



Diversité des prasinovirus (phycodnaviridae) et contrôle par les facteurs environnementaux.

Hugo Lebretonchel

► **To cite this version:**

Hugo Lebretonchel. Diversité des prasinovirus (phycodnaviridae) et contrôle par les facteurs environnementaux.. Microbiologie et Parasitologie. Université Pierre et Marie Curie - Paris VI, 2016. Français. <NNT : 2016PA066005>. <tel-01345234>

HAL Id: tel-01345234

<https://tel.archives-ouvertes.fr/tel-01345234>

Submitted on 13 Jul 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE DE DOCTORAT DE L'UNIVERSITÉ PIERRE ET MARIE CURIE

Spécialité : Complexité du Vivant (ED 515)

Pour obtenir le grade de

DOCTEUR de l'UNIVERSITÉ PIERRE ET MARIE CURIE



**Diversité des *Prasinovirus*
(Phycodnaviridae) et contrôle par les
facteurs environnementaux**

Présentée et soutenue publiquement par

M. Hugo Lebretonchel

Le 21 janvier 2016

Devant un jury composé de :

Stéphan Jacquet, Directeur de recherche INRA, Thonon-les-Bains

Purificación López-García, Directeur de recherche CNRS, Université Paris-Sud

Laure Guillou, Directeur de recherche CNRS, Station Biologique de Roscoff

Hélène Montanié, Maître de conférences, Université de La Rochelle

Yves Desdevises, Professeur, UPMC, Observatoire Océanologique de Banyuls-sur-Mer

Rapporteur

Rapporteur

Examineur

Examineur

Directeur de thèse

THÈSE DE DOCTORAT DE L'UNIVERSITÉ PIERRE ET MARIE CURIE

Spécialité

Complexité du Vivant

Présentée par

M. Hugo Lebretonchel

Pour obtenir le grade de

DOCTEUR de l'UNIVERSITÉ PIERRE ET MARIE CURIE

**Diversité des *Prasinovirus*
(*Phycodnaviridae*) et contrôle par les
facteurs environnementaux**

Le 21 janvier 2016

Devant un jury composé de :

Stéphane Jacquet, Directeur de recherche INRA, Thonon-les-Bains

Rapporteur

Purificación López-García, Directeur de recherche CNRS, Université Paris-Sud

Rapporteur

Laure Guillou, Directeur de recherche CNRS, Station Biologique de Roscoff

Examineur

Hélène Montanié, Maître de conférences, Université de La Rochelle

Examineur

Yves Desdevises, Professeur, UPMC, Observatoire Océanologique de Banyuls-sur-Mer

Directeur de thèse

Remerciements

Trois ans ! Au départ, chaque jour marqué par l'ambition, la motivation, un enthousiasme extrême où l'on met en place une quantité astronomique de projets.

Mais après plus de 2200 litres d'eau de mer filtrée, 300 litres d'eau de mer concentrés, des litres de cultures et de lysats, des kilos de produits chimiques, le temps passe si vite qu'il faut se remettre en question mais aussi et surtout en poser.

Une thèse ne peut pas se résumer seulement en terme d'hypothèses, de manips, de résultats et de conclusions, car il faut y ajouter les discussions, les conseils, les réconforts et les encouragements.

En synthèse, la préparation d'un doctorat n'est pas qu'une aventure scientifique, c'est aussi une aventure humaine, marquée par la rencontre de personnes parfois déroutantes, formidables ou ennuyeuses, mais bien souvent inoubliables. Ce sont ces personnes que je souhaite ici, très sincèrement, remercier.

Je souhaite commencer par remercier l'ensemble des membres du jury, qui me fait l'honneur de juger ce travail de doctorat.

Un remerciement particulier pour mon directeur de thèse, Yves Desdevises, qui m'a donné l'opportunité de réaliser cette thèse et fait confiance pour la conduite de ce projet. J'ai appris beaucoup et j'apprends encore. Je sais que je te dois beaucoup, et pour ça merci encore.

Merci à toute l'équipe Genophy (Nigel Grimsley, Gwenael Piganeau, Hervé Moreau, Evelyne Derelle, Sophie Sanchez-Ferandin, Élodie Desgranges, Sheree Yau, Marc Krasovec) pour votre accueil, nos discussions et remarques durant les bons moments et votre soutien durant les moments plus stressants.

Je tiens aussi à saluer et remercier Télésphore Sime-Ngando et Pascal Hingamp pour leurs conseils et remarques lors des comités de suivi de thèse. Le temps que vous m'avez consacré m'a aidé à prendre du recul sur les points les plus complexes.

Merci à l'équipe SOMLIT (Eric Maria et Aurore Gueux) pour leurs conseils lors de la mesure des paramètres environnementaux. Ainsi que l'équipe de la plateforme d'imagerie et de cytométrie (Christophe Salmeron, David Pecqueur et Marie-Line Escande) pour leurs conseils et coups de main.

Le laboratoire ne pourrait pas fonctionner sans l'ensemble du personnel, qui chaque jour, nous permet de travailler dans des conditions agréables et que l'on oublie trop vite. Merci donc à Vanda Mesquita-Martins pour les bureaux et les petites discussions très tôt le matin, merci aussi à Didier Peuzé pour les factures, contrats et autres péripéties administratives. Jean-Michel et Émilie pour leur accueil à la cantine du labo. Merci à tous les autres pour les petites discussions pendant les pauses cafés ou simplement dans les couloirs.

Ces pensées sont aussi pour mes collègues mais aussi amies de bureau, Sophie et Evelyne. Merci pour ces moments passés en votre compagnie, pour m'avoir laissé exprimer mes « craquages » avec autant de force que mes plus grandes joies. Votre écoute et vos conseils m'ont apporté beaucoup, j'aimerais bien un jour que nous partagions à nouveau ces petites chansons, qui arrivent et restent loooongtemps.

Un grand merci aussi à Anne-Claire Baudoux et Nathalie Simon, qui m'ont fait découvrir le monde des virus et m'ont transmis leur dynamisme pour continuer dans ce fabuleux univers. J'espère que nous aurons à nouveau l'occasion de travailler ensemble.

Que serais-je sans les « collègues », qui sont devenus bien plus que ça, durant ces années à Banyuls. Nous avons partagé bien plus que de la science pendant nos discussions autour d'une Chouf. Merci à Matthias et Marianna (pour leur réconfort, leur bonne humeur et tous les moments que l'on a partagé ensemble), à Marc (pour les discussions semi-sérieuses sur le monde et les 95%), à Mathieu (j'espère arriver à ton niveau en chasse-sous marine et en course, je t'aurai un jour, je t'aurai !), à Sandrine (pour les chats, et son entrain), à Hugo (je te vaincrai un jour prochain). Mais aussi aux amis qui sont loin maintenant (Romain, Eriko, Remy, Tatiana, Elsa, Brian, Marine, Jessi, Mareike), ou qui sont loin depuis un moment (Aurélien et Chacha, par exemple).

J'oublie de remercier certains d'entre vous, n'y voyez pas de raison particulière, lorsque l'on vous dira « Tiens ! C'est toi le prochain maintenant ! », vous comprendrez ! Et je suis sûr que vous saurez aussi vous reconnaître dans ces quelques lignes ! Vous êtes ma famille de cœur.

Je gardais les plus importants pour la fin, merci à ma famille pour le soutien qu'elle m'a apporté, m'apporte et m'apportera. Donc merci à mes parents pour leurs encouragements et leurs petites questions toujours un peu délicates « Comment ça avance ta thèse ? », je ne serais pas là sans vous. Merci à Cécile et Jean-Claude, pour leur conseil, confiance et soutien.

Un homme n'est rien sans la femme qu'il aime, merci à toi Océane, pour ta patience surtout, avoir un doctorant à la maison n'est pas toujours drôle, mais cette période s'achève. De nouveaux projets sont prévus, et d'autres fourmillent déjà d'impatience.

A mes proches, famille et amis

How inappropriate to call this planet « Earth », when it is quite clearly « Ocean ».

(Arthur C. Clarke, 1990)

Table des matières

<i>Chapitre I - Introduction</i>	1
I HISTORIQUE DE LA VIROLOGIE ET DECOUVERTE DES VIRUS MARINS	1
II LES VIRUS	3
1) DESCRIPTION ET DIVERSITE	3
2) ORIGINE ET PHYLOGENIE.....	5
3). LES VIRUS AQUATIQUES.....	8
III INFLUENCE DES VIRUS SUR L'ENVIRONNEMENT	11
1) INFLUENCE DES VIRUS SUR LES CYCLES BIOGEOCHIMIQUES	11
2) IMPLICATION DES VIRUS SUR LA COMPOSITION DES COMMUNAUTES PHYTOPLANCTONIQUES.....	13
3) IMPACT DES VIRUS SUR LE CLIMAT.....	15
IV INFLUENCE DE L'ENVIRONNEMENT SUR LES VIRUS	16
V LE SYSTEME HOTE-VIRUS UTILISE MAMIELLOPHYCEAE-PRASINOVIRUS	21
1) LES MAMIELLOPHYCEAE.....	21
2) LES PRASINOVIRUS	25
VI OBJECTIF ET ORGANISATION DE LA THESE	29
<i>Chapitre II - Matériel et Méthodes</i>	35
I ZONE D'ECHANTILLONNAGE	35
II STRATEGIE D'ECHANTILLONNAGE	41
1) ECHANTILLONNAGE ET ISOLEMENT.....	41
a. <i>Les prasinovirus</i>	41
b. <i>La fraction phytoeucaryotique</i>	44
2) EXTRACTION ET AMPLIFICATION.....	45
a. <i>Extraction d'ADN</i>	45
b. <i>Amplification et séquençage</i>	47
c. <i>PCR quantitative sur les Mamiellophyceae</i>	50
III TRAITEMENT DES SEQUENCES	53
1) LES VIRUS ISOLES.....	53
a. <i>Les communautés environnementales</i>	53
b. <i>Les communautés virales</i>	54

2) LES COMMUNAUTES EUCARYOTES.....	54
IV MESURES DES VARIABLES ENVIRONNEMENTALES.....	55
1) MESURE DE L'AMMONIUM	57
2) VARIABLES SATELLITAIRES	57
3) ANOMALIE DE DENSITE	57
V ANALYSES STATISTIQUES ET MULTIDIMENSIONNELLES.....	58
1) ANALYSES STATISTIQUES.....	58
2) ANALYSES MULTIDIMENSIONNELLES	59
 <i>Chapitre III – Interplay between Micromonas species and their viruses in the Western English Channel.....</i>	 <i>63</i>
 <i>Chapitre IV - Green microalgae of the Mamiellophyceae dominate picoeukaryotic communities throughout the year in a coastal lagoon revealed by qPCR.....</i>	 <i>83</i>
 <i>Chapitre V – Impact of environmental factors on Prasinovirus communities</i>	 <i>115</i>
 <i>Chapitre VI – Discussion générale et perspectives.....</i>	 <i>151</i>
I DISCUSSION GENERALE.....	151
1) DESCRIPTION DE NOUVEAUX PRASINOVIRUS	151
2) DYNAMIQUE DES PRASINOVIRUS ET MAMIELLOPHYCEAE DANS DES ENVIRONNEMENTS CONTRASTES 155	
<i>a. Interaction hôte-virus et hypothèse de distribution.....</i>	<i>155</i>
<i>b. Diversité et distribution saisonnières des communautés de Mamiellophyceae</i>	<i>156</i>
<i>c. Diversité et variations saisonnières des communautés de Prasinovirus.....</i>	<i>159</i>
<i>d. Comparaison des hypothèses sur la distribution des Prasinovirus avec les observations spatio-temporelles.....</i>	<i>164</i>
<i>e. Vers l'identification d'un contrôle environnemental.....</i>	<i>165</i>
3) LIMITE METHODOLOGIQUE POUR L'ETUDE DES COMMUNAUTES HOTES-VIRUS	169

a. <i>Utilisation de la qPCR et du séquençage haut débit pour l'étude des communautés eucaryotiques</i>	169
b. <i>L'étude des communautés de Prasinovirus par séquençage haut débit</i>	171
II CONCLUSION ET PERSPECTIVES	172
<i>Références</i>	177
<i>Annexes</i>	211

Liste des abréviations

Les abréviations sont indiquées en français et/ou en *anglais* suivant l'ordre alphabétique.

18SV4	Région variable V4 du 18S
18SV9	Région variable V9 du 18S
ACC (<i>CCA</i>)	Analyse Canonique des Correspondances
ACP (<i>PCA</i>)	Analyse en Composantes Principales
ADNr	ADN ribosomique
AFC (<i>CA</i>)	Analyse Factorielle des Correspondances
AIC	<i>Akaike Information Criterion</i>
ARN (RNA)	Acide ribonucléique
ATCV	<i>Acanthocystis turfacea Chlorella virus</i>
ATP (<i>PTA</i>)	Analyse Triadique Partielle
BET	Bromure d'éthidium
BLAST	<i>Basic Local Alignment Search Tool</i>
BpV	Virus de <i>Bathycoccus prasinos</i>
CCMP	anciennement le NCMA
CeV	Virus de <i>Chrysochromulina ericina</i>
Chl	Chloroplaste
Chla	Chlorophylle <i>a</i>
COP	Carbone Organique Particulaire
CroV	Virus de <i>Cafeteria roenbergensis</i>
CruV	Virus de <i>Crustomastix</i>
CTAB	Bromure de cétrimonium
Cyt	Cytoplasme
db (ds)	double brin
DGGE	<i>Denaturing Gradient Gel Electrophoresis</i>
DMS	Diméthylsulfone
DMSP	Diméthylsulfoniopropionate
dNTP	DésoxyNucléotide triphosphate
DTT	Dithiothréitol
ECMWF	<i>European Centre for Medium-Range Weather Forecasts</i>

EDTA	Acide éthylène diamine tétraacétique
FCM	Cytométrie en flux
FISH	<i>Fluorescent In Situ Hybridization</i>
FSC	<i>Forward Scatter</i>
HaV	Virus de <i>Heterosigma akashiwo</i>
INRA	Institut National de la Recherche Agronomique
ITS	<i>Internal Transcribed Spacer</i>
Kb	Kilo base
KDa	Kilo dalton
LBBE	Laboratoire de Biométrie et Biologie Evolutive
LUCA	<i>Last Universal Cellular Ancestor</i>
M. pusilla	<i>Micromonas pusilla</i>
MaV	Virus de <i>Mamiella</i>
Mb	Mégabase
MCP	<i>Major Capsid Protein</i>
MES	Matière en Suspension
MID	<i>Multiplex Identifiers</i>
ML (LM)	Maximum de vraisemblance
MOD (DOM)	Matière organique dissoute
MOLA	<i>Microbial Observatory Laboratoire Arago</i>
MOOSE	<i>Mediterranean Ocean Observing System on Environment</i>
MOP	Matière Organique Particulaire
MpV	Virus de <i>Micromonas pusilla</i>
NCBI	<i>National Center for Biotechnology Information</i>
NCLDV	<i>Nucleo-Cytoplasmic Large DNA Virus</i>
NCMA	<i>National Center for Marine Algae and Microbiota</i>
NIES	<i>National Institute for Environmental Studies</i>
NOP	Azote Organique Particulaire
OIV	Virus de <i>Ostreococcus lucimarinus</i>
OmV	Virus de <i>Ostreococcus mediterraneus</i>
OTU	<i>Operational Taxonomic Units</i>
OtV	Virus de <i>Ostreococcus tauri</i>
pb (bp)	Paire de base

PBCV	Virus de <i>Paramecium bursaria Chlorella</i>
PC	Polycarbonate
PCR	<i>Polymerase Chain Reaction</i>
PES	Polyéthersulfone
PFGE	<i>Pulsed Field Gel Electrophoresis</i>
PFU	<i>Page Forming Unit</i>
PgV	Virus de <i>Phaeocystis globosa</i>
PolB	ADN Polymérase virale type B
PoV	Virus de <i>Pyramimonas orientalis</i>
PPE	<i>Photosynthetic PicoEukaryote</i>
PpV	Virus de <i>Phaeocystis pouchetii</i>
qPCR	PCR quantitative
RCC	<i>Roscoff Culture Collection</i>
RT	Rétro-Transcription
sb (ss)	simple brin
Sg	grain d'amidon
SOLA	Service d'observation du Laboratoire ARAGO
SOMLIT	Service d'Observation en Milieu LItoral
SSC	<i>Side SCatter</i>
T.P.I.	Temps Post-Infection
TAE	Tris-Acétate-EDTA
UPGMA	<i>Unweighted Pair-Group Method using arithmetic Average</i>
UV	Ultraviolet

Liste des figures et tableaux

Chapitre I – Introduction

Figure 1 : Nombre de publications référençant le terme « Virus ».....	2
Figure 2 : Structure interne et externe d'une capsid virale, exemple d'un Prasinovirus.....	4
Figure 3 : Système de degrés de vie, du vivant au non-vivant.....	6
Figure 4 : Proportion en termes de biomasse et d'abondance des procaryotes, eucaryotes et virus.....	8
Figure 5 : Schéma de la boucle microbienne.....	12
Figure 6 : Illustration des différents phénotypes d' <i>Emiliana huxleyi</i>	14
Figure 7 : Relation entre la salinité et l'abondance virale dans les eaux de surface...	19
Figure 8 : Phylogénie consensuelle des principaux groupes d'eucaryotes.....	21
Figure 9 : Phylogénie moléculaire des <i>Mamiellophyceae</i>	22
Figure 10 : Photographie des trois genres dominants des <i>Mamiellophyceae</i> , <i>Ostreococcus</i> , <i>Micromonas</i> et <i>Bathycoccus</i>	22
Figure 11 : Cycle lytique de OtV5.....	25
Figure 12 : Reconstruction phylogénétique des membres de NCLDV.....	26
Figure 13 : Phylogénie des NCLDV.....	27

Chapitre II – Matériel et Méthodes

Figure 1 : Localisation des sites de prélèvement.....	35
Tableau 1 : Stations d'échantillonnage.....	36
Figure 2 : Lagune de Salses-Leucate.....	38
Figure 3 : Circulation par vents d'Autan dans la Lagune de Salses-Leucate	39
Tableau 2 : Culture de <i>Mamiellophyceae</i> utilisée.....	42

