); MicAV17 (KF378570); MicAV8 (KF378574); MicBV10 (KF378575); MicAV16 (KF378569); MicAV29 (KF378572); MicBV39 (KF378578); MicBV30 (KF378577); MicBV26 (KF378576); MpV1 (HM004429); 6 *Ostreococcus* viruses (OV): O356V303 (KF378581); OlV537 (GQ412097); OtV63 (FJ267501); OtV09_559 (KF378582); OlV464 (GQ412092); OtV6 (JN225873); and 2 *Chloroviruses* (PBCV): PBCV_CVK2 (AB011500); PBCV_NY2A (DQ491002). The annotation was done using each genotype as a query for a blastp search against the 23 reference sequences-containing dataset. The percentages of identities (% ID) were then compared to the values of a similarity matrix computed using the protein sequences of the 23 references. As a consequence, an environmental genotype was considered as a putative *Bathycoccus* virus if % ID ≥ 90.35 with BpV87; a putative *Micromonas* virus if % ID ≥ 80.95 with a MpV; a putative *Ostreococcus* virus if % ID ≥ 92.38 with a OV; a putative *Chlorovirus* if % ID ≥ 90.35 with a PBCV. Furthermore, a phylogenetic reconstruction was computed from an alignement of the 23 reference sequences and 75 environmental OTU (defined using a cutoff of 75%). This cutoff was found to give the clearest tree showing phylogenetic diversity. Phylogenetic reconstructions were based on DNA sequences, using Bayesian inference (BI), with an evolutionary model selected via Akaike Information Criterion and jModelTest v2 (Darriba et al., 2012). Sequences were partitioned according to codon position, and the estimation of model parameters was unlinked across partitions. Bayesian analysis was carried out with MrBayes 3.2 (Ronquist et al., 2012), with 4 chains of 2,000,000 generations, trees sampled every 1000 generations, and burnin value set to 20% of the sampled trees. We checked that standard deviation of the split frequencies fell below 0.01 to ensure convergence in tree search.

Statistical and multivariate analysis

All of the statistical analyses were done using R v3.0.2 (R: a language and environment for statistical computing, 2008; R Development Core Team, R Foundation for Statistical Computing, Vienna, Austria [http://www.R-project.org]). Principal component analysis (PCA) was used to discriminate the different sampling stations according to the environmental variables using the R "FactoMineR" package and the "PCA" function. PCA were done using the environmental variables standardized with the R "vegan" package and the "decostand" function. The significant variables (those that were significantly correlated with the different components) were selected using the R "FactoMineR" package and the "dim.desc" function. Canonical correspondence analysis (hereafter named CCA) was used to investigate the variations in the proportions of the different *Prasinovirus* OTU under the constraint of the environmental variables. We assumed that the response of species to environmental variations was unimodal. We first selected the variables that best explain the variation of *Prasinovirus* using a correspondence analysis and an environmental vector fitting, implemented in the R "vegan" package and in the "cca", "envfit" functions. The significant variables (i.e. variables that significantly explained changes in the distribution of *Prasinovirus*) of the new dataset were then chosen using CCA and a forward-selection procedure (999 permutations), implemented in the R "vegan" package and in the "cca", "ordistep" functions. Finally, in order to disentangle the influence of geography and the environment on the composition of *Prasinovirus* communities, we used Mantel tests (Mantel, 1967) to infer the significance of correlations between virus community distances (based on the proportion of OTU for a 90% cutoff) and geographical distances (GPS coordinates) (accounting for spatial structure), virus community distances and environmental distances (in situ temperature, salinity, ammonium, phosphates, silicates, nitrites, nitrates) (accounting for environmental selection), virus community distances and sampling time distances (indirect inference of environmental selection, according to seasons). The distance matrices were computed using the Bray–Curtis dissimilarity for *Prasinovirus* communities and the Euclidean metric for the environmental variables using the R "vegan" package and the "vegdist" function after a standardization step (R "vegan" package and the "decostand" function). The sampling time distances were computed with month numbers and the Euclidean metric. Mantel tests were computed ultimately using the R "vegan" package and the "mantel" function.

Data deposition footnote

The 454-sequence datasets were submitted to the Sequence Read Archive of the European Nucleotide Archive (accession number: PRJEB6059).