Figure 4 : Plage de lyses sur une culture de <i>Mamiellales</i> étalée avec de l'agarose....	43
Tableau 3 : Séquences et caractéristiques des amorces pour les virus	47
Tableau 4 : Séquences et caractéristiques des amorces pour les hôtes.....	49
Tableau 5 : Séquences et caractéristiques des amorces en qPCR.....	51
Tableau 6 : Ensemble des descripteurs environnementaux mesurés	56
Figure 5 : Evolution temporelle de la salinité et de la température dans les 4 sites échantillonnés	56

Chapitre III – Interplay between *Micromonas* species and their viruses in the Western English Channel

Figure 1 : Dynamics of total picoeukaryotes and <i>Micromonas</i> cells and <i>Micromonas</i> clades at the station SOMLIT-ASTAN.....	67
Figure 2 : Dynamics of putative <i>Micromonas</i> virus and plaque-forming units formed on <i>Micromonas</i> clades at the station SOMLIT-ASTAN.....	68
Table 1 : Host range of 45 <i>Micromonas</i> viruses isolated from SOMLIT-Astan throughout the year 2009.....	69
Figure 3 : Growth characteristics of representative <i>Micromonas</i> viruses isolated at SOMLIT-Astan in 2009.....	70
Figure 4 : Phylogenetic relationships between <i>Micromonas</i> viruses.....	71
Table S1 : List of <i>Micromonas</i> cultures used for the determination of MicV host range.....	74
Table S2 : AVS1 and AVS2 primer sequences.....	75
Table S3 : Accession numbers of MicV isolates polB.....	76
Figure S1 : Flow cytograms of <i>Micromonas</i> virus RCC 4265 and natural virus community at SOMLIT-Astan on August 2009.....	78
Figure S2 : Modularity analysis of <i>Micromonas</i> virus infection networks using the complete virus infection matrix.....	79

Figure S3 : Relationship between viral infection kinetics and host growth rates during the one step growth experiments.....	80
--	----

Chapitre IV – Green microalgae of the Mamiellophyceae dominate picoeukaryotic communities throughout the year in a coastal lagoon revealed by qPCR

Figure 1 : Locations of sampling sites.....	88
Table 1 : Algal stains used in this study.....	89
Table 2 : Cell genomic DNA content.....	93
Table 3 : Primers designed for this study.....	95
Table 4 : <i>in silico</i> control of primer specificity	96
Table 5 : Primer specificity	97
Figure S1 : Correlation between cytometry measurements and qPCR estimations	98
Figure 2 : Phytoplankton community abundances.....	99
Figure S2 : <i>Micromonas</i> and <i>Ostreococcus</i> clade distribution	101
Figure 3 : Linear regression between phytoplanktonic communities and qPCR estimations on environmental samples	102

Chapitre V – Impact of environmental factor on Prasinovirus communities

Figure 1 : Sites sampling.....	119
Table 1 : Primer sequences characteristics.....	121
Figure 2 : PCA based on environmental factors	126
Figure 3 : Abundance OTUs percentage for host (V4; 50 first) (right) and Prasinovirus (PolB; 100 first) (left) in the different sampling stations.	128
Figure 4 : UPGMA dendrograms based on Bray-Curtis dissimilarity indices on Mamiellophyceae OTUs.....	131

Figure 5 : UPGMA Tree based on Bray-Curtis dissimilarity indices on Prasinovirus OTUs relative abundance.	133
Figure 6 : Venn diagrams of representing Prasinovirus OTU percentages shared between stations.	134
Figure 7 : Trajectories factor maps from STATICO program representing the structure of the variation in the same factorial space for Mamiellophyceae OTUs and environmental variables	135
Figure 8 : Trajectories factor maps from STATICO program which represented variation structure in same factorial space of Prasinovirus OTUs and environmental variables.....	136
Figure 9 : Canonical correspondence analysis on normalized virus OTUs and environmental variables for all sites	137
Figure 10 : Canonical correspondence analysis on normalized virus OTUs and environmental variables at Leucate A.....	138
Figure 11 : Canonical correspondence analysis on normalized virus OTUs and environmental variables at SOLA.....	139
Figure 12 : Canonical correspondence analysis on normalized virus OTUs and environmental variables at MOLA.	140
Figure 13 : Correspondence analysis on the relative abundances of Prasinovirus and Mamiellophyceae.	141

Chapitre VI – Discussion générale et perspectives

Figure 1 : Cliché de microscopie électronique de CruV.....	152
Figure 2 : PFGE des virus CruV1, Mav1 et MaV2	152
Figure 3 : Cinétique de lyse de CruV.....	152
Figure 4 : Phylogénie des prasinovirus et des Mamiellophyceae	154

Figure 5 : dynamiques théoriques des groupes de Prasinovirus dans les environnements étudiés	160
Figure 6 : abondances relatives (%) des différents groupes de virus aux sites A et B de la lagune.....	162

Chapitre I - Introduction

I Historique de la virologie et découverte des virus marins

Les virus ont eu au cours de l'évolution de la vie de nombreux impacts sur le développement des organismes et des communautés. Les premières traces de la présence de ces entités biologiques microscopiques dans nos sociétés humaines sont identifiées dans des hiéroglyphes égyptiens. L'hypothèse selon laquelle ces scènes décriraient des symptômes associés au virus de la poliomyélite fut confirmée par la découverte de traces de la maladie sur un squelette de 3400 ans avant J-C.

La virologie classique telle que nous la connaissons a vraiment pris son essor en 1796. Malgré le manque d'information sur ces entités, c'est à cette époque qu'une solution, que l'on appellera par la suite « vaccin », fut mise au point contre le virus de la variole. Les travaux du botaniste russe Dimitri Ivanovski sur le tabac, *Nicotiana tabacum*, complétés par ceux du hollandais Martinus Beijerinck, ont permis d'identifier ces agents infectieux, ou plutôt, de définir les virus comme n'étant ni des bactéries, ni des toxines. Les précédentes hypothèses, décrivant les virus comme des toxines ou des bactéries, furent soutenues par de nombreux scientifiques comme Louis Pasteur qui refusait de croire en l'existence de ces entités, et qui par ironie du sort découvrit le vaccin contre le virus de la rage (1885). C'est à cette époque que le terme latin « virus », signifiant « poison », commença à être utilisé.

Les virus infectant des bactéries (bactériophages) furent découverts par Félix d'Hérelle (1917) durant la première Guerre Mondiale, et c'est à cette époque que le processus de lyse cellulaire (décrivant la destruction des cellules par l'agent pathogène) fût décrit.

L'apparition dans les années 30 de la microscopie électronique a permis de voir ces entités biologiques méconnues. Il faudra attendre les années 50 pour qu'une définition précise des virus soit proposée par le biologiste français A. M. Lwoff, décrivant les virus de la façon suivante :

- 1/ les virus ne contiennent qu'un seul type d'acide nucléique (ADN ou ARN) qui constitue le génome viral.
- 2/ les virus se reproduisent à partir de leur matériel génétique et par réplication.
- 3/ les virus sont doués de parasitisme intracellulaire absolu.

C'est aussi à cette époque que les premiers virus marins sont observés par Spencer (1955). L'avènement de la virologie marine moderne pris son essor en 1990 (Figure 1), au moment où la prédominance numérique des virus a été montrée, et environ 50 ans après le début de l'étude de ces entités. Les virus n'ont plus été considérés comme des toxines ou même des « entités non vivantes » provoquant simplement des maladies, mais bel et bien comme des entités biologiques importantes dans les écosystèmes, exerçant un rôle clé dans l'évolution.

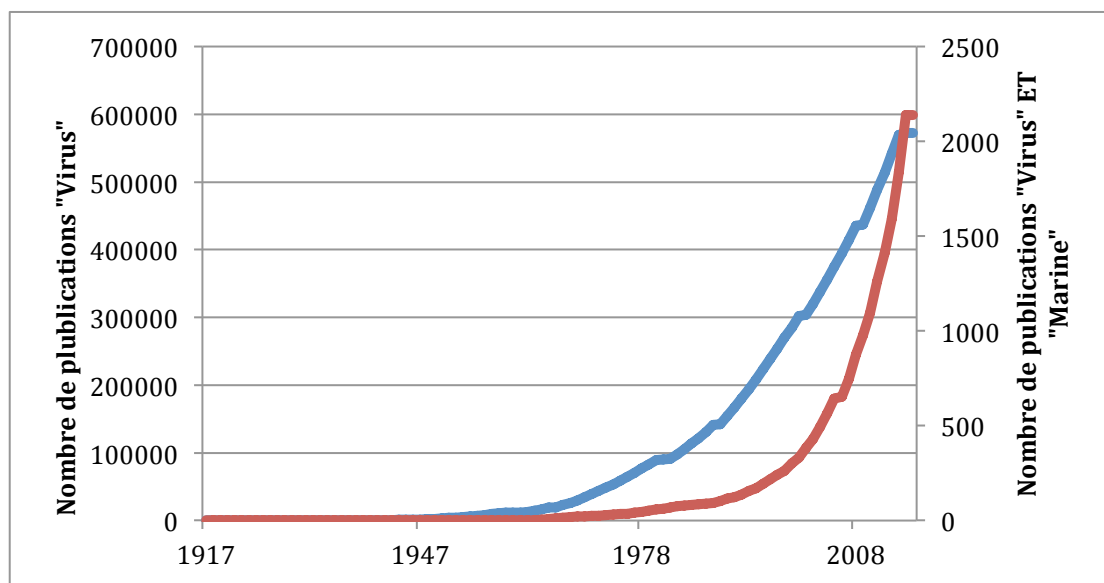


Figure 1 : Quantité cumulée de publications référençant le terme "Virus" (courbe bleue), "Virus" ET "Marine" (courbe rouge) (source : NCBI)

Le développement d'outils comme la microscopie, la cytométrie ou la biologie moléculaire, a rendu possible un approfondissement des connaissances et des concepts de la virologie. A l'heure actuelle, les virus sont beaucoup mieux connus, ainsi que de nombreux mécanismes liés à leur nature si particulière.

Les scientifiques les utilisent maintenant comme outil. Grâce au génie génétique les virus deviennent des vecteurs pour les thérapies géniques, mais aussi des traitements contre les infections de bactéries antibiorésistantes. Félix d'Hérelle qui essaya d'imposer la phagothérapie ne rencontra pas un fort succès en Europe, et c'est en Géorgie (pays du Caucase), qu'il créa le « Phage Therapy Center » (aujourd'hui appelé « Institut Eliava ») où, depuis 1930 sont développées des thérapies phagiques (utilisation des virus comme traitement médical ; Dublanche, 2009). Cet institut possède aujourd'hui une virothèque très développée en phage thérapeutique mais aussi un savoir-faire dans toutes les spécialités médicales, notamment en néonatalogie et pédiatrie (Kutateladze et Adamia, 2008)

II Les virus

1 Description et diversité

La définition des virus proposée par A.M. Lwoff n'a pas significativement évolué. Les virus sont des parasites obligatoires, car à l'opposé des cellules procaryotes ou eucaryotes, ils ne possèdent pas la machinerie nécessaire à la réplication de leur génome et de leur enveloppe protéique. Ils parasitent donc une cellule hôte afin de se répliquer en utilisant les composés organiques, les enzymes et l'énergie de ce même hôte.

Il est communément admis que chaque organisme possède un virus qui lui est propre, virus qui peut lui aussi être attaqué par un virus ou virophage (La Scola, 2008).

Pour résumer leurs particularités structurales (Figure 2), les virus sont formés d'une capsid, sorte de coque protéique parfois elle-même incluse dans une enveloppe phospholipidique. Cette structure close peut se décliner en des formes extrêmement diverses. Elle renferme le génome viral qui existe lui aussi sous diverses formes, et peut être composé d'ADN ou d'ARN, en brin simple ou double, linéaire ou circulaire, segmenté ou non.

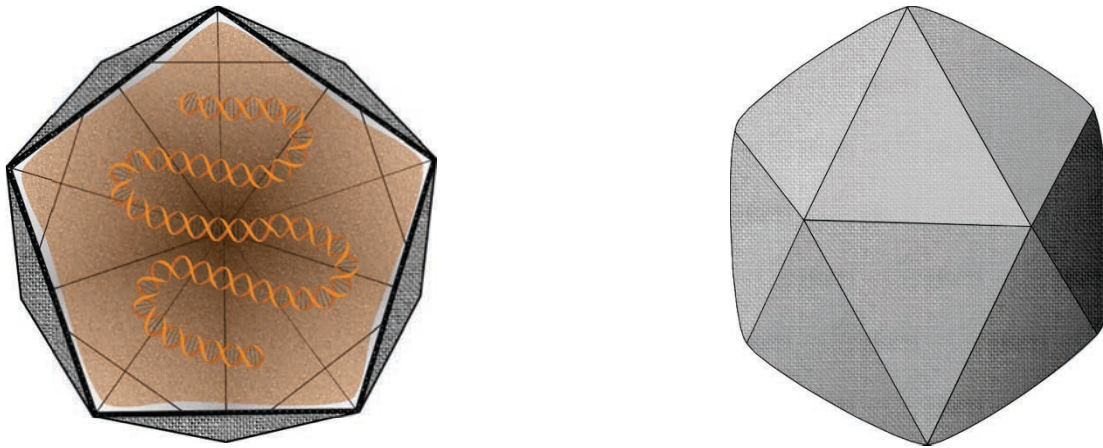


Figure 2 : Structure interne et externe d'une capsid virale, exemple d'un Prasinovirus

Bien que leur mode de réplication soit varié, deux principaux processus peuvent être décrits : la réplication lytique et lysogénique (Berdjed et Jacquet, 2009).

Dans le cas d'une réplication lytique, le virus, après attachement et injection de son génome au sein de la cellule hôte, va détourner l'ensemble du métabolisme de l'hôte pour la production de son propre génome et des différents composants de sa capsid. Suite à l'assemblage des structures produites, la lyse de la cellule est initiée et la libération des virions dans l'environnement s'effectue par la destruction de la cellule.

Le mode de réplication lysogénique se distingue du mode lytique par une phase dite de latence avec l'intégration du génome viral à celui de l'hôte, après l'injection du génome. Le génome viral est ainsi répliqué avec le génome de l'hôte lors des phases de reproduction-multiplication. Dans la majorité des cas, une modification des conditions environnementales, engendrant un stress, déclenche une reprise du cycle viral. Ainsi les descendants de l'hôte originellement infectés vont produire des virus par lyse cellulaire.

Leur réplication est tout de même sujette à des limitations. Lors d'une infection les mécanismes de réplication peuvent être mis en échec par des mécanismes de résistance développés par l'hôte. De plus, la plupart des virus ont une spécificité limitée à une espèce d'hôte, voire parfois limitée à quelques souches au sein d'une espèce (cyanophages : Chen *et al.*, 1996 ; Wang et Chen, 2008, phycodnavirus : Bellec et Desdevises, 2015 ; Baudoux *et al.*, 2015) . Cet aspect implique que pour une « espèce » de virus, il peut exister un grand nombre de souches virales différentes.

Cette diversité globale extraordinaire est un aspect particulier, pouvant peut être témoigner de leur origine ancienne et multiple.

2) Origine et Phylogénie

L'origine des virus, comme leur appartenance au domaine de la vie, fait l'objet d'un débat dont certains aspects peuvent être discutés.

La majorité des virus possèdent un petit génome, l'un des plus petits (Circovirus) possède un génome de 1.7 kb codant pour 2 gènes (Finsterbusch et Mankertz, 2009). Quant au plus gros (Pandoravirus), il possède un génome de 2.5Mb codant 2500 gènes (Philippe *et al.*, 2013). Les 83 familles de virus possèdent leurs caractéristiques propres. Du fait de leur diversité, la classification de Baltimore proposée en 1971 regroupe les virus suivant 7 classes différentes basées sur leurs caractéristiques génomiques (dsDNA, ssDNA, dsRNA, ss(+)RNA, ss(-)RNA, ssRNA(RT) et SSDNA(RT)) (<http://www.ictvonline.org>).

Les virus n'ont probablement pas les mêmes ancêtres communs, comme le met en évidence le fait qu'ils ne forment pas un groupe monophylétique.

A l'opposé des organismes cellulaires, qui ont comme ancêtre commun supposé LUCA (*Last Universal Cellular/Common Ancestor* (Glansdorff *et al.*, 2008)), les virus n'ont pas d'ancêtre commun présumé et pourraient précéder LUCA, avec donc une apparition antérieure à 3,8 milliards d'années (Forterre, 2006). En revanche d'autres auteurs ont estimé leur apparition à 100 millions d'années, comme le suggère l'étude de séquences d'anciens virus endogènes (Johnson, 2010).

Certains auteurs proposent l'hypothèse que les virus à ARN seraient les premières entités biologiques et auraient possiblement un rôle dans la création du noyau des eucaryotes (Claverie et Abergel, 2010; Takemura, 2001), voire dans création des trois domaines du vivant (Forterre, 2006).

Cet état de fait qui place les virus à l'origine de la vie, et donc aux frontières du vivant, implique un débat sur l'aspect vivant ou non de ces entités particulières.

Ce débat, non résolu jusqu'à présent, repose principalement sur la conception même de la vie. La première conception moderne de la vie fut proposée par Kant. Selon lui, la vie ne serait qu'une forme spéciale d'organisation dotée de sa propre finalité (critique de la faculté de juger). Mais cette définition philosophique n'est pas réellement compatible avec la conception biologique. Dans ce cadre, deux approches sont alors en concurrence. La première est basée sur l'aspect individuel des organismes, qui pose le postulat qu'un être vivant peut se résumer à un système autopoïétique du grec « *auto* », soi-même, et « *poièsis* » : production, c'est-à-dire un système qui se reproduit et se régénère. La deuxième approche est basée sur l'aspect évolutif : dans cette optique la capacité d'évolution devient une caractéristique essentielle de la vie.