Acknowledgments

We thank the Genophy team in Banyuls-sur-Mer, especially Sheree Yau, for stimulating discussions and anonymous reviewers for useful comments. We are grateful to Séverine Boyer and Emilie Le Floch (ECOSYM, University of Montpellier 2) for helping with sampling in the Thau lagoon, Fabien Joux and Ingrid Obernosterer for lending us a peristaltic pump, the crew of the Néreis II boat for the sampling at SOLA and the SOMLIT network, particularly Ian Salter, who supplied the environmental data of SOLA. This work was supported by the CNRS-EC2CO Grant 'BIOVIR' AO2011-693921/AO2012-765127 (coordinator Y. Desdevises), and the 'Agence Nationale de Recherche' grants 'TARA-GIRUS' ANR-09-PCS-GENM-218 (coordinator H. Ogata) and 'PHYTADAPT' ANR n° NT09_567009 (coordinator H. Moreau). Camille Clerissi benefited from a doctoral fellowship from the AXA Research Fund.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.virol.2014.07.016.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Arnaud, P., Raimbault, R., 1969. L'étang de Salses-Leucate ses principaux caracteres physico-chimiques et leurs variations (en 1955-1956 et de 1960 à 1968). Rev. des Trav. de l'Inst. des Pêches Marit. 33, 335–443.

Bec, B., Husseini-Ratrema, J., Collos, Y., Souchu, P., Vaquer, A., 2005. Phytoplankton seasonal dynamics in a Mediterranean coastal lagoon: emphasis on the picoeukaryote community. J. Plankton Res. 27, 881–894.

Bellec, L., Clerissi, C., Edern, R., Foulon, E., Simon, N., Grimsley, N., Desdevises, Y., 2014. Cophylogenetic interactions between marine viruses and eukaryotic picophytoplankton. BMC Evol. Biol. 14, 59.

Bellec, L., Grimsley, N., Derelle, E., Moreau, H., Desdevises, Y., 2010a. Abundance, spatial distribution and genetic diversity of *Ostreococcus tauri* viruses in two different environments. Environ. Microbiol. Rep. 2, 313–321.

Bellec, L., Grimsley, N., Desdevises, Y., 2010b. Isolation of prasinoviruses of the green unicellular algae *Ostreococcus spp.* on a worldwide geographical scale. Appl.i Environ. Microbiol. 76, 96–101.

Bellec, L., Grimsley, N., Moreau, H., Desdevises, Y., 2009. Phylogenetic analysis of new Prasinoviruses (Phycodnaviridae) that infect the green unicellular algae *Ostreococcus*, *Bathycoccus* and *Micromonas*. Environ. Microbiol. Rep. 1, 114–123.

Boehme, J., Frischer, M.E., Jiang, S.C., Kellogg, C.A., Pichard, S., Rose, J.B., Steinway, C., Paul, J.H., 1993. Viruses, bacterioplankton, and phytoplankton in the southeastern Gulf of Mexico: distribution and contribution to oceanic DNA pools. Mar. Ecol. Prog. Ser. 97, 1.

Bratbak, G., Egge, J.K., Heldal, M., et al., 1993. Viral mortality of the marine alga *Emiliania huxleyi* (Haptophyceae) and termination of algal blooms. Mar. Ecol. Prog. Ser. 93, 39–48.

Bratbak, G., Wilson, W., Heldal, M., 1996. Viral control of *Emiliania huxleyi* blooms? J. Mar. Syst. 9, 75–81.

Breitbart, M., Rohwer, F., 2005. Here a virus, there a virus, everywhere the same virus? Trends Microbiol. 13, 278–284.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., et al., 2010. QIIME allows analysis of high-throughput community sequencing data. Nat. Methods 7, 335–336.

Charles, F., Lantoine, F., Brugel, S., Chrétiennot-Dinet, M.-J., Quiroga, I., Rivière, B., 2005. Seasonal survey of the phytoplankton biomass, composition and production in a littoral NW Mediterranean site, with special emphasis on the picoplanktonic contribution. Estuar., Coast. Shelf Sci. 65, 199–212.