De ces deux approches dérivent une importante quantité de définitions. Richard Dawkins (1976) proposa de considérer la vie comme une simple sélection naturelle de gènes ou de manière plus générale, de réplicateurs. La NASA, en 1992, pour un programme d'exobiologie, proposa que « la vie est un système auto-entretenu, capable d'évolution darwinienne ».

Dans cette multitude de définitions, Malaterre a proposé en 2010 que la vie est une affaire de « plus ou moins » (degrés de vie, Figure 3), et c'est dans ce contexte que les virus s'inscrivent.

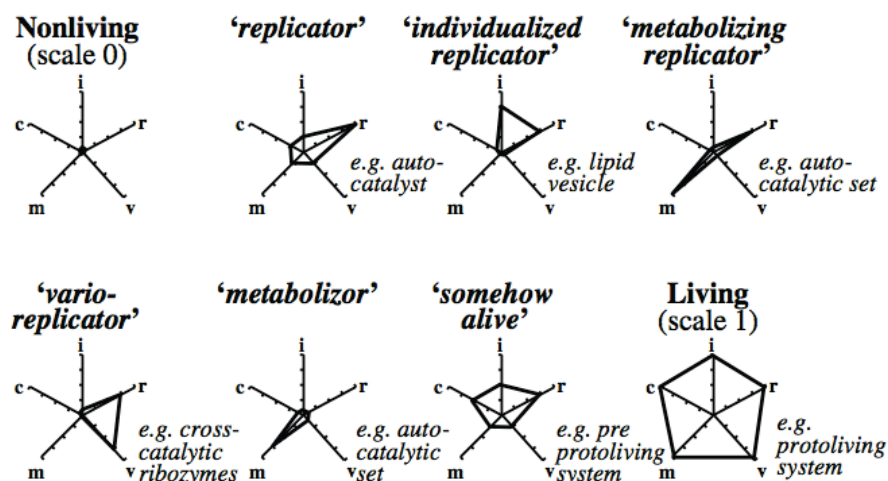


Figure 3 : Système de degrés de vie, du non-vivant au vivant.

Les degrés du vivant sont représentés par un polygone représentant des scores de performance du système le long d'un axe représentant une fonction. Les axes représentent : « i » l'individualisation, « r » la réplication, « v » la variation, « m » le métabolisme, « c » l'accouplement. Chaque échelle va du centre (valeur 0) à la périphérie (valeur 1). Modifié de Malaterre (2010).

Dans un contexte antagoniste (vivant/non vivant) strict, les virus sont dans leur apparente « simplicité » assimilés au monde du non vivant. En effet, les virus ne possèdent pas d'outil leur permettant de se répliquer seul. Ils n'ont ni mitochondrie pour leur production d'énergie, ni ribosome leur permettant de synthétiser leurs protéines ainsi qu'une absence d'activité métabolique (Luisi, 1998 ; Ruiz-Mirazo *et al.*, 2004). Leur génome ne contiendrait donc que le « manuel » pour permettre leur réplication.

Cependant, si l'on inclut des degrés de vie comme l'a proposé Malaterre (2010), les virus peuvent être assimilés au « somehow alive ». En effet, certains auteurs proposent une vision plus permissive en posant l'hypothèse que le virion sous sa forme primaire ne serait qu'une partie du cycle du virus, assimilant ainsi les virions à des « gamètes ». Ainsi la seconde phase appelée « virocell » représentant l'étape d'infection d'une cellule par un virus serait la partie du cycle où le virus serait « vivant » (Claverie et Abergel, 2010; Forterre, 2010). En effet, lors de ce passage, le virus agit de manière autonome et donc obéit aux définitions précédemment citées du vivant. Ils auraient ainsi les propriétés des entités « somehow alive » puisqu'ils sont capables d'individualisation, et par le couplage avec leur hôte, de « reproduction » (lysogénie) et de processus métaboliques ou de réplifications.

Dans un contexte vivant/non vivant, seuls les organismes strictement vivants peuvent être parasités. Or, dans cette vision manichéenne, comment peut-on définir le virus Sputnik (La Scola *et al.*, 2008), un virophage infectant le virus Mimivirus ? Pearson a tranché dans un article intitulé « *virophage* » *suggests viruses are alive*. (Pearson, 2008). Ces différents constats n'ont valeur que d'hypothèses controversées actuellement puisque ces deux visions (vivant/non-vivant) font appel à des arguments d'importance dépendant du concept même de la vie (voir Moreira et López-García, 2009). Toutefois, les avancées permanentes de la science vont permettre la sélection de ces théories et bien sûr permettront l'émergence de nouvelles théories, qui seront elles aussi sujettes à débat.

3). Les virus aquatiques

La première observation dans les années 50 (Spencer, 1955) et la confirmation de leur prédominance numérique dans les années 90 (Bergh *et al.*, 1989; Proctor et Fuhrman, 1990), ont conduit à considérer les virus marins comme un compartiment important, appelé le virioplancton. Le plancton, du grec « *planktos* » : errant, désigne les organismes aquatiques passivement en suspension dans l'eau et inaptes à lutter contre les courants (Hensen, 1887). La majorité des organismes de ce compartiment est microscopique mais représentent à eux seuls une part extrêmement importante de la biomasse (Figure 4).

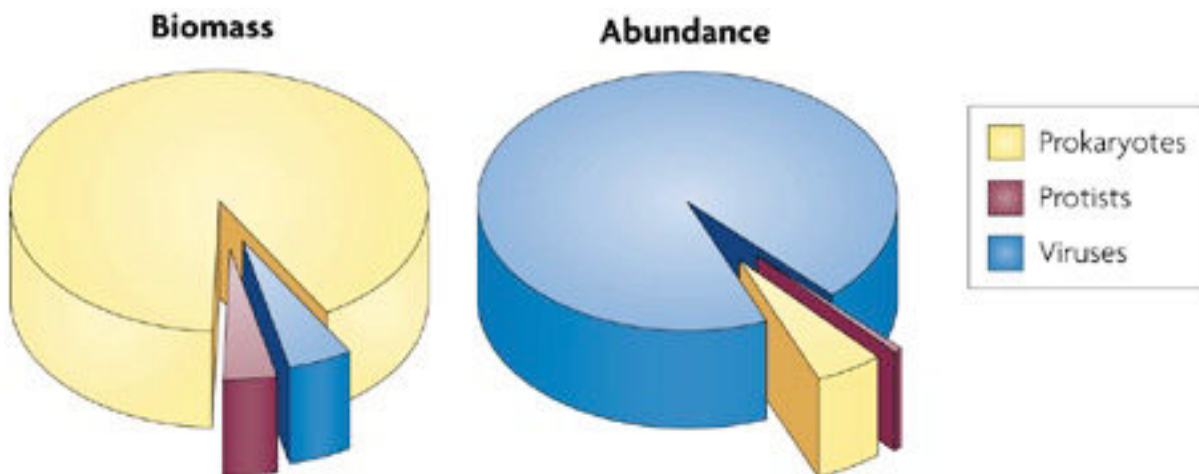


Figure 4 : Proportion en termes de biomasse et d'abondance des procaryotes, eucaryotes et virus dans les océans. Adaptée de Suttle (2007).

Bien que l'essentiel de cette biomasse soit principalement imputable aux bactéries, et que numériquement celles-ci sont 10 fois supérieures aux organismes eucaryotes, les virus sont dix fois plus abondants. Leur concentration est estimée entre 10^6 et 10^8 virus.mL⁻¹ (Suttle, 2007). Cette abondance virale est pourtant approximativement 100 fois moins importante que dans les sédiments où la concentration virale peut atteindre environ 10^{10} virus.g⁻¹ de sédiment (Danovaro et Serresi, 2000). Ces mesures ont permis une estimation de leur nombre total à 10^{31} virus dans l'ensemble des océans (Suttle, 2005; Wommack et Colwell, 2000) représentant approximativement 94 % des acides nucléiques retrouvés dans les entités biologiques (Suttle, 2007). Le nombre total de

virus a toutefois été ramené à 10^{30} , dû à la présence de microvesicules induisant une surestimation de l'abondance des particules virales (Biller *et al.*, 2014 ; Forterre *et al.*, 2013). Les virus restent tout de même les entités biologiques les plus abondantes sur terre, et nous savons aujourd'hui que 99 % d'entre eux infectent des organismes unicellulaires. Les entités virales sont retrouvées en abondance dans tous les types de milieu aquatique, à l'exception des lacs de haute montagne ou quelques lacs polaires (Sawstrom *et al.*, 2007 ; Hofer et Sommaruga, 2001) Les virus à ADN simple ou double brins peuvent être détectés dans tous les types d'hôtes (des bactéries aux microalgues) (Rosario *et al.*, 2010). L'utilisation de la microscopie électronique à transmission a montré, outre leur importante abondance, les virus varient énormément en taille et en forme (Flint *et al.*, 2000). Dans les environnements aquatiques, la plupart des virus possèdent une taille comprise entre 30 et 60 nm (Sandaa, 2008), car ils sont majoritairement composés de bactériophages. Cet aspect est lié au fait que les bactéries sont les organismes vivants les plus abondants dans les milieux aquatiques (Rohwer et Edwards, 2002). La diversité virale a été largement étudiée par PFGE (*Pulsed Field Gel Electrophoresis* pour électrophorèse en champs pulsé) que ce soit en mer (Boujelben *et al.*, 2012 ; Ory *et al.*, 2011 ; Filippini et Middelboe, 2007), en estuaire (August *et al.*, 2006 ; Jamindar *et al.*, 2012, Filippini et Middelboe, 2007), en rivière (August *et al.*, 2006), en zone lacustre (Filippini et Middelboe, 2007 ; Tijdens *et al.*, 2008 ; Parvathi *et al.*, 2012) ainsi que dans les sédiments (Filippini et Middelboe, 2007 ; Leroy *et al.*, 2008). Ces études ont révélé un assemblage viral naturel comprenant de 7 à 35 tailles distinctes de génome, entre 12 et 560 kb (Auguet *et al.*, 2006, 2005; Larsen *et al.*, 2001; Steward, 2001). Malgré cette prédominance des bactériophages, certaines études utilisant la méthode du MPN (*Most Probable Number*) montrent que les virus infectant les eucaryotes photosynthétiques peuvent être eux-aussi abondants (supérieur à 10^5 virus par mL^{-1}), avec une importante diversité (Waterbury et Valois, 1993 ; Suttle et Chan, 1994, Cottrel et Suttle, 1995a ; Tomaru *et al.*, 2004).

D'autres études ont permis de mettre en évidence l'augmentation du nombre de virus lors d'une efflorescence phytoplanctonique (Bratbak *et al.*, 1990 ; Castberg *et al.*, 2001), ainsi que les variations de cette augmentation le long d'un gradient trophique (Noble et Fuhrman, 2000a ; Weinbauer *et al.*, 1993) ou bathymétrique (Clasen *et al.*, 2008 ; Culley et Welschmeyer, 2002 ; Hara *et al.*, 1996). Ces études ont montré que les concentrations virales mesurées sont plus fortes en zone eutrophe qu'en zone oligotrophe. De plus, un

contraste important est observé au niveau des différentes limites physico-chimiques présentes dans l'océan comme au niveau de la thermocline (Cochlan *et al.*, 1993 ; Hwang et Cho, 2002 ; Riemann et Middelboe, 2002 ; Weinbauer *et al.*, 1995), de la halocline (Winter *et al.*, 2008), et de la chemocline (i.e. entre eaux oxiques/anoxiques ; Taylor *et al.*, 2003, 2001).

Les ratios du nombre virus/procaryotes révèlent aussi des différences entre environnements. En effet, les eaux de surface des océans Pacifique et Arctique présentent des ratios compris entre 10 et 40 (Clasen *et al.*, 2008), mais les eaux profondes de l'Atlantique montrent un ratio excédant souvent 100 (Parada *et al.*, 2007). L'hypothèse avancée par les auteurs est que le taux de conservation des virions dans les eaux froides et obscures serait supérieur (Clasen *et al.*, 2008). Dans les environnements froids, les virions peuvent être préservés durant de nombreuses années comme en témoigne la découverte du Pandoravirus dans des sédiments datés de 30000 ans (Philippe *et al.*, 2013).

Ces caractéristiques numériques couplées à leur mode de réplication confèrent au virioplancton un rôle particulièrement important au sein des écosystèmes (voir la revue de Jacquet *et al.*, 2010), avec comme dans tout système biologique une forte dynamique spatio-temporelle. Toutefois, malgré un développement important des connaissances sur les virus marins, de nombreux points restent mal compris, voire inconnus. Nous pouvons donc supposer une sous-estimation de leur importance dans les écosystèmes marins. Cependant trois aspects peuvent être résumés de manière simple, 1) leur impact sur les cycles biogéochimiques, 2) leur impact sur les communautés par l'évolution des communautés d'organismes et 3) le stockage/préservation du matériel génétique par la séquestration des virus dans les sédiments.

III Influence des virus sur l'environnement

1) Influence des virus sur les cycles biogéochimiques

C'est le biologiste Charles Elton qui en 1927, mit les réseaux trophiques en évidence. Au sein des écosystèmes marins, le phytoplancton, composé des eucaryotes et procaryotes photosynthétiques, permet l'entrée de la matière organique dans le système. Cette biomasse nouvellement créée par les producteurs primaires va permettre aux échelons trophiques supérieurs (consommateurs) de se développer. Toutefois à chaque échelon trophique, une perte de matière organique dissoute et particulaire (MOD-P) a lieu. Cette matière organique ne peut pas (dans sa grande majorité) être réintégrée directement dans la chaîne trophique. La présence de décomposeurs, représentant le dernier maillon de la chaîne, est alors indispensable. Les bactéries assurent cette fonction par le phénomène de « boucle microbienne » (Azam, 1983). Les bactéries hétérotrophes qui participent à cette boucle vont alors reminéraliser la MOD-P. Ce processus va permettre à cette MOD-P de retourner dans la chaîne trophique par la consommation des ces bactéries mais va aussi permettre le retour de nutriments nécessaires aux producteurs primaires.

Par la lyse de leur hôte, les virus influencent le cycle de la matière organique (Figure 5) puisqu'ils convertissent la biomasse vivante en MOD (Bratbak *et al.*, 1990 ; Fuhrman, 1992 ; Proctor, 1991 ; Wilhelm et Suttle, 1999, Figure 5). Les virus détournent donc la biomasse nouvellement produite et destinée aux niveaux trophiques supérieurs et stimulent, via la production de MOD, la croissance du compartiment bactérien ainsi que le recyclage de la matière organique (Brussaard *et al.*, 1996 ; Middelboe *et al.*, 2003, 1996). La lyse virale pousse donc le système à fonctionner en un cycle plus « régénéré ».

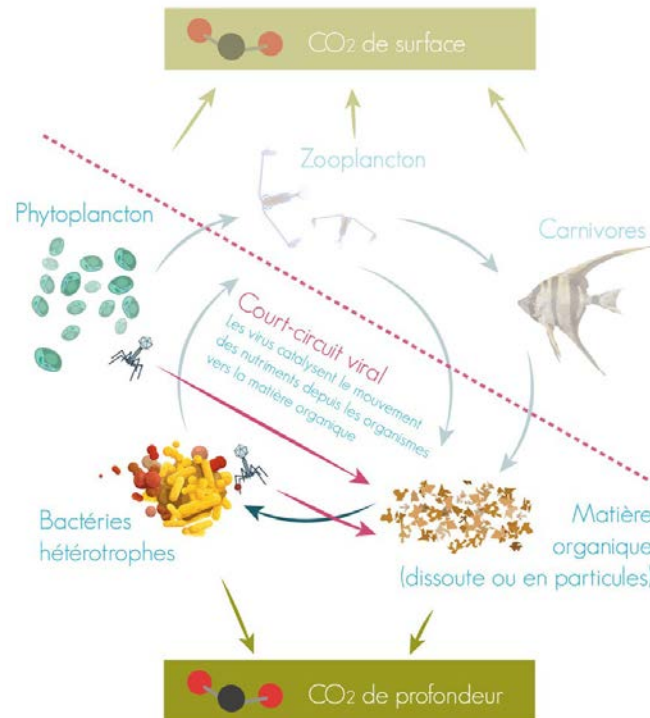


Figure 5 : Schéma de la boucle microbienne, intégrant le court-circuit viral. Adaptée de Suttle (2005) par e-marin'Lab (www.emarinlab.obs-banyuls.fr).

L'utilisation d'un modèle mathématique a permis à Wilhelm et Suttle (1999) de démontrer que 2 à 26 % du carbone organique fixé lors de la photosynthèse est recyclé en matière organique dissoute grâce à la lyse virale. Lors de la modélisation d'un bloom de *Phaeocystis globosa*, basée sur des données empiriques, Ruardij *et al.* (2005) ont démontré que la mortalité des microalgues par les virus constituait un facteur de régulation essentiel dans le cycle des nutriments. L'observation par microscopie électronique à transmission a également révélé qu'une proportion importante (10-50%) des cellules algales étaient clairement infectées à la fin d'une efflorescence d'*Aureococcus anophagefferens* (Sieburth *et al.*, 1988 ; Gastrich *et al.*, 2004), *Heterosigma akashiwo* (Nagasaki *et al.*, 1994) et *Emiliania huxleyi* (Bratbak *et al.*, 1993, 1996 ; Brussaard *et al.*, 1996b). D'autres approches, déterminant le taux des lyses cellulaires à partir du nombre de virus produits et en le divisant pour la *burst size* estimée empiriquement, suggèrent que les virus sont responsable d'une part importante de la mortalité (7 - 100%) durant des efflorescences de *Phaeocystis globosa* (Brussaard *et al.*, 2005a ; Ruardij *et al.*, 2005) et *Emiliania huxleyi* (Jacquet *et al.*, 2002).

Ainsi le compartiment viral, par la libération de matière organique dissoute, provoque un apport significatif de nutriments comme le carbone, l'azote, le phosphore et le fer, pouvant favoriser d'autres microorganismes autotrophes ou hétérotrophes (Gobler *et al.*, 1997 ; Middelboe *et al.*, 2003, 1996 ; Poorvin *et al.*, 2004).

Ce « court-circuit viral » aurait donc un effet significatif sur les communautés aquatiques en induisant une modification de la composition des éléments nutritifs dans l'environnement.

2) Implication des virus sur la composition des communautés phytoplanctoniques

Outre cette implication dans le contrôle des cycles de la matière, les virus jouent un rôle direct sur les communautés planctoniques. Ils peuvent avoir un impact majeur sur la dynamique du phytoplancton, et la lyse virale peut représenter 75 % des pertes cellulaires lors de la terminaison d'un bloom de *Phaeocystis pouchetti* ou *Emiliana huxleyi* (Bratbak *et al.*, 1993a; Brussaard, 2005; Brussaard *et al.*, 1995; Martínez *et al.*, 2007; Nagasaki *et al.*, 1994; Tarutani *et al.*, 2000; Van Boekel, 1992). La majorité des virus marins ont un spectre d'hôtes réduit (Clerissi *et al.*, 2012), ainsi seule une partie de la communauté planctonique serait affectée par un virus donné. L'une des théories sur cette sélection spécifique d'hôtes est appelée « killing the winner ». Suivant ce principe, on admet que l'hôte le plus abondant dans la communauté planctonique sera l'organisme le plus infecté et donc lysé. Cette théorie est assez similaire aux modèles classiques proie-prédateur (modèle Lotka-Volterra), montrant une réponse densité-dépendante, à ceci près qu'elle considère plus spécifiquement une fraction particulière des souches d'hôtes présentes dans une population cible. Ainsi par ce processus les virus maintiennent la compétition intraspécifique en régulant la souche la plus abondante (Thingstad, 2000) et en permettant aux autres membres de la communauté de se développer.

La pression des virus sur leurs hôtes ne s'arrête pas au niveau de l'espèce. En effet les virus ont un effet sur la composition clonale de la communauté (Nagasaki *et al.*, 1999;

Tomaru *et al.*, 2004), comme montré chez des populations d'*Heterosigma akashiwo* (Tarutani *et al.*, 2000). Les virus peuvent modifier la composition clonale de leurs hôtes par la sélection de souches d'hôtes résistantes, phénomène qui a été décrit chez les cyanobactéries (Waterbury et Valois, 1993) mais aussi chez les picoeucaryotes (Thomas *et al.*, 2011b).

Cependant, la résistance des hôtes aux virus représente généralement un coût métabolique supplémentaire (Thomas *et al.*, 2011) et ne semble donc pas fixée au sein des communautés d'hôtes, car désavantageuse pour la compétition. Ainsi en l'absence de pression de sélection, l'abondance des souches d'hôtes résistantes va tendre à diminuer en faveur des souches plus compétitives.

En revanche, certaines espèces ont des mécanismes réversibles provoquant une résistance face aux attaques virales. Une étude sur des populations d'*Emiliania huxleyi* a montré que les virus ont un impact sur le phénotype habituel de l'hôte, correspondant aux individus diploïdes. Le passage à une population haploïde rend les individus insensibles aux infections virales. La pression de sélection ne peut alors survenir que sur la fraction diploïde restante (stratégie du « Cheshire cat », Frada *et al.*, 2008, vue artistique Figure 6).

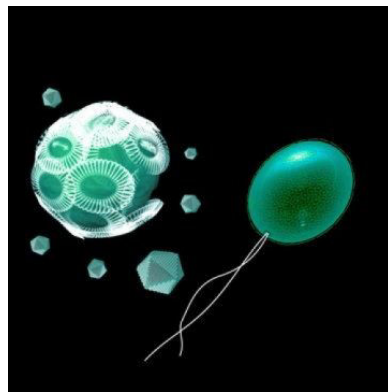


Figure 6 : Illustration des différents phénotypes d'*Emiliania huxleyi* en réponse à une attaque virale. Phase diploïde sensible au virus (gauche), phase haploïde résistante au virus (droite), Miguel Frada.

Les virus ont aussi la capacité d'opérer des transferts horizontaux de gènes, c'est-à-dire de capter du matériel génétique d'une cellule hôte en l'intégrant à leur génome, soit un transfert d'hôte à virus. Cette transmission est aussi possible dans le sens inverse, lors de la résistance de l'hôte par intégration du génome viral. Par extension, la transmission entre virus semble aussi tout à fait probable. Ainsi, ils peuvent avoir un rôle dans le maintien

de la diversité génétique des populations, à travers l'hétérogénéisation des gènes dans une population hôte donnée, et peuvent même participer à sa diversification (Ackermann, 2001). De plus, la séquestration des virus dans les eaux océaniques profondes mais aussi au sein des sédiments suggère la présence d'un réservoir de gènes viraux. Ces gènes pourraient avoir disparu au cours de l'évolution et revenir au sein des communautés par des processus de transduction (Thurber *et al.*, 2009) à travers des phénomènes de remise en suspension (*upwelling*).

L'effet des virus sur l'évolution de leur hôte ne s'arrête pas à la diversification des populations. Ils favorisent ce que l'on a nommé « la course aux armements » (Van Valen, 1973) ou « l'hypothèse de la reine rouge ». En effet, cette hypothèse propose que « la sélection naturelle favorise les prédateurs les plus rapides, elle favorise aussi les proies les plus rapides, ce qui aurait pour résultat un rapport de forces inchangé ». Pour simplifier, ce concept stipule que la coévolution de deux espèces antagonistes devrait accélérer leur évolution moléculaire grâce aux innovations constantes chez ces mêmes espèces destinées à les maintenir à l'état d'équilibre l'une vis-à-vis de l'autre.

3) Impact des virus sur le climat

Bien que les virus soient des entités biologiques infinitésimales, leur impact peut influencer des processus à plus larges échelles. Malin *et al.* en 1998 ont montré que les virus peuvent avoir un impact sur le climat et en particulier sur l'effet de serre. Le diméthyl sulfide (DMS) va après oxydation dans l'atmosphère, avoir une influence sur l'absorption des radiations dans les nuages, et ainsi avoir une influence sur la retenue de température dans l'atmosphère (Malin *et al.*, 1992). Le diméthylsulfoniopropionate (DMSP), l'un des principaux précurseurs du DMS ainsi que ce dernier, sont produits naturellement par les communautés phytoplanctoniques. Malin *et al.* (1992) ont montré grâce à des expériences en mésocosme qu'une infection virale sur une culture de l'algue *Phaeocystis pouchetii* augmente par 8 fois la quantité de DMS. La lyse virale pourrait donc avoir un rôle important sur l'environnement à différentes échelles. Cependant, dans le sens inverse, quel pourrait être l'impact global de l'environnement sur les communautés virales ?

IV Influence de l'environnement sur les virus

Il est particulièrement délicat de discuter l'impact de l'environnement sur un système hôte-parasite, et plus encore lorsque le parasite présente une activité métabolique nulle jusqu'à la rencontre avec l'hôte. A l'exception de certains facteurs, il est difficile de faire la distinction entre un effet direct de ces variables sur le virion lui-même ou un effet indirect par l'intermédiaire de son hôte. Cependant, depuis la première culture de virus infectant une algue eucaryote, *Micromonas pusilla*, par J. A. Mayer et Taylor (1979), des éléments supportant l'impact de l'environnement sur les communautés virales se sont accumulés. La plupart de ces observations proviennent d'études en laboratoire ou en mésocosme sur un système hôte-virus déterminé. Les études réalisées en milieu naturel ne tiennent principalement compte que de l'abondance de la communauté virale.

Lors de ces études visant à quantifier l'abondance virale, les méthodes employées ne permettaient pas de distinguer les groupes viraux. En revanche, elles nous fournissent des informations sur les dynamiques spatio-temporelles des populations et ont également permis de montrer que si l'environnement est favorable, les populations virales sont généralement assez stables (Clokie *et al.*, 2011).

Le développement de méthodes indépendantes de la culture d'hôte a permis d'aborder la diversité virale *in situ*, par exemple, l'électrophorèse en champs pulsé permet de discriminer les génomes suivant leur taille (Auguet *et al.*, 2006, 2005; Larsen *et al.*, 2001; Steward, 2001). Cette technique a permis de confirmer les dynamiques spatio-temporelles précédemment montrées par des études basées sur des techniques de microscopie électronique et en épifluorescence. Ainsi, les dynamiques des communautés virales lors d'un bloom phytoplanctonique (Castberg *et al.*, 2001; Larsen *et al.*, 2001), ou lors de la stratification de la colonne d'eau (Jiang *et al.*, 2003; Wommack *et al.*, 1999) ont pu être confirmées.

Les progrès techniques et scientifiques pour décrire la diversité virale au sein de communautés naturelles par l'utilisation de la métagénomique ont permis de mettre en évidence entre 1000 et plus d'un million de génotypes différents dans des échantillons d'eau côtière (Breitbart *et al.*, 2002).

Des auteurs s'intéressant aux communautés de viriophages ont montré l'impact de facteurs abiotiques sur ces virus. Par exemple, il semble que les stratégies lysogéniques soient dominantes dans les eaux anoxiques, supposant ainsi que l'**oxygène** aurait un impact sur les stratégies virales (Muck *et al.*, 2014). Ceci est cohérent avec une étude révélant que, dans un système oligotrophique marin en condition d'anoxie, les infections virales sont relativement faibles (Taylor *et al.*, 2003). Paradoxalement, les plus fortes concentrations et productions virales ont été rapportées dans les sites marins anoxiques, mais il semblerait aussi que la destruction virale y soit la plus forte, démontrant ainsi un *turn over* beaucoup plus important (Corinaldesi *et al.*, 2007a, 2007b).

La **température** a également un impact sur les communautés virales mais aussi sur leur influence dans l'environnement. Rowe *et al.*, (2012) ont montré une corrélation entre l'abondance virale et la température, illustrée par le fait que dans les eaux froides, la mortalité des communautés est influencée par les brouteurs, alors que dans les eaux plus chaudes, la lyse virale serait le mécanisme dominant (Mojica et Brussaard, 2014). Dans les eaux plus chaudes cependant, l'efficacité de croissance des organismes est réduite, diminuant ainsi la production virale (Danovaro *et al.*, 2011). Cela suggère un équilibre entre la conservation des virions avec une faible production dans les eaux froides, et une activité lytique plus forte mais une durée de vie des virus plus faible dans les eaux chaudes (Nagasaki et Yamaguchi, 1998 ; Wells et Deming, 2006). L'augmentation de la **température** serait en effet associée à une augmentation exponentielle de perte virale dans la colonne d'eau (Danovaro et Middelboe, 2010). Cette destruction virale liée à l'augmentation de la température peut s'expliquer par une stimulation des activités exoenzymatiques qui seraient largement responsables de la dégradation de la capsid virale (Corinaldesi *et al.*, 2010). En revanche, les communautés de virus présentes dans les sédiments ne réagissent pas de la même manière à une augmentation de la température. L'hypothèse serait que la matrice sédimentaire pourrait jouer le rôle d'une « gangue protectrice » et ainsi limiter l'action des exoenzymes (Corinaldesi *et al.*, 2010). Ces différents effets pourraient expliquer l'abondance plus importante des virus dans les eaux froides et dans les sédiments (Danovaro *et al.*, 2011).

Comme dans tout système hôte-parasite, l'abondance virale est liée à l'**abondance des hôtes**, représentant un lien densité-dépendant (Danovaro *et al.*, 2011), mais aussi à leur **productivité**. Une production plus importante (phase exponentielle de croissance) des cellules hôtes aurait pour conséquence une augmentation du nombre de virus produits (*burst size*) et une diminution de la période de latence (Clasen *et al.*, 2008 ; Wilhelm et Matteson, 2008). L'abondance des hôtes, principalement les organismes photosynthétiques, sont dépendant de la **luminosité** pour leur croissance. Les virus infectant des organismes photosynthétiques semblent eux aussi avoir besoin de l'activité photosynthétique pour leur réplication. Adolph et Haselkorn (1971) ont montré que l'inhibition de la photosynthèse chez *Nostoc muscorum* lors d'une infection par le cyanophage BGA N-1 réduit de 98 % la production virale. Il a aussi été montré chez un virus de microalgue (MpV-SP1) qu'une activité photosynthétique est nécessaire pour le déclenchement de la lyse virale (Brown *et al.*, 2007). En dépit de cela, pour les virus des organismes photosynthétiques, il a été montré que les UV sont responsables d'une dégradation importante des communautés virales (Noble et Fuhrman, 1997 ; Wilhelm *et al.*, 1998 ; Jeffrey *et al.*, 2000), principalement à travers les dommages à l'ADN (Suttle, 2000). En outre, ces dommages resteraient significatifs jusqu'à 10 m de profondeur (Wilhmen *et al.*, 2002). Néanmoins, même si cette composante lumineuse s'avère néfaste pour les communautés virales, une étude portant sur un virus de *Micromonas* a montré qu'une activité photosynthétique de l'hôte est nécessaire pour le déclenchement de la lyse virale (Brown *et al.*, 2007). L'abondance des hôtes est un facteur essentiel dans le développement des communautés virales, mais le développement et la diversification des communautés d'organismes peut être un facteur limitant pour les populations virales. En effet, ils peuvent être phagocytés et donc dégradés par **les nanoflagellés** dits « viriovores » (Suttle et chen, 1992 ; Gonzales et Suttle, 1993 ; Manage *et al.*, 2002 ; Bettarel *et al.*, 2005). **La salinité** joue aussi un rôle. Malgré le fait que les populations virales semblent être génétiquement distinctes (Clasen et Suttle, 2009 ; Sano *et al.*, 2004), les *burst sizes* moyens chez les bactériophages sont moins forts dans les eaux marines que dans les eaux douces, respectivement 20-25 et 28-40 (Parada *et al.*, 2006 ; Wommack et Colwell, 2000).

Le rôle direct de la salinité et du **pH** n'est pas encore clair (Danovaro *et al.*, 2011). La littérature révèle que l'abondance virale est plus forte au sein des environnements de basse salinité. Les auteurs mettent l'accent sur le fait que la plupart des données

analysées proviennent d'échantillonnages dans des estuaires et donc davantage liées à des caractéristiques de ces zones. En effet, ces estuaires sont des régions particulières, riches en matière organique et donc caractérisées par des communautés d'hôtes à la fois plus nombreux et plus productifs. (Figure 7).

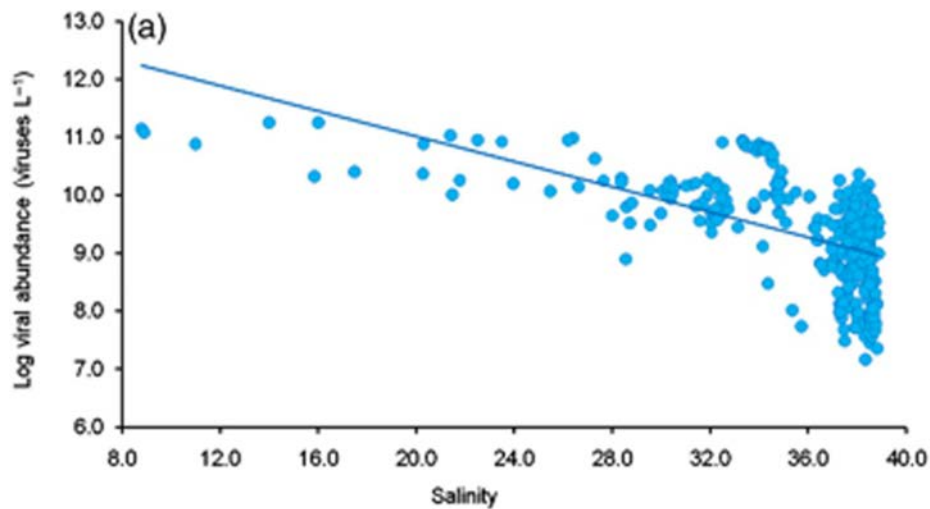


Figure 7 : Relation entre la salinité et l'abondance virale dans les eaux de surface (Danovaro *et al.* 2011)

D'autre part, certains aspects en relation avec la **salinité et le pH** peuvent être discutés quant à leur rôle dans l'élimination des virions. Un changement de salinité et de pH influence l'adsorption (l'attachement même temporaire des virus) en modifiant les forces électro-statiques entre les virus et les particules du milieu (Harvey et Ryan, 2004). Cette propriété d'adsorption se renforce généralement avec une augmentation du pH (Charles, 1984 ; Grant *et al.*, 1993) ou des sels dissous (Sobsey *et al.*, 1975). Ainsi par ce phénomène, il a été montré une inactivation, même temporaire, de l'infectiosité (Brussaard *et al.*, 2009), mais aussi une sédimentation plus importante des virus (Proctor et Furhman, 1991). En revanche cet effet ne s'applique pas uniformément aux virus, et les sensibilités virales au pH présentent de grandes variations. Les phages d'eau douce y sont sensibles mais les cyanophages marins semblent simplement affectés par les valeurs de pH et non par les conditions d'adsorptions (Weinbauer, 2004). Les cyanophages d'eau douce en revanche semblent montrer une tolérance plus large au pH (5-11) que les bactériophages en général (Suttle, 2002). Très peu d'études sur les virus d'hôtes eucaryotes ont été conduites mais une perte d'infection a été montrée pour un virus infectant *Micromonas pusilla* à pH 5 (Brussaard *et al.*, 2005). Certains virus

semblent insensibles à un pH faible (3, voir moins) (Jin *et al.*, 2005 ; Krueger et Fong, 1937), ce qui rend difficile la compréhension des mécanismes expliquant l'influence du pH sur ces communautés, résultant peut-être de mécanismes particuliers d'adaptation (Rueckert, 1996).