Clerissi, C., Desdevises, Y., Grimsley, N., 2012. Prasinoviruses of the marine green alga *Ostreococcus tauri* are mainly species-specific. J. Virol. 86, 4611.

Clerissi, C., Grimsley, N., Desdevises, Y., 2013. Genetic exchanges of inteins between prasinoviruses (Phycodnaviridae). Evolution 67, 18–33.

Clerissi, C., Grimsley, N., Ogata, H., Hingamp, P., Poulain, J., Desdevises, Y., 2014. Unveiling of the diversity of prasinoviruses (*Phycodnaviridae*) in marine samples by using high-throughput sequencing analyses of PCR-amplified DNA polymerase and major capsid protein genes. Appl. Environ. Microbiol. 80, 3150–3160.

Corinaldesi, C., Crevatin, E., Del Negro, P., Marini, M., Russo, A., Fonda-Umani, S., Danovaro, R., 2003. Large-scale spatial distribution of virioplankton in the Adriatic Sea: testing the trophic state control hypothesis. Appl. Environ. Microbiol. 69, 2664–2673.

Courties, C., Vaquer, A., Troussellier, M., Lautier, J., Chretiennot-Dinet, M.J., Neveux, J., Machado, C., Claustre, H., 1994. Smallest eukaryotic organism. Nature 370, 255.

Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9, 772.

Derelle, E., Ferraz, C., Escande, M.-L., et al., 2008. Life-cycle and genome of OtV5, a large DNA virus of the pelagic marine unicellular green alga *Ostreococcus tauri*. PLoS One 3, e2250.

Derolez, V., Leurion, A., Fiandrino, A., 2012. Directive Cadre Eau. Estimation du degré d'échange avec la mer pour les lagunes du bassin Rhône Méditerranée & Corse (Ref RST. DOP/LER-LR/12-10). Agence de l'Eau, 36.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797.

Fuller, N.J., Tarran, G.A., Yallop, M., Orcutt, K.M., Scanlan, D.J., 2006. Molecular analysis of picocyanobacterial community structure along an Arabian Sea transect reveals distinct spatial separation of lineages. Limnol. Oceanogr., 2515–2526.

Galtier, N., Gouy, M., Gautier, C., 1996. SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. Comput. Appl. Biosci. 12, 543–548.

García-Pintado, J., Martínez-Mena, M., Barberá, G.G., Albaladejo, J., Castillo, V.M., 2007. Anthropogenic nutrient sources and loads from a Mediterranean catchment into a coastal lagoon: Mar Menor, Spain. Sci. Total Environ. 373, 220–239.

Gouy, M., Guindon, S., Gascuel, O., 2010. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol. Biol. Evol. 27, 221–224.

Hershey, A.D., Chase, M., 1952. Independent functions of viral protein and nucleic acid in growth of bacteriophage. J. Gen. Physiol. 36, 39–56.

Ifremer. 2008. Réseau de Suivi Lagunaire du Languedoc-Roussillon: Bilan des résultats 2007. Rapport RSL-08/2008, 363 p.

Jacquet, S., Heldal, M., Iglesias-Rodriguez, D., Larsen, A., Wilson, W., Bratbak, G., 2002. Flow cytometric analysis of an *Emiliana huxleyi bloom* terminated by viral infection. Aquat. Microb. Ecol. 27, 111–124.

John, S.G., Mendez, C.B., Deng, L., et al., 2010. A simple and efficient method for concentration of ocean viruses by chemical flocculation. Environ. Microbiol. Rep. 3, 195–202.

Kelly, L., Ding, H., Huang, K.H., Osburne, M.S., Chisholm, S.W., 2013. Genetic diversity in cultured and wild marine cyanomyoviruses reveals phosphorus stress as a strong selective agent. ISME J. 7, 1827–1841.

Khan, F.A., Ansari, A.A., 2005. Eutrophication: an ecological vision. Bot. Rev. 71, 449–482.

Kirkham, A.R., Jardillier, L.E., Tiganescu, A., Pearman, J., Zubkov, M.V., Scanlan, D.J., 2011. Basin-scale distribution patterns of photosynthetic picoeukaryotes along an Atlantic Meridional Transect. Environ. Microbiol. 13, 975–990.