V Le système hôte-virus utilisé : Mamiellophyceae-Prasinovirus

1) Les Mamiellophyceae

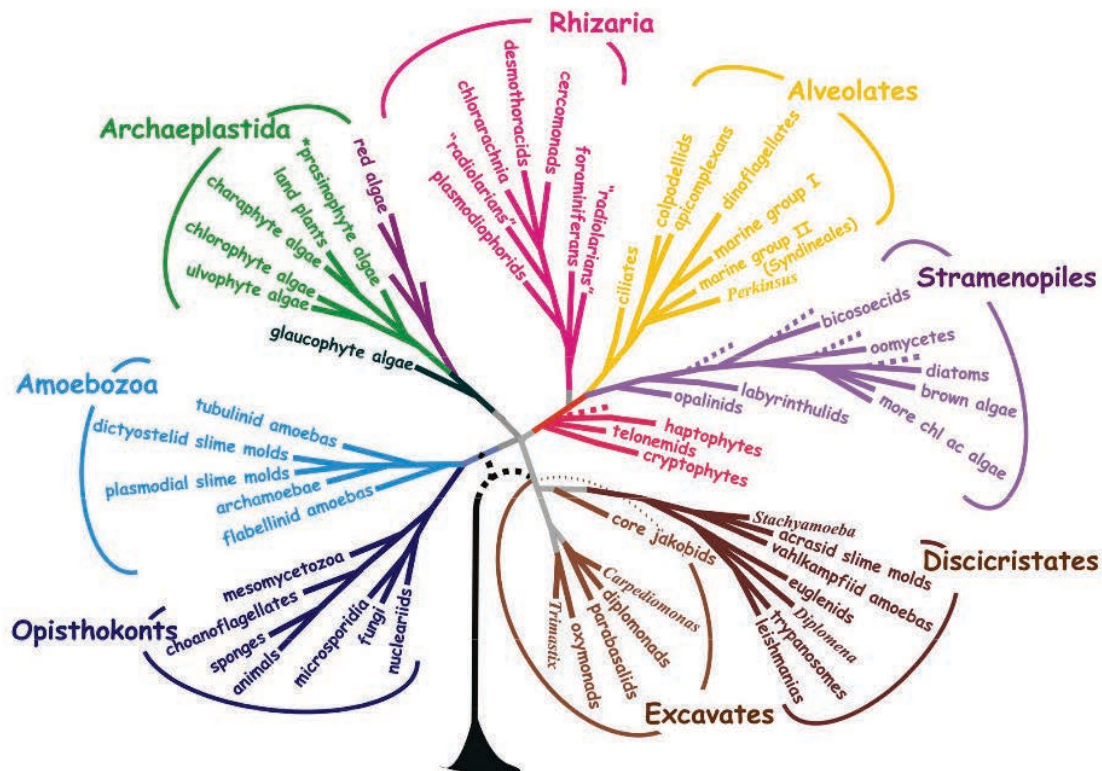


Figure 8 : Phylogénie consensuelle des principaux groupes d'eucaryotes, * le groupe des Mamiellophyceae est inclus dans « Prasinophyte Algae », adaptée de Baldauf (2008)

La classe des Mamiellophyceae, incluse à l'origine dans le groupe paraphylétique des Prasinophyceae (Marin et Melkonian, 2010a), est composée d'organismes photosynthétiques unicellulaires situés à la base de la lignée verte (Baldauf, 2008, Figure 8).

Elle contient trois ordres (Figure 9), les Monomastigales (inféodés au milieu dulçaquicole), les Dolichomastigales et les Mamiellales (tous deux en milieu marin, (Marin et Melkonian, 2010)) qui sont classés suivant leur taille dans le pico- (<2 μm) et nanoplancton (2-20 μm).

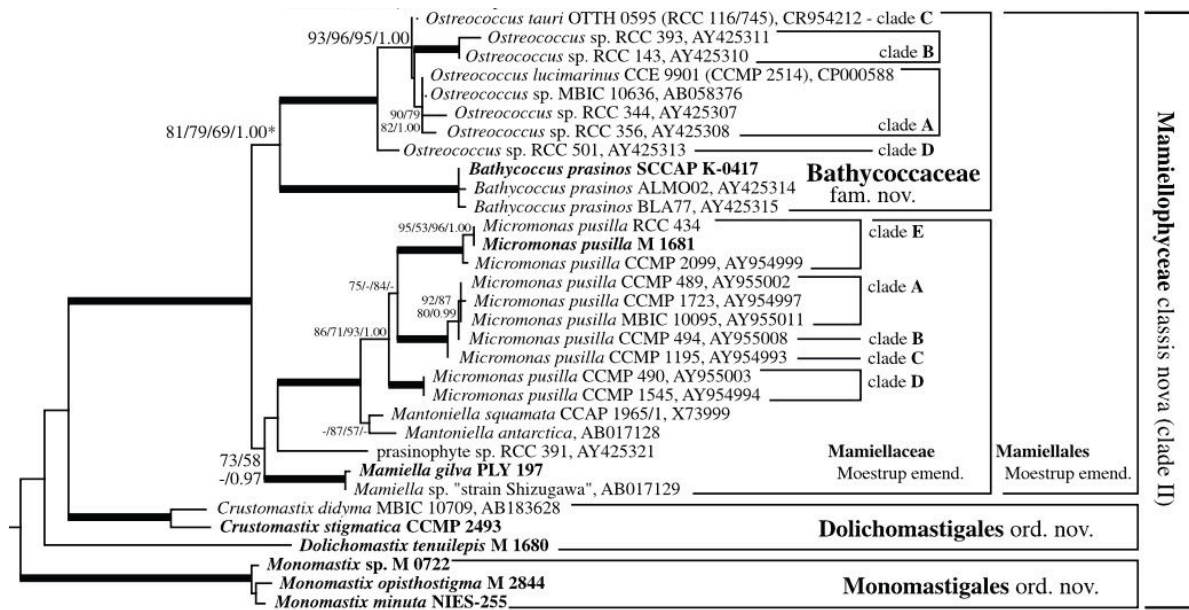


Figure 9 : Phylogénie moléculaire des Mamiellophyceae basée sur le « gène » ribosomal 18S. Adaptée de Marin et Melkonian (2010).

Les deux genres composant les Dolichomastigales (*Crustomastix* et *Dolichomastix*) montrent une abondance relative plus importante en milieu oligotrophe alors que les membres des Mamiellales dominent généralement la fraction picoeucaryotique des différents milieux (Viprey *et al.*, 2008 ; Wu *et al.*, 2013 ; Zhu *et al.*, 2005). Les Mamiellophyceae picoeucaryotiques appartiennent à trois genres dominants (tous regroupés dans l'ordre des Mamiellales). Ce sont les genres *Micromonas* (Bucher, 1952; Manton et Parke, 1960), *Bathycoccus* (Eikrem et Throndsen, 1990) et *Ostreococcus* (Courties *et al.*, 1994).

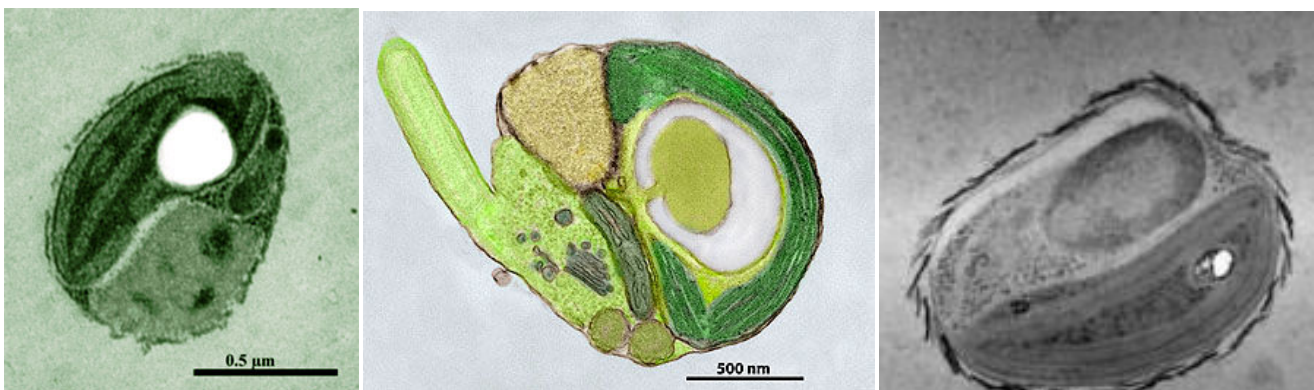


Figure 10 : Les trois genres dominants de Mamiellophyceae. Les barres représentent 500 nm. A. *Ostreococcus* : <http://genome.jgi.doe.gov> ; B. *Micromonas* : <http://sb-roscoff.fr/> ; C. *Bathycoccus* : <http://bioinformatics.psb.ugent.be/plaza/versions/pico-plaza/organism/view/Bathycoccus+prasinos>

Ces trois genres (Figure 10) et en particulier *Ostreococcus* sont considérés comme les plus petits eucaryotes photosynthétiques connus avec une taille inférieure à 2 µm (Courties *et al.*, 1994). Leur structure cellulaire est simple avec des organites en exemplaire unique pour le chloroplaste, la mitochondrie ainsi que l'appareil de Golgi (Derelle *et al.*, 2006 ; Worden *et al.*, 2009). En revanche, des différences morphologiques sont visibles au sein de cette classe, avec des cellules nues ou pourvues d'écailles, et la présence d'un, deux ou aucun flagelle(s).

Le genre *Micromonas* a été le premier picoeucaryote planctonique décrit, avec l'espèce *Micromonas pusilla* (Manton et Parke, 1960), et est pourvu d'une seule flagelle. Malgré la description d'une seule espèce, ce genre semble être un complexe d'espèces cryptiques (Šlapeta *et al.*, 2006 ; Simon *et al.*, in prep.). Marin et Melkonian (2010) ont défini cinq clades en se basant sur l'analyse des séquences de l'ADNr 18S. Le genre *Micromonas* semble avoir divergé au sein des Mamiellophyceae il y a environ 65 millions d'années (Šlapeta *et al.*, 2006), divergence qui aurait induit l'apparition d'un nombre important de souches à travers une radiation adaptative favorisant ainsi l'ubiquité de ce genre. Ces adaptations n'ont cependant pas abouti à une spéciation claire, montrée par une faible divergence entre les gènes ribosomiques de deux souches provenant de l'Atlantique Nord (CCMP1545) et du Sud-Ouest Pacifique (RCC 299). Malgré cette similarité ribosomique forte, 10% des gènes de leur génome leur sont spécifiques (Worden *et al.*, 2009), suggérant une plasticité génique pouvant être attribuée à leur adaptation environnementale mais suggérant également que ce genre serait potentiellement un complexe d'espèce cryptique. Quoi qu'il en soit, *Micromonas* est un genre particulièrement ubiquiste et dominant les picoeucaryotes dans les régions riches en nutriments (Viprey *et al.*, 2008 ; Wu *et al.*, 2013) comme la Manche (Not *et al.*, 2004).

Les deux autres genres dominants, *Bathycoccus* et *Ostreococcus*, sont exempts de flagelle. En revanche *Bathycoccus* est caractérisé par la présence d'écailles autour de la cellule. Malgré son abondance, les études du genre *Bathycoccus* sont très peu nombreuses et à notre connaissance, seulement cinq études portent sur sa distribution. Johnson et Sieburth en 1982, avant la description de ce genre par Eikrem et Throndsen (1990), ont décrit des cellules morphologiquement similaires à *Bathycoccus*, particulièrement abondantes dans des échantillons provenant de l'Atlantique Nord. En 2008, Viprey *et al.*, ont montré qu'il était souvent observé en Méditerranée et Vaultot *et al.* (2012) qu'il était

présent au niveau de l'upwelling du Chili. Enfin, Wu *et al.* (2014) ont montré grâce à un transect réalisé en Mer de Chine que ce genre était particulièrement abondant vers 100 m de profondeur dans une zone oligotrophe. Malgré le faible nombre d'études sur ce genre, l'analyse de l'ensemble des séquences génétiques disponibles indique une faible divergence par rapport aux génomes de référence disponibles (Moreau *et al.*, 2012 ; Piganeau *et al.*, 2011).

Ostreococcus est un genre particulièrement ubiquiste mais découvert très récemment, dans l'étang de Thau en 1994 par Courties *et al.*. La première espèce décrite, *Ostreococcus tauri*, a été nommée ainsi car elle a été isolée dans l'étang de Thau (« *tauri* »), qui est un environnement riche en culture ostréicole (« *Ostreo* ») et ressemblant à une perle (« *coccus* »). Il reste à ce jour le plus petit eucaryote photosynthétique décrit (Courties *et al.*, 1994) et de nombreuses souches ont été isolées dans de nombreuses régions océaniques. Ainsi, basé sur l'analyse des séquences de l'ADNr 18S, quatre clades sont aujourd'hui distingués dont trois sont décrits comme des espèces (Marin et Melkonian, 2010).

Le clade A correspond à l'espèce *Ostreococcus lucimarinus* (Palenik *et al.*, 2007) et contient plusieurs souches de surface adaptées aux fortes luminosités (Rodríguez *et al.*, 2005). Le clade B, l'espèce encore à décrire *Ostreococcus sp.*, est un écotype de profondeur adapté à la faible luminosité (Cardol *et al.*, 2008; Rodríguez *et al.*, 2005). Le clade C est composé de souches polyvalentes liées à la première espèce décrite *Ostreococcus tauri*. Enfin le clade D, l'espèce *Ostreococcus mediterraneus*, isolée en Méditerranée (Guillou *et al.*, 2004 ;Subirana *et al.*, 2013) semble être abondant dans cette zone. En revanche, la similarité de la composition pigmentaire, (Rodríguez *et al.*, 2005) ainsi que des séquences de l'ADNr 18S et des *Internal Transcribed Spacer* (ITS) semble suggérer une diversité réduite au sein des espèces.

2) Les Prasinovirus

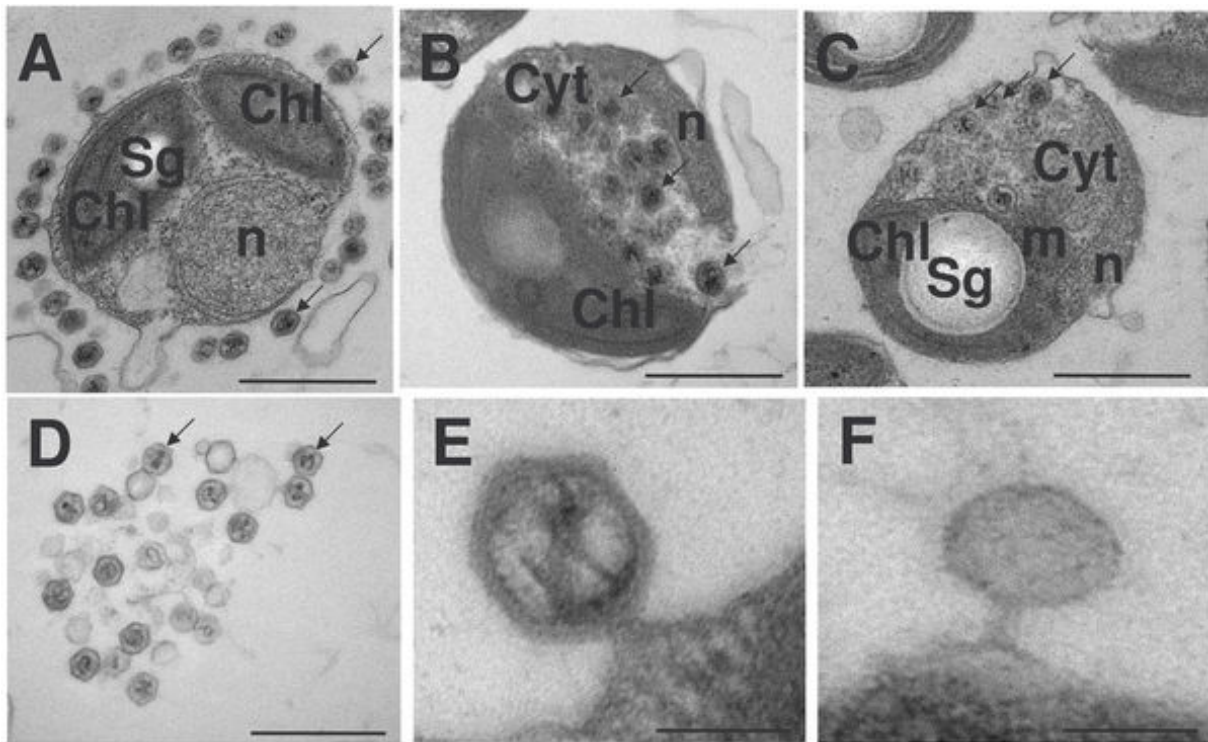


Figure 11 : Cycle lytique du virus OtV5. Les barres A-D représentent 500 nm et E et F représentent 50 nm. Les flèches indiquent les particules virales. Chl : chloroplaste ; Cyt : cytoplasme ; n : noyau ; Sg : grain d'amidon. A-C = les particules s'adsorbent sur la membrane d'*Ostreococcus tauri* et se répliquent à l'intérieur de l'hôte, D = virus libre, E-F : injection du génome du virus dans l'hôte. Derelle et al. (2008).

Les Prasinovirus font partie de la famille des Phycodnaviridae (Figure 12), qui appartiennent au groupe des *Nucleo-Cytoplasmic Large DNA Virus* (NCLDV). Les NCLDV semblent former un groupe monophylétique infectant des animaux et divers eucaryotes unicellulaires, et a été décrit comme tel pour la première fois en 2001 (Iyer *et al.*, 2001). Ce groupe est composé des familles *Poxviridae*, *Asfarviridae*, *Iridoviridae*, *Ascoviridae*, et les *Phycodnaviridae* (Yutin *et al.*, 2009), trois groupes de virus isolés à partir d'amibes, *Marseilleviridae*, *Mimiviridae* et les *Pandoravirus* (Philippe *et al.*, 2013; Claverie et Abergel, 2009 ; Raoult *et al.*, 2004) ainsi qu'un autre groupe en attente de classification, les *Pithovirus* (Claverie et Abergel, 2014). Le groupe des NCLDV a pour particularité de se répliquer à l'intérieur du cytoplasme même si une initiation au sein du noyau est possible (Iyer *et al.*, 2006, Figure 11). Les *Phycodnaviridae* sont caractérisés par une capsidie icosaédrale (à l'exception des *Pandoravirus*), sans queue ou enveloppe, ainsi qu'un grand génome d'ADN double brin allant de 160 à 560 kb (Van Etten *et al.*, 2002 ;

Van Etten et Meints, 1999 ; Wilson *et al.*, 2009). Les *Phycodnaviridae* sont constitués de 6 genres, nommés suivant les organismes qu'ils sont capables d'infecter : les *Chlorovirus*, les *Coccolithovirus*, les *Prymnesiovirus*, les *Phaeovirus*, les *Raphidovirus* et les *Prasinovirus* (Dunigan *et al.*, 2006 ; Iyer *et al.*, 2006 ; Wilson *et al.*, 2009).

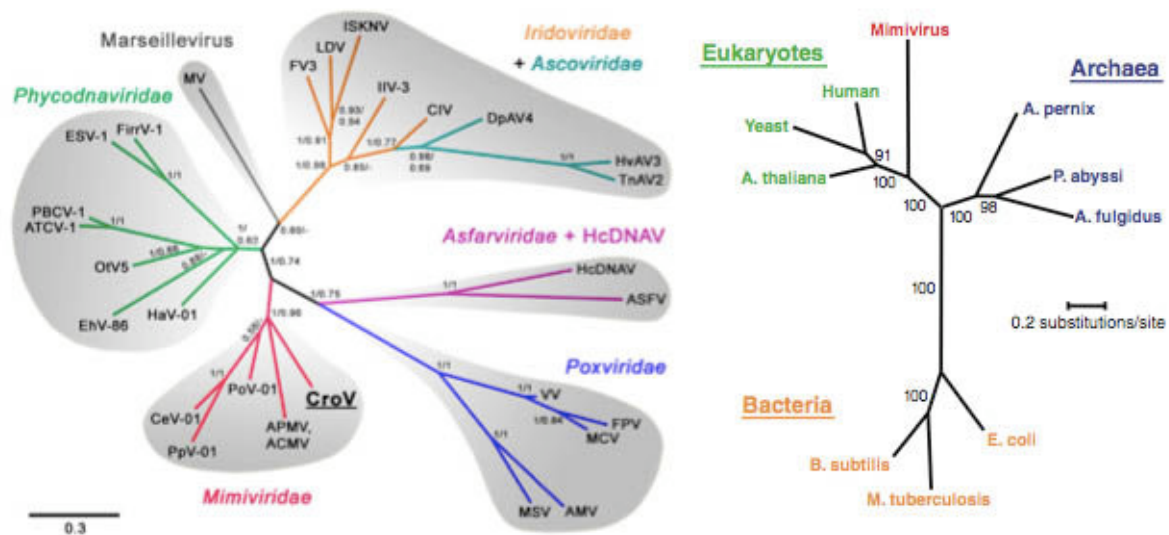


Figure 12 : Reconstruction phylogénétique des membres des NCLDV. Arbre généré à partir de l'analyse de la région conservée de l'ADN polymérase B (Fischer *et al.* 2010)

Jusqu'à présent la diversité des *Phycodnavirus* n'a été estimée qu'avec l'aide de deux principaux marqueurs moléculaires : la séquence codante partielle de la polymérase B (PolB) et le gène codant pour la protéine majoritaire de la capsid (MCP). Ces primers ont déjà fourni de nombreux résultats sur la distribution et la diversité de ces virus (Schroeder *et al.*, 2003 ; Short, 2012). Plus particulièrement, l'utilisation du marqueur PolB a montré une distribution étendue et variable des Prasinovirus dans les environnements marins (Bellec *et al.*, 2009 ; Chen *et al.*, 1996 ; Clerissi *et al.*, 2015 ; Short et Suttle, 2003) ainsi que dans les environnements lacustres (Clasen et Suttle, 2009 ; Short et Short, 2009). Une seule étude conduite sur des échantillons provenant d'un lac ultra-oligotrophique limnopolaire a révélé que les phycodnavirus pouvaient dominer la fraction des viriomes de virus à ADN double brin en période estivale mais pas en période hivernale (Lopez *et al.*, 2009). Dans ces différentes études, réalisées en

eau douce ou marine, la plupart, voire toutes les séquences obtenues en ciblant la PolB, sont extrêmement proches des *Prasinovirus* (Larsen *et al.*, 2008).

La présence d'une intéine (élément génétique parasite retrouvé au sein des motifs conservés des gènes de leurs hôtes, (Gogarten *et al.*, 2002)) pourrait inhiber l'amplification du gène de la polymérase virale de certains Prasinovirus. Partant de cette hypothèse, il a été montré récemment que l'utilisation des amorces « classiques » AVS1-AVS2 (Chen et Suttle, 1995) peut conduire à sous-estimer la diversité virale réelle (Clerissi *et al.*, 2012).

Les *Prasinovirus* décrits forment un groupe monophylétique constitué de virus marins infectant les trois genres dominants des *Mamiellophyceae* (Figure 10). Ces virus forment 3 groupes (Figure 13) composés des virus infectant *Ostreococcus* (OtV pour l'espèce *tauri* et OIV pour l'espèce *lucimarinus*, OmV pour l'espèce *mediterraneus*), les virus infectant *Bathycoccus* (BpV), et les virus infectant *Micromonas* (MpV) qui fut le premier décrit en 1979 par Mayer et Taylor. Leur taille est d'environ 100 nm et ils possèdent une capsidie icosahédrale, généralement commune au NCLDV (*Pandoravirus* étant l'exception).

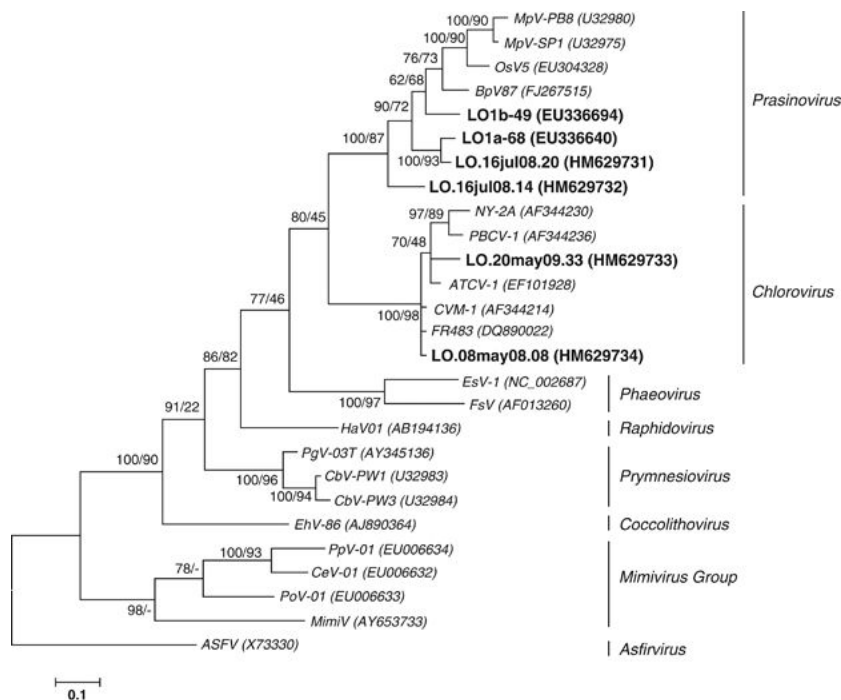


Figure 13 : Phylogénie des NCLDV basée sur PolB

A l'heure actuelle, environ 700 souches de Prasinovirus ont été isolées (Derelle *et al.*, 2008 ; Bellec *et al.*, 2009, 2010a, b , Clerissi *et al.*, 2012), mais seulement 13 génomes ont été séquencés, annotés et publiés (Derelle *et al.*, 2008 ; Weynberg *et al.*, 2009, 2010, Moreau *et al.*, 2010, Derelle *et al.*, 2015). Ces génomes ont une taille assez peu variable (200 kb +/- 30kb). Ces prasinovirus sont lytiques (Figure 11) et possèdent un spectre infectieux limités à quelques souches, voire quelques clades (correspondant probablement à des espèces au sein d'un genre (Clerissi *et al.*, 2012 ; Derelle *et al.*, 2015). Les différentes études réalisées sur ce groupe suggèrent une distribution spatiale variable, même s'ils sont abondants dans une majorité de milieux aquatiques. En effet, les virus de *Micromonas* semblent posséder une large distribution avec une prédilection pour les environnements eutrophes (Baudoux *et al.*, 2015). Les virus *Ostreococcus tauri* sont majoritaires dans les environnements côtiers, plus particulièrement lagunaires, où leurs hôtes sont abondants (Bellec *et al.*, 2010b). En revanche, ceux infectant l'espèce *O. lucimarinus* ont été isolés d'échantillons d'eaux provenant de plusieurs océans sans distribution spatiale spécifique (Bellec *et al.*, 2010a). Les 7 virus tirés de l'étude précédente dont les génomes ont été publiés montrent une séparation en deux groupes bien distincts basée sur des caractéristiques génomiques (Derelle *et al.*, 2015).

Les dynamiques classiques des phycodnavirus décrites par Short *et al.* (2011) grâce à l'utilisation d'une approche de PCR quantitative, montrent une abondance maximale durant la période estivale et minimale le reste de l'année. Cette étude montre la persistance de prasinovirus et de chlorovirus durant la période froide à des valeurs inférieures à 100 copies mL⁻¹. Ce résultat supporte la théorie de la « seed bank », qui propose qu'une partie seulement des virus reste active à une période donnée, dans l'attente d'un développement de leur hôte (Breitbart et Rohwer, 2005). L'activité des virus serait alors influencée par l'abondance de leurs hôtes.

VI Objectif et organisation de la thèse

La majorité des études disponibles visant à comprendre les dynamiques virales se basent encore sur des méthodes dites « classiques », impliquant l'utilisation d'hôtes en culture. L'un des désavantages majeurs de ces techniques culture-dépendantes est la faible diversité des souches d'hôtes utilisables, considérant la diversité naturelle retrouvée dans l'environnement (Massana, 2011). Ce biais limite d'autant nos connaissances sur les dynamiques intrinsèques des communautés virales. L'utilisation des techniques de séquençage à haut débit offre une nouvelle manière d'appréhender ces questions de diversité et de dynamique. Elles permettent de s'affranchir des techniques culture-dépendantes qui restent néanmoins indispensables pour comprendre cette diversité et étudier plus en détail des souches de virus données.

L'objectif de ce travail de thèse est de mieux comprendre l'impact des facteurs environnementaux sur les communautés virales, en utilisant le modèle Prasinovirus-Mamiellophyceae.

Je présente d'abord une étude à laquelle j'ai pris part, étude qui illustre les relations qui unissent les communautés de Prasinovirus et de Mamiellophyceae, en particulier les interactions d'abondances entre trois clades-espèces de *Micromonas* et leurs virus. Grâce à ce travail, il nous a été possible de comprendre comment, dans un environnement donné, sont associées les dynamiques d'abondance en fonction des spécificités virales.

La seconde partie expose le suivi de la dynamique de souches clés de Mamiellophyceae par PCR quantitative au sein d'environnements contrastés dans le Golfe du Lion suivi réalisé grâce à un échantillonnage mensuel d'un an (mars 2013 à avril 2014). Nous avons conçu et optimisé 12 armorces spécifiques pour suivre de façon quantitative 12 membres (clade et espèces) des Mamiellophyceae. Ces observations, couplées aux interactions mises en valeur par l'étude précédente, nous ont amené à proposer des hypothèses sur l'évolution théorique des communautés de Prasinovirus, pour les comparer ensuite aux dynamiques réelles observées au sein d'autres environnements sur les mêmes échantillons.

Enfin, dans une troisième partie, l'utilisation des techniques de séquençage haut débit nous a permis de suivre les communautés de prasinovirus et de leurs hôtes dans ces mêmes milieux contrastés. L'acquisition en parallèle d'un grand nombre de descripteurs environnementaux, nous a également permis d'étudier comment ils influencent ces communautés.

Nous pouvons résumer les objectifs de cette thèse en quatre questions :

- Quelles sont les relations entre les Prasinovirus et les *Mamiellophyceae* ?
- Quelles sont les variations spatio-temporelles des communautés de *Mamiellophyceae* dans les environnements contrastés du Golfe du Lion?
- Quelles sont les variations spatio-temporelles des communautés de prasinovirus dans les environnements contrastés du Golfe du Lion ? Tout est partout ?
- Existe-t-il une différence entre les dynamiques théoriques et les dynamiques observées : vers l'identification d'un contrôle par les facteurs environnementaux.

Au cours de ce travail, nous avons initié, en collaboration avec le Laboratoire de Biométrie et Biologie Evolutive (LBBE) de Lyon, le développement d'une méthode visant à identifier à partir des séquences environnementales les souches virales et hôtes possiblement associées à travers une relation hôte-parasite.

Enfin, de nouveaux virus des microalgues *Mamiellophyceae* *Crustomastix* et *Mamiella* ont été isolés durant ce travail de thèse, et leurs génomes sont en cours de séquençage. Bien que cette partie soit importante pour l'étude des Prasinovirus, principalement pour leur identification dans les bases de séquences environnementales, elle ne sera abordée que dans la discussion générale.

Cette thèse se présente sous la forme de publications scientifiques.

- **Article 1** : Chapitre 3 - *Interplay between Micromonas species and their viruses in the Western English Channel*

- A-C. Baudoux, H. Lebretonchel, H. Dehmer, M. Latimier, R. Edern, F. Rigaut-Jalabert, P. Gei, L. Guillou, E. Foulon, Y. Bozec, T. Cariou, Y. Desdevises, E. Derelle, N. Grimsley, H. Moreau, and N. Simon. (Environ Microbiol Rep. 2015 7(5): 765-73. doi: 10.1111/1758-2229.12309)

Bien que je ne sois pas premier auteur de ce travail, nous avons jugé pertinent de l'intégrer dans cette thèse car il permet de mettre en avant la connexion existante entre les hôtes et leur virus, tant d'un point de vue dynamique, qu'en termes de stratégie de réplication

- **Article 2** : Chapitre 4 - *Green microalgae of the Mamiellophyceae dominate picoeukaryotic communities throughout the year in a coastal lagoon revealed by qPCR*
- H. Lebretonchel, M. Ros, N. Grimsley, Y. Desdevises (soumis à Applied and Environmental Microbiology)

L'étude des dynamiques de Mamiellophyceae au sein des écosystèmes étudiés est une approche nécessaire afin de mieux comprendre les dynamiques des virus associés. L'utilisation de la PCR quantitative nous a permis de suivre les principaux groupes de Mamiellophyceae et ainsi de montrer qu'au sein d'un environnement semi-fermé (deux sites lagunaires), la dynamique des communautés est extrêmement variable.

- **Article 3** : Chapitre 5 - *Impact of environmental factors on Prasinovirus communities*
- H. Lebretonchel, B. Alric, E. Derelle, N. Grimsley, S. Dray, Y. Desdevises (en préparation.)

Ce travail vise à comprendre les relations entre les communautés de Prasinovirus et les facteurs environnementaux. Il met en évidence une forte relation entre les dynamiques des hôtes et des virus, en lien avec les facteurs environnementaux, avec un impact fort de ces facteurs en période estivale.

Chapitre II - Matériel et Méthodes

I Zone d'échantillonnage

Le but de ce travail est d'étudier l'impact des facteurs environnementaux sur les communautés virales, pour cela nous avons choisi des sites présentant un fort contraste entre eux. Les stations d'échantillonnage ont été sélectionnées en Méditerranée, et plus particulièrement dans le Golfe du Lion. Cette région a comme particularité de présenter des environnements contrastés dans un espace géographique restreint.

Ces échantillonnages font partie du projet ANR Blanc DECOVIR, Diversity and Environmental Control of Prasinoviruses (ANR-12-BSV7-0009).



Figure 1 : Localisation des sites de prélèvement
(source : maps.google.fr/)

Quatre sites ont été choisis : 2 lagunaires, 1 côtier et 1 hauturier (Figure 1, Tableau 1), parmi ceux utilisés lors des précédents travaux sur les prasinovirus (Bellec *et al.*, 2010 ;

Derelle *et al.*, 2008 ; Moreau *et al.*, 2010 ; Thomas *et al.*, 2011). Chacun de ces sites a été échantillonné mensuellement en surface.

Tableau 1 : Stations d'échantillonnage

Site	Longitude (WGS84)	Latitude (WGS84)	Profondeur (m)
Leucate A	3,021	42,805	1,5
Leucate B	2,998	42,873	1
SOLA	3,145	42,489	27
MOLA	3,543	42,453	600

Deux sites ont été choisis dans la lagune de Salse-Leucate qui s'étire sur une longueur de 14 km et atteint 6,5 km dans sa plus grande largeur, représentant un volume d'eau de 95 millions de m³, un au sud (Leucate A) et un au nord-ouest (Leucate B) (Figure 1, Tableau 1). Bien que situées toutes les deux dans le bassin de Leucate Nord, ces deux stations constituent des environnements différents. Le site Leucate A est situé près du grau (point de connexion entre la lagune et la mer) sans zone de résurgence d'eau douce, et le site nord-ouest est situé à l'opposé du grau et au sein d'une zone de résurgence d'eau douce. Les sites lagunaires sont des zones à forte variabilité intra- et inter-annuelle. Cette lagune est un milieu présentant de variations environnementales de plus grande amplitude au cours de l'année que les deux sites marins. Cette région est caractérisée par un forçage climatique important. En effet d'octobre à mai, la Tramontane, qui est un vent violent du Nord-Ouest, exerce une influence particulièrement forte sur l'hydrodynamisme des zones lagunaires. Lors de cette période, le vent et les apports d'eau douce favorisent la sortie d'eau. Inversement, en été, lors de l'arrêt de ce vent, le vent d'Autan (sud-ouest) couplé à l'évaporation induit une entrée d'eau de mer à l'intérieur de la lagune (Ladagnous et Le Bec, 1998). Malgré ces dynamiques, la lagune de Leucate est globalement un bassin de dilution caractérisé par des salinités inférieures

aux sites marins, mais aussi par des variations de températures plus importantes que la mer ouverte.