Lami, R., Ghiglione, J., Desdevises, Y., West, N.J., Lebaron, P., 2009. Annual patterns of presence and activity of marine bacteria monitored by 16S rDNA–16S rRNA fingerprints in the coastal NW Mediterranean Sea. Aquat. Microb. Ecol. 54, 199–210.

Lindell, D., Jaffe, J.D., Johnson, Z.I., Church, G.M., Chisholm, S.W., 2005. Photosynthesis genes in marine viruses yield proteins during host infection. Nature 438, 86–89.

Lindell, D., Sullivan, M.B., Johnson, Z.I., Tolonen, A.C., Rohwer, F., Chisholm, S.W., 2004. Transfer of photosynthesis genes to and from Prochlorococcus viruses. Proc. Natl. Acad. Sci. U.S.A 101, 11013.

Llas, Picot B., Illes, S., Bontoux, J., 1990. Structure spatiale des sels nutritifs au sein d'un ecosysteme lagunaire: l'étang de Thau. Water Res. 24, 1479–1489.

Maat, D.S., Crawfurd, K.J., Timmermans, K.R., Brussaard, C.P.D., 2014. Elevated CO2 and phosphate limitation favor *Micromonas pusilla* through stimulated growth and reduced viral impact. Appl. Environ. Microbiol. 80, 3119–3127.

Mackey, K.R.M., Rivlin, T., Grossman, A.R., Post, A.F., Paytan, A., 2009. Picophytoplankton responses to changing nutrient and light regimes during a bloom. Mar. Biol. 156, 1531–1546.

Mann, N.H., Cook, A., Millard, A., Bailey, S., Clokie, M., 2003. Bacterial photosynthesis genes in a virus. Nature 424, 741.

Manrique, J.M., Calvo, A.Y., Jones, L.R., 2012. Phylogenetic analysis of *Ostreococcus* virus sequences from the Patagonian Coast. Virus Genes 45, 316–326.

Mantel, N., 1967. The detection of disease clustering and a generalized regression approach. Cancer Res. 27, 209.

Marin, B., Melkonian, M., 2010. Molecular phylogeny and classification of the Mamiellophyceae class. nov.(Chlorophyta) based on sequence comparisons of the nuclear-and plastid-encoded rRNA operons. Protist 161, 304–336.

Monier, A., Welsh, R.M., Gentemann, C., Weinstock, G., Sodergren, E., Armbrust, E.V., Eisen, J.A., Worden, A.Z., 2012. Phosphate transporters in marine phytoplankton and their viruses: cross-domain commonalities in viral-host gene exchanges. Environ. Microbiol. 14, 162–176.

Moreau, H., Piganeau, G., Desdevises, Y., Cooke, R., Derelle, E., Grimsley, N., 2010. Marine Prasinovirus genomes show low evolutionary divergence and acquisition of protein metabolism genes by horizontal gene transfer. J. Virol. 84, 12555–12563.

O'Kelly, C.J., Sieracki, M.E., Thier, E.C., Hobson, I.C., 2003. A transient bloom of *Otreococcus* (chlorophyta, prasinophyceae) in west neck bay, long island, New York. J. Phycol. 39, 850–854.

Palenik, B., Grimwood, J., Aerts, A., et al., 2007. The tiny eukaryote *Ostreococcus* provides genomic insights into the paradox of plankton speciation. Proc. Natl. Acad. Sci. U.S.A 104, 7705–7710.

Paterson, S., Vogwill, T., Buckling, A., et al., 2010. Antagonistic coevolution accelerates molecular evolution. Nature 464, 275–278.

Quince, C., Lanzen, A., Davenport, R.J., Turnbaugh, P.J., 2011. Removing noise from pyrosequenced amplicons. BMC Bioinform. 12, 38.