L'enclavement et la faible profondeur (en moyenne 1,5 m) de ces environnements les rendent particulièrement sensibles aux facteurs météorologiques. En particulier, l'action des vents engendre une dissymétrie entre les rives occidentale et orientale du bassin. Cette dernière est exposée aux vents et subit un effondrement avec transport actif des sédiments dans le bassin, phénomène qui n'a pas lieu sur la rive occidentale protégée du vent par sa topographie des hauteurs qui la borde. De plus, l'importance du bassin versant, couvrant une superficie de 160 km², favorise l'apport de matière organique contribuant au caractère eutrophe de cet endroit et la mise en place d'un gradient physico-chimique fort au cours de l'année, ce dernier existant également le long d'un gradient géographique allant des zones de résurgence jusqu'au grau. Les précipitations, les vents, les marées ainsi que les différences de pression atmosphérique provoquent des fluctuations des masses d'eau entraînant des variations d'échanges entre la lagune et la mer (Figure 2 ; Hervé, 1978). La température de l'eau est similaire à celle de l'air et Arnaud et Raimbault (1969) ont mis en évidence que les fluctuations thermiques saisonnières et journalières sont liées pour 88 % à celles de l'air. Les moyennes maximales sont régulièrement enregistrées en juillet (23°C) et les minimales en décembre (8,5°C). Cette lagune n'est pas limitante au niveau des sels nutritifs, les valeurs moyennes en phosphate et azote sont proches de 0,3 µmol/L, soulignant le caractère eutrophe de la lagune (Boutière *et al.*, 1982). Cette lagune est également une zone conchylicole importante et la présence de cette activité économique a un impact écologique sur les communautés de micro-organismes (Ladagnous et Le Bec, rapport IFREMER, 1998) mais principalement dans le bassin nord où est localisée l'activité. Cinq classes d'organismes photosynthétiques ont été rapportées comme jouant un rôle important dans la production primaire : les Chlorophycées, les Diatomées, les Cryptophycées, les Dinoflagellés et les Coccolithophoridés (Heinrich, 1988). Malgré tout, comparée aux autres lagunes, celle de Leucate est considérée comme pauvre en phytoplancton avec une biomasse chlorophyllienne comprise entre 0,3 et 2,6 mg/m³ (Anonyme, 1984).

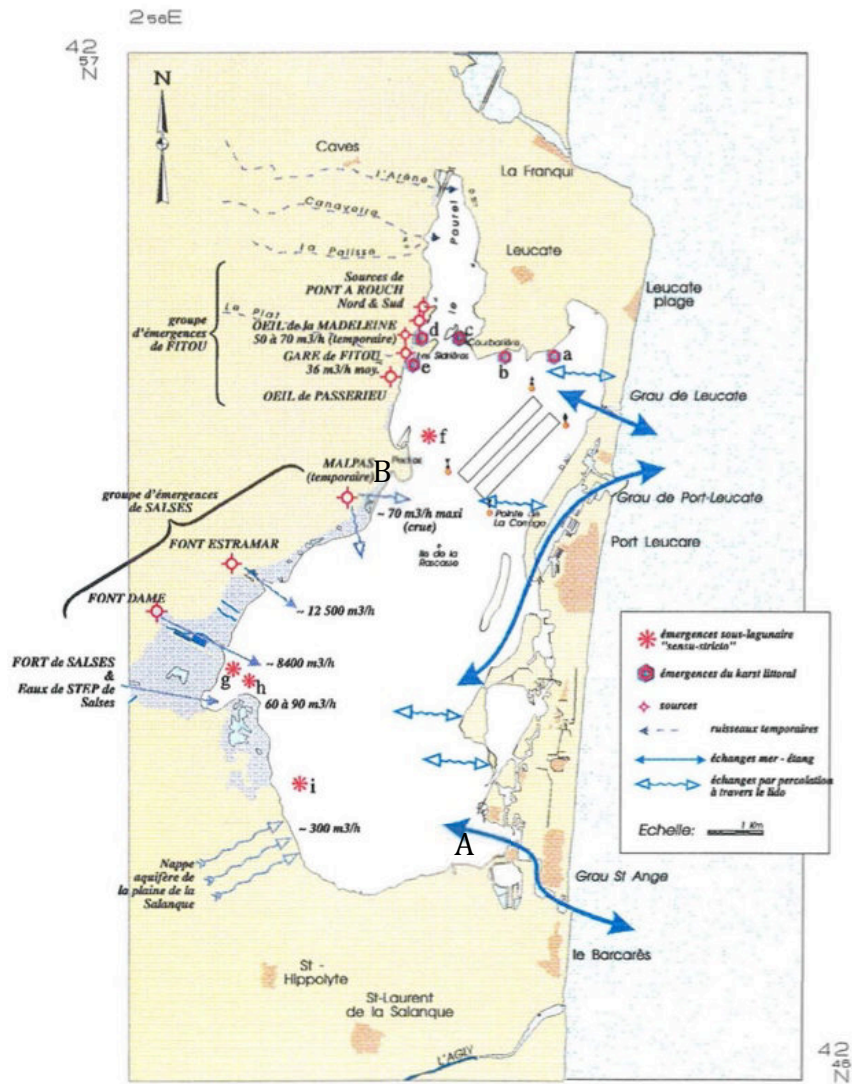


Figure 2 : Lagune de Salses-Leucate (source : rapport IFREMER, Analyse bibliographique 1997)

Malgré le caractère semi-fermé de la lagune, des différences importantes sont observables entre les deux stations échantillonnées (Figure 3). Le point B possède une température et une salinité plus importante que le point A résultant de la présence d'une zone de concentration au nord de cette lagune. Ces caractéristiques salines plus proches de l'eau de mer pourraient influencer les populations algales en favorisant le recrutement de souches marines dans cette zone lors de l'entrée d'eau dans la lagune.

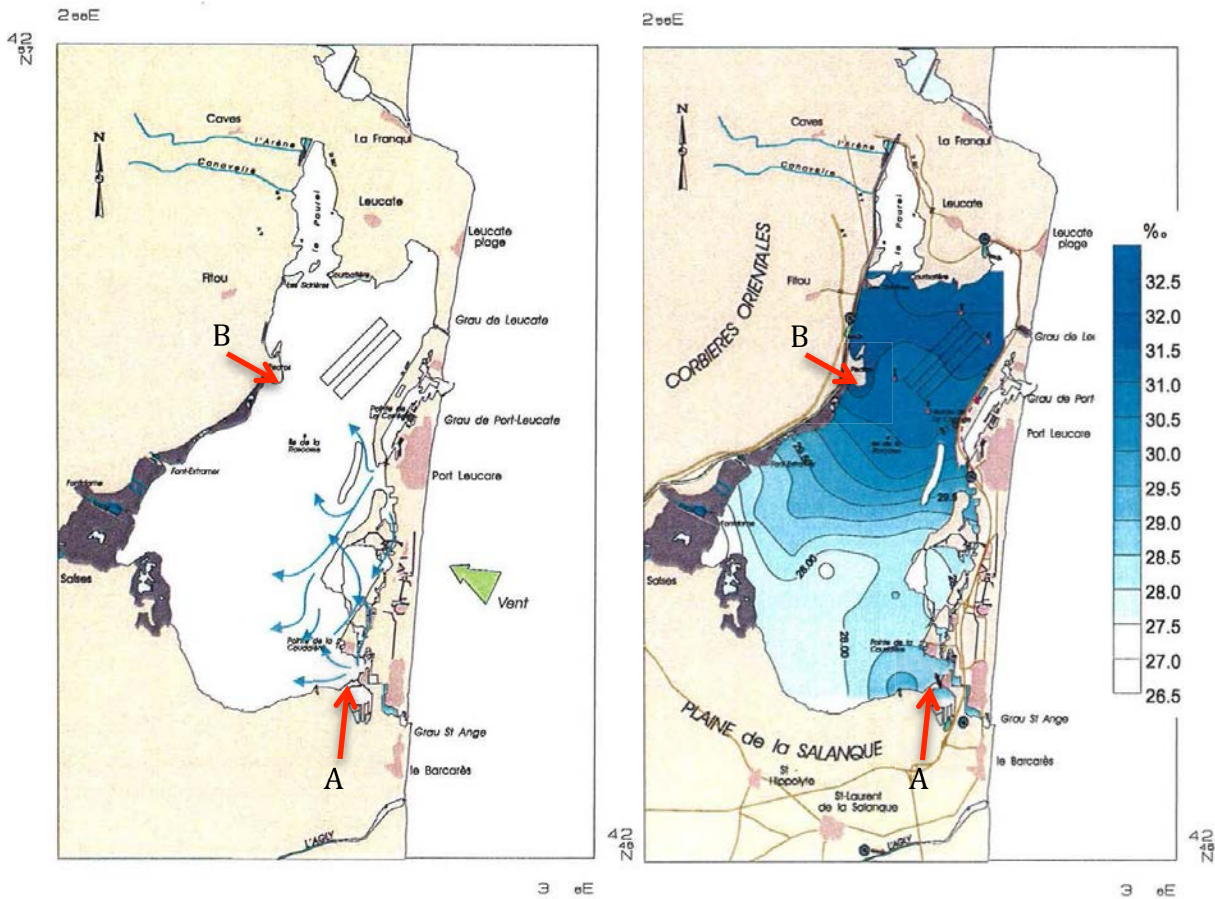


Figure 3 : Circulation par vents d’Autan (Sud-Est) dans la Lagune de Salses-Leucate (gauche). Salinité dans la lagune de Salses-Leucate en juin 1996 (droite). Source : Ladagnous et Le Bec, Rapport IFREMER Lagune de Salses-Leucate, 1998.

En comparaison, outre le gradient trophique distinguant la lagune eutrophe et l’environnement marin, les sites marins présentent une variabilité annuelle réduite en terme de température et de salinité.

Le site côtier est la station SOLA (Service d’Observation du Laboratoire Arago) située dans la baie de Banyuls-sur-Mer et faisant partie du réseau SOMLIT (Service d’Observation en Milieu LIToral, <http://somalit.epoc.u-bordeaux1.fr>). Ce site présente une profondeur de 24 m et contrairement aux lagunes, l’apport du bassin versant est principalement réduit à une rivière active de la fin de l’automne jusqu’au printemps. Ce site est à priori moins impacté par les événements météorologiques du fait de sa faible profondeur. Il présente une faible variation de salinité avec une valeur moyenne d’environ 38, avec une baisse légère durant l’hiver à 37,5 correspondant à l’activité de la rivière se déversant dans la baie. La température moyenne est de 15°C avec des

maximum de 22-23°C atteints régulièrement pendant l'été (base de données SOMLIT). Ce site est considéré comme mésotrophique (Salter *et al.*, 2015 ; Bethoux *et al.*, 2002), comme en témoigne la concentration moyenne de biomasse chlorophyllienne de 1 mg/m³ , avec une augmentation maximale atteignant 1,5 mg/m³ au printemps, mais réduite à 0,2 en période estivale (base de donnée SOMLIT). Cette limitation de biomasse est principalement due à la concentration en phosphate, avec des valeurs moyennes de 0,04 µmol/L. En comparaison, l'azote et principalement les nitrates sont retrouvés en forte concentration moyenne (1,2 µmol/L), due aux fortes valeurs mesurées en hiver qui sont la conséquence du lessivage du bassin versant viticole enrichi en nitrates par les viticulteurs. En période estivale, ces valeurs sont de l'ordre de 0,09 µmol/L (base de données SOMLIT).

Le site hauturier est la station MOLA (*Microbial Observatory Laboratory Arago*) du réseau MOOSE (<http://www.insu.cnrs.fr/environnement/atmosphere/moose-mediterranean-ocean-observing-system-on-environment>) située à environ 20 milles au large. Il présente une profondeur de 600 m et est parfois influencé par le courant Liguro-Provençal. En comparaison des autres sites, ce site est stable au cours de l'année puisque faiblement impacté par les apports terrestres (Laghdass *et al.*, 2012 ; Krom *et al.*, 1991), comme en témoigne les moyennes de température (15,9 °C +/- 3) ou de salinité (38 +/- 0,26). Ce site est considéré comme oligotrophique avec des valeurs moyennes en azote et phosphate de 0,7 µmol/L et 0,03 µmol/L respectivement. Ces conditions trophiques ont un impact direct sur les communautés, puisque la concentration moyenne en chlorophylle est de 0,3 mg/m³ avec un minimum de juin à décembre de 0,13 mg/m³ (base de données MOOSE). A l'opposé des lagunes, les stations SOLA et MOLA sont dominées par les cyanobactéries et non par les picoeucaryotes.

II Stratégie d'échantillonnage

1) Echantillonnage et isolement

a. Les Prasinovirus

Nous avons utilisé différents protocoles pour recueillir les particules virales libres dans l'environnement : 1) protocole des plages de lyse basée sur l'infection de cultures de Mamiellophyceae afin d'isoler de nouvelles souches de Prasinovirus, 2) protocole de précipitation et de filtration afin de récupérer la fraction virale libre pour l'étude de la diversité virale.

1) Isolement de nouvelles souches virales

i) Culture de Mamiellophyceae

Notre but a été d'isoler de nouveaux virus sur les Mamiellophyceae n'ayant pas encore de virus identifiés, à savoir *Mantoniella*, *Mamiella*, *Crustomastix*, *Dolichomastix* et *Monomastix*. Dans la mesure du possible, les souches utilisées ont été isolées en Méditerranée. L'ensemble de ces souches (Tableau 2) est disponible soit à la Roscoff Culture Collection (<http://www.sb-roscoff.fr/Phyto/RCC/>), soit au National Center for Marine Algae and Microbiota (USA ; <https://ncma.bigelow.org/>), soit au National Institute for Environmental Studies (Japon, <https://www.nies.go.jp/index-e.html>).

Les souches ont été maintenues dans un milieu de culture adapté à leur croissance : L1, MK (milieu salin) ou MC (milieu d'eau douce). Afin de garantir un taux de croissance optimum, ces souches ont été repiquées chaque semaine et placées en incubateur à 20°C, avec une photopériode de 12-12 (jour-nuit) à 200 $\mu\text{E.m}^2.\text{s}^{-1}$.

RCC : Roscoff Culture Collection, NCMA (ancien CCMP), National Center for Marine Algae and Microbiota, USA; NIES : National Institute for Environmental Studies, Japon.

Ordre	Famille	Genre	espèce	Origine	Milieu utilisé	Noms	Taille (µm)
Mamiellales	Bathycoccaceae	<i>Ostreococcus</i>	sp. (Clade B)	Tropical Atlantic (+21° 2', -31° 8')	K	RCC 809	1
		<i>Ostreococcus</i>	<i>tauri</i>	Mediterranean Sea (+42° 29', +3° 8')	K	RCC 1108	1
		<i>Ostreococcus</i>	<i>mediterraneus</i>	Mediterranean Sea (+43° 3', +3° 0')	K	RCC 2572	1
		<i>Ostreococcus</i>	<i>lucimarinus</i>	Mediterranean Sea (+43° 3', +3° 0')	K	CCMP2490	1
		<i>Bathycoccus</i>	<i>prasinos</i>	North Sea (+54° 11', +7° 54')	K	RCC 682	1
	Mamiellaceae	<i>Micromonas</i>	sp. Clade A	Pacific ocean (-22° 20', +166° 20')	K	RCC 299	1-2
		<i>Micromonas</i>	sp. Clade B	Mediterranean Sea (+41° 40', +2° 48')	K	RCC 434	1-2
		<i>Micromonas</i>	sp. Clade C	Baltic sea (+58° 11', +9° 6')	K	RCC 373	1-2
		<i>Mantoniella</i>	<i>squamata</i>	North sea (+52° 58', +0° 39')	K	RCC 417	4
		<i>Mamiella</i>	sp.	Mediterranean Sea (+30° 8', +10° 3')	L1	RCC 391	4
Dolichomastigales	Crustomastigaceae	<i>Crustomastix</i>	<i>stigmata</i>	Mediterranean Sea (+40° 54', +14° 9')	L1	CCMP 3273	5
	Dolichomastigaceae	<i>Dolichomastix</i>	<i>tenuilepis</i>	Mediterranean Sea (+40° 30', +14° 15')	L1	CCMP 3274	5
Monomastigales	Monomastigaceae	<i>Monomastix</i>	<i>minuta</i>	Tsuchiura, Ibaraki (Japan)	C	NIES 255	8-10

ii) Méthode d'isolement culture-dépendante

L'isolement de nouvelles souches virales consiste à mélanger un échantillon environnemental filtré sur 0,45 µm avec une culture d'hôte spécifique. Si une clarification de la culture est observée après 7 à 15 jours, 100 µl de ce mélange sont prélevés pour réinfecter une nouvelle culture. Cette opération est répétée 3 fois pour augmenter la quantité de virus potentiellement présents dans l'échantillon. Après la troisième infection, 10 à 20 µl de ce « potentiel » lysat clarifié, et centrifugé à 7500 g pendant 15 min, sont prélevés pour infecter une nouvelle culture. Après une heure d'incubation afin de favoriser l'adsorption des virus sur les hôtes, cette culture (10-20 millions de cellules.mL) est mélangée à une solution d'agarose (0,15 % final). Après une semaine d'incubation en chambre humide dans les mêmes conditions de lumière que les souches, des plages de lyse ou *Plaque Forming Unit* (PFU) (Figure 4) dans l'agarose ont

parfois été observées et prélevées. Les virus contenus dans l'agarose de ces plaques sont remis en suspension dans du tampon SM à 4°C puis utilisés pour réinfecter une culture. Cette étape de plage de lyse est répétée trois fois afin d'obtenir une culture clonale du virus isolé.

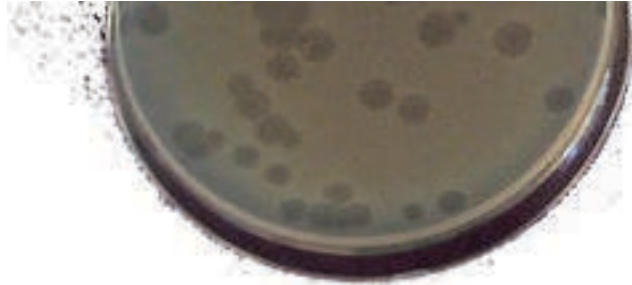


Figure 4: Plages de lyse sur une culture étalée avec de l'agarose

2) Echantillonnage des virions libres

Pour étudier la diversité de la fraction virale libre au sein des écosystèmes, nous avons mis en place un protocole indépendant des cultures d'hôtes disponibles dans les différentes collections.