Raven, J.A., Finkel, Z.V., Irwin, A.J., 2005. Picophytoplankton: bottom-up and top-down controls on ecology and evolution. Vie Milieu 55, 209–215.
Ronquist, F., Teslenko, M., van der, Mark P., et al., 2012. 2012. MrBayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542.
Schloss, P.D., Westcott, S.L., Ryabin, T., et al., 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl. Environ. Microbiol. 75, 7537–7541.
Schroeder, D.C., Oke, J., Hall, M., Malin, G., Wilson, W.H., 2003. Virus succession observed during an *Emiliania huxleyi* bloom. Appl. Environ. Microbiol. 69, 2484–2490.
Short, S.M., Suttle, C.A., 2002. Sequence analysis of marine virus communities reveals that groups of related algal viruses are widely distributed in nature. Appl. Environ. Microbiol. 68, 1290–1296.
Snyder, J.C., Wiedenheft, B., Lavin, M., Roberto, F.F., Spuhler, J., Ortmann, A.C., Douglas, T., Young, M., 2007. Virus movement maintains local virus population diversity. Proc. Natl. Acad. Sci. U.S.A 104, 19102.
Subirana, L., Péquin, B., Michely, S., et al., 2013. Morphology, genome plasticity, and phylogeny in the genus *Ostreococcus* reveal a cryptic species, *O. mediterraneus* sp. nov. (Mamiellales, Mamiellophyceae). Protist 164, 643–659.
Suttle, C.A., 2005. Viruses in the sea. Nature 437, 356–361.
Taylor, G.T., Hein, C., Iabichella, M., 2003. Temporal variations in viral distributions in the anoxic Cariaco Basin. Aquat. Microb. Ecol. 30, 103–116.
Thomas, R., Grimsley, N., Escande, M., Subirana, L., Derelle, E., Moreau, H., 2011. Acquisition and maintenance of resistance to viruses in eukaryotic phytoplankton populations. Environ. Microbiol. 13, 1412–1420.
Van Etten, J.L., Graves, M.V., Müller, D.G., Boland, W., Delaroque, N., 2002. Phycodnaviridae – large DNA algal viruses. Arch. Virol. 147, 1479–1516.
Van Valen, L., 1973. A new evolutionary law. Evol. Theory 1, 1–30.
Vidussi, F., Marty, J.-C., Chiavérini, J., 2000. Phytoplankton pigment variations during the transition from spring bloom to oligotrophy in the northwestern Mediterranean sea. Deep Sea Res. Part I 47, 423–445.
Viprey, M., Guillou, L., Ferreol, M., Vaulot, D., 2008. Wide genetic diversity of picoplanktonic green algae (Chloroplastida) in the Mediterranean Sea uncovered by a phylum-biased PCR approach. Environ. Microbiol. 10, 1804–1822.
Weinbauer, M.G., Fuks, D., Peduzzi, P., 1993. Distribution of viruses and dissolved DNA along a coastal trophic gradient in the northern Adriatic sea. Appl. Environ. Microbiol. 59, 4074–4082.
Weynberg, K.D., Allen, M.J., Ashelford, K., Scanlan, D.J., Wilson, W.H., 2009. 2009. From small hosts come big viruses: the complete genome of a second *Ostreococcus tauri* virus, OtV-1. Environ. Microbiol. 11, 2821–2839.
Weynberg, K.D., Allen, M.J., Gilg, I.C., Scanlan, D.J., Wilson, W.H., 2011. Genome sequence of *Ostreococcus tauri* Virus OtV-2 throws light on the role of picoeukaryote niche separation in the ocean. J.Virol. 85, 4520–4529.
Wilson, W.H., Carr, N.G., Mann, N.H., 1996. The effect of phosphate status on the kinetics of cyanophage infection in the oceanic cyanobacterium *Synechococcus sp.* WH78031. J. Phycol. 32, 506–516.
Wilson, W.H., Etten, J.L., Allen, M.J., 2009. The phycodnaviridae: the story of how tiny giants rule the world, Lesser Known Large dsDNA Viruses, pp. 1–42 (Available from) http://dx.doi.org/10.1007/978-3-540-68618-7_1.
Winnepenninckx, B., Backeljau, T., Wachter, R., 1993. Extraction of high molecular weight DNA from molluscs. Trends Genet. 9, 407.
Wommack, K.E., Colwell, R.R., 2000. Virioplankton: viruses in aquatic ecosystems. Microbiol. Mol. Biol. Rev. 64, 69–114.
Zhu, B., Clifford, D.A., Chellam, S., 2005. Virus removal by iron coagulation–microfiltration. Water Res. 39, 5153–5161.