La méthode que nous avons utilisée est basée sur la précipitation des virus à l'aide de chlorure de fer (FeCl_3). Ce procédé est couramment utilisé dans le nettoyage des eaux usées et a été adapté par John *et al.*, (2011). Il est basé sur la modification des charges ioniques, ayant pour effet l'agglomération des particules. Ce procédé, en plus d'être peu onéreux, présente l'avantage d'être plus efficace que les techniques précédemment mises au point pour l'étude des communautés virales (comme la filtration tangentielle).

Ce protocole a été conçu pour le traitement d'un volume d'eau conséquent (20L). Une filtration en série de l'eau prélevée sur des filtres en Polyéthylène Sulfone (PES, Millipore) (5, 3, 1, 0,45 et 0,2 μm) a été réalisée, en plus de la pré-filtration (à 200 μm) réalisée sur le terrain. L'utilisation de filtres en PES est préconisée car à l'opposé des filtres en polycarbonate (PC) ou en cellulose, les filtres en PES ne favorisent pas l'adsorption des particules virales.

Par la suite, une solution de FeCl_3 a été ajoutée pour une concentration finale de 1g.L^{-1} . Après une incubation d'une nuit à température ambiante, les 20 litres d'échantillon pré-traités sont refiltrés sur un filtre en PC de $1\ \mu\text{m}$ afin de récupérer le floculat. Le filtre est ensuite stocké à 4°C et traité dans un délai maximum de 4 mois. Les auteurs de cette méthode ont montré une absence de destruction significative avec ce délai de stockage avant utilisation.

L'agrégat contenu sur le filtre est resuspendu dans 10 ml de tampon Tris-HCl, Mg_2EDTA (pH 6) et concentré en un volume de 2 ml grâce à une étape d'ultracentrifugation ($180000\ \text{g}$, 2 h). L'échantillon obtenu est immédiatement utilisé pour extraire l'ADN des particules virales ainsi concentrées.

b. La fraction phytoeucaryotique

Afin d'extraire la fraction phytoeucaryotique des échantillons, 5 litres d'eau de mer ont été filtrés sur un filet à plancton de $10\ \mu\text{m}$, excluant les particules non ciblées. Ces 5 litres ont ensuite été concentrés par VivaFlow 200 (50 KDa, Sartorius) en 50 ml, cryopréservés par l'ajout de Glycine-Bétaine (concentration finale 8 %) et stockés à -80°C en vue d'un tri par cytométrie en flux. Cette technique est connue pour concentrer également les virus, mais l'utilisation de tri cellulaire et d'amorces spécifiques pour le 18S nous a permis de négliger cette fraction lors des étapes suivantes.

Le tri cellulaire a pour but d'augmenter la quantité d'ADN des organismes d'intérêt, dans notre cas la communauté des Mamiellophyceae. Il a été effectué en triplicat pour chaque échantillon, avec 200 000 cellules isolées par replicat. Les paramètres de tri utilisés ont été la fluorescence des particules (FL3 et FL2, respectivement pour mesurer l'émission à 670 et 585 nm), le Side Scatter (SSC) afin de discriminer la fraction phytoeucaryotique avec un critère de sélection élevé (*fine tune*) et des critères de tri stricts (Mask : purity, amplitude : 22,1 V , fréquence : 87 kHz, drop 1 : 192 pixels, gap : 7 pixels, drop delay : 44,52, pression : 85 psi). Grâce à ce tri, les cellules photosynthétiques récupérées ont été stockées à -80°C en vue de l'extraction de l'ADN.

2) Extraction et amplification

En fonction des objectifs, les techniques utilisées ont été diverses. Ainsi, nous distinguerons les techniques suivant les entités étudiées (virus isolés ou provenant des échantillons environnementaux). Nous distinguerons aussi les techniques visant à amplifier les virus isolés, les communautés d'hôtes et de virus, ainsi que la PCR quantitative (qPCR) effectué sur les hôtes.

a. Extraction d'ADN

i) Visualisation et extraction des génomes viraux par PFGE

Pour les virus isolés, après infection de 2L de culture, les virus ont été concentrés à l'aide de VIVACELL 100 (50 KDa, Sartorius). Nous avons utilisé la technique de PFGE (*Pulsed Field Gel Electrophoresis* = Electrophorèse en champ pulsé) afin de déterminer les tailles génomiques des virus. En effet, les techniques conventionnelles d'électrophorèses ne permettent pas d'obtenir une résolution pour des fragments génomiques supérieurs à 50 kb, contrairement à la PFGE. Afin d'utiliser cette méthode, l'échantillon de virus clonal concentré a été inclus dans des *plugs* d'agarose, inclusion permettant de préserver l'ADN lors des étapes de destruction de la capsid. Cette destruction est effectuée à travers l'agarose grâce à un tampon Tris Hcl 10 mM, EDTA 0,5M (pH 8), protéinase K (1mg.mL), lauroyl sarcosinate (1 %) lors d'une incubation de 24h à 37°C avec agitation. Les *plugs* sont ensuite lavés dans trois bains successifs d'EDTA 0.5M pendant 2h puis stockés dans la nouvelle solution d'EDTA 0,5M à 4°C jusqu'à leur utilisation. Les *plugs* sont déposés sur un gel d'agarose (1% agarose- TBE 0,5x) de PFGE et mis à migrer pendant 24 h :

block 1 : initial switch time = final switch time = 60s, Temps de migration = 15h, 6V/cm ;
angle = 120° ,

block 2 : initial switch time = final switch time = 90s, Temps de migration = 9h ; 6V/cm ;
angle = 120°.

Pour visualiser l'ADN viral par exposition aux UV, le gel est incubé dans un bain de BET (Bromure d'éthidium) pendant 20 min.

Lors de l'extraction de l'ADN viral par cette technique, le protocole de PFGE est le même avec 2 différences : l'agarose utilisé est de l'agarose LMP (Low Melting Point) et seuls les cotés du gel sont placés dans un bain de BET après découpe. Ce procédé permet de visualiser et de marquer la bande virale. La bande d'agarose contenant l'ADN viral "BET-free" est ensuite découpée, l'ADN est extrait de l'agarose par de GELase (EPICENTRE Biotechnologie) (Derelle *et al.*, 2008) et resuspendu dans de l'eau *free DNA*.

ii) Les communautés naturelles (virus et hôtes)

L'ADN de chacune des communautés naturelles concentrées par le tri a été extrait à l'aide d'un protocole bromure de cétrimonium (CTAB-DTT) modifié (Winnepeninckx *et al.*, 1994). Les échantillons ont été mis en contact avec un tampon CTAB (volume final de 10 ml ; concentration finale : 2 % CTAB (hexadecyltrimethylammonium bromide) ; 100 mM TrisHCl (pH = 8) ; 20 mM EDTA ; 1,4 M NaCl ; 0,2 β -mercaptoethanol ; 0,1 mg/mL proteinase K ; 10 mM DTT (dithiothreitol)), puis incubés 1 h à 60 °C. Suite à la lyse des entités biologiques, l'ADN a été purifié à l'aide de 10 mL de Chloroforme-Isomyalalcohol (24:1). Une étape de digestion des ARN a été réalisée par l'ajout d'ARNase dans la partie aqueuse récupérée par centrifugation, puis incubée 1 h à 37 °C. L'ADN purifié a ensuite été précipité par l'ajout d'isopropanol (2/3 du volume récupéré après purification). Cet ADN précipité est nettoyé avec de l'éthanol 76%-Acétate d'ammonium (10 mM). L'ADN est ensuite resuspendu dans 100 μ l d'eau ultra pure et stocké à -20 °C.

b. Amplification et séquençage

i) Les virus isolés

Afin de caractériser les séquences génomiques des Prasinovirus isolés, plusieurs couples d'amorces ont été utilisés pour amplifier les marqueurs choisis : la séquence partielle de l'ADN polymérase virale (PolB) et celle codant pour la sixième copie de la séquence codante de la protéine majoritaire de la capsid (Major Capsid Protein ; MCP) (Tableau 3). Pour la polymérase virale, le couple d'amorces AVS1-AVS2 défini par Chen et Suttle (1995) a été utilisé. Nous avons aussi utilisé un couple d'amorces défini au sein de notre équipe (VpolAS4-VpolAAS1), permettant d'amplifier un fragment de la PolB ne contenant pas de séquence d'intéine (voir Clerissi *et al.*, 2012). Ces séquences génomiques particulières peuvent atteindre ici une taille de 900 bp et inhibent parfois la réaction d'amplification. Pour la séquence codante de la MCP, nous avons utilisé le couple d'amorces VmcpAS3-VmcpAAS1.

Tableau 3 : Séquences et caractéristiques des amorces utilisées pour la production des amplicons viraux à partir de l'ADN de la fraction virale

Nom	Séquence (5'-3')	Taille (pb)	sens	Gène/Groupe cible
AVS1	GARGGIGCIACIGTIYTIGAYGC	23	Sens	PolB/Phycodnaviridae
AVS2	GCIGCRTAICKYTTYTTISWRTA	23	Anti-sens	
VpolAS4	GARGGIGCIACIGTIYTNGA	20	Sens	PolB/Prasinovirus
VpolAAS1	CCIGTRAAICCRTAIACISWR TTCAT	26	Anti-sens	
VmcpAS3	GGIGGICARMGIRTIGAYAA	20	Sens	MCP/Prasinovirus
VmcpAAS1	TGIACYTGYTCDATIARRTAYTCRTG	26	Anti-sens	

Pour l'amplification des séquences génomiques d'intérêt, les PCR ont été réalisées sur 5 μ L de lysat chauffé à 95°C pendant 10 min dans le but de détruire la capsid virale et ainsi rendre accessible le génome viral. Par la suite, nous avons ajouté 45 μ L de mix PCR contenant le tampon Promega de l'ADN Polymérase GoTaq Flexi 2 ; 0,2 mM de

chaque dNTP ; 1,5mM de MgCl₂ ; 100 mM d'amorces, ainsi que 0,5 U d'ADN polymérase GoTaq Flexi 2 (Promega).

L'amplification par PCR a ensuite été réalisée dans un Mastercycler (Eppendorf®) avec une étape initiale de 95°C (3 min), suivie de 35 cycles à 95°C (30 s), 50°C pour la PolB et 48°C pour la MCP (60 s), 72°C (90 s) ainsi qu'une extension finale à 72°C (4 min).

Chaque amplification a été contrôlée par migration sur gel d'agarose (0,8%) contenant 0,05 % de Bromure d'éthydiuim (BET) à 100V pendant 25 min, puis l'ADN a été révélé par ultraviolet (UV). Les amplicons ont été stockés à -20°C.

Ces séquences ont ensuite été séquencées en Sanger dans les deux sens par la plateforme BIO2MAR du Laboratoire ARAGO.

ii) Les communautés virales naturelles

Les séquences génomiques des marqueurs PolB et MCP ont été amplifiées et séquencées par illumina MiSeq. Nous avons été en mesure de multiplexer les banques d'amplicons au moyen d'un *multiplex identifiers* (MID), correspondant à une séquence nucléotidique spécifique, permettant de reconnaître les banques d'amplicons obtenues lors de la PCR. Un adaptateur a donc été intégré aux séquences des amorces afin d'y ajouter à posteriori une séquence MID. Du fait de cette technologie, une très grande quantité de séquences sont produites et rend impossible le contrôle manuel de la qualité des séquences. Nous avons ainsi choisi d'utiliser une ADN polymérase haute fidélité afin de réduire le nombre d'erreurs lors de l'amplification. Le protocole d'amplification et de contrôle sur gel est le même que lors de l'amplification des virus isolés.

Les banques d'amplicons produites ont été multiplexées et séquencées par la plateforme GeT-PlaGe (génomique et transcriptome) du centre INRA de Toulouse en « paired-end » de 2 x 250 bp et été séquencées en 5' et 3'.

iii) Les communautés phytoplanctoniques naturelles

Afin d'étudier les populations d'hôtes eucaryotes des différentes stations, deux séquences génomiques (régions V4 et V9 de l'ADN ribosomique 18S) ont été sélectionnées pour l'amplification et le séquençage suivant la même technologie que les séquences virales. L'amplification a été faite sur les échantillons d'hôtes concentrés et triés. Pour la région V4, le couple d'amorce TAREuk454FWD1 - TAREukREV3 (Stoeck *et al.*, 2010) a été sélectionné et pour la région V9, le couple d'amorce 1389f-1510r (Amaral-Zettler *et al.*, 2009) (Tableau 4). Un adaptateur a été ajouté à chaque couple d'amorces permettant le multiplexage des banques d'amplicons.

Tableau 4 : Séquences et caractéristiques des amorces pour les hôtes (eucaryotes)

Nom	Séquence (5'-3')	Taille (pb)	sens	Gene/Groupe cible
TAREuk454FWD1	CCAGCASCYGCGGTAATTCC	20	Sens	18SV4/Eucaryotes
TAREukREV3	ACTTTCGTTCTTGATYRA	18	Anti-sens	
1389f	TTGTACACACCGCCC	15	Sens	18SV9/Eucaryotes
1510r	CCTTCYGCAGGTTACCTAC	20	Anti-sens	

Le mix PCR est identique à celui utilisé pour les virus. L'amplification PCR a également été conduite dans un mastercycler (Eppendorf) à l'aide d'un programme PCR spécifique pour les régions V4 et V9. Pour la V4, le programme comporte une étape d'initiation de 95°C (5 min), 10 cycles (94°C pendant 30 s ; 57°C pendant 45 s ; 72°C pendant 1 min) puis 25 cycles (94°C pendant 30 s ; 48°C pendant 45 s ; 72°C pendant 1 min), suivi d'une phase d'élongation à 72°C pendant 10 min. Pour la V9, l'amplification a consisté en une phase d'initiation 95°C pendant 5 min, suivi de 30 cycles de 94°C (30 s), 57°C (45 s) et 72°C (1 min) suivi d'une phase d'élongation de 2 min à 72°C. Le protocole d'amplification et de contrôle sur gel est le même que lors de l'amplification des virus isolés.

Les banques d'amplicons générées ont été multiplexées et séquencées par la plateforme GeT-PlaGe (génomique et transcriptome) du centre INRA de Toulouse en « paired-end » de 2 x 250 bp pour la V4 et 2 x 150 bp pour la V9.

c. PCR quantitative sur les Mamiellophyceae

Nous avons développé des amorces de PCR quantitative (qPCR) afin de suivre les dynamiques d'abondance des Mamiellophyceae, principalement dans la région V4 qui permet une bonne discrimination du groupe étudié. Ce travail a fait l'objet d'une publication Chapitre 4, et cette partie ne fait que résumer les grandes lignes. La qPCR permet de suivre spécifiquement l'amplification du gène cible afin de déterminer le nombre initial de copies de ce gène dans un échantillon d'ADN. Outre l'avantage d'être rapide et relativement économique, la qPCR peut, après optimisation et calibrage, être appliquée à un grand nombre d'échantillons simultanément. En revanche, elle présente certains inconvénients. Premièrement, la spécificité repose entièrement sur les amorces, la température d'hybridation ainsi que les concentrations des différents composés utilisés. Il est donc crucial de mettre en place une phase de tests préliminaires afin de vérifier à la fois la bonne spécificité et la qualité de l'amplification. Deuxièmement, il existe un risque non négligeable de faux positifs, impliquant une surestimation de la quantification du nombre de copies. Une analyse scrupuleuse des courbes de fusion est nécessaire afin de contrôler la formation de dimères et l'amplification de séquences non-spécifiques. L'ensemble des résultats d'optimisation de qPCR est présenté chapitre 4.

i) Mise au point d'amorces spécifiques

Les séquences génomiques des différents genres et espèces des Mamiellophyceae ont été extraites de la base de données GenBank (<http://www.ncbi.nlm.nih.gov/genbank/>), données comprenant à la fois la séquence des souches utilisées par la suite lors du test de spécificité mais aussi les séquences de souches du même groupe taxonomique.

L'alignement des séquences a été réalisé en utilisant ClustalW implémenté dans BioEdit (<http://www.mbio.ncsu.edu/Bioedit/bioedit.html>). L'examen visuel des séquences alignées nous a révélé 3 régions variables que nous avons utilisées pour définir la séquence des amorces (Tableau 5). Quatre contraintes ont été choisies pour la conception des amorces, i) une taille d'amplicon comprise entre 150 et 200 bp afin d'optimiser l'amplification de la séquence, ii) une queue 3' riche en A et T afin de limiter les hybridations non spécifiques, iii) des amorces de 20 bp afin de favoriser la spécificité et iiiii) une température d'hybridation commune (58 °C).

Tableau 5 : Séquences et caractéristiques des amorces utilisées en qPCR. f = Sens, r = Anti-sens

Nom	Séquence (5'>3')	Groupe cible	Taille de l'amorce (pb)	Tm de l'amorce (°C)	Amorce GC%	Taille de l'amplicon (pb)	Tm de l'Amplicon (°C)	Amplicon GC%
OstreT-f	TGTGCTTCACGGCCACTTA	<i>Ostreococcus tauri</i>	20	58	53	162	79	45
OstreB-f	GGATTTTGCTGAGAACGAA	<i>Ostreococcus</i> clade B	20	58	45	223	81	48
OstreM-f	CCCTTAACGAGGATCCATTGG	<i>Ostreococcus mediterraneus</i>	21	58	52	196	83	56
OstreL-f	TGGTGAAGGAAGGTGTGCTT	<i>Ostreococcus lucimarinus</i>	20	58	50	183	79	45
Bath682-f	TCTGTACTGCGTTGTCTTGA	<i>Bathycoccus</i>	20	58	53	197	79	48
MicroA-f	CGTGCTCTGGGTTAACGAC	<i>Micromonas</i> clade A	20	59	55	179	79	46
MicroB-f	ACGCGCTCTGGCTTCA	<i>Micromonas</i> clade B	16	58	62	185	80	46
MicroC-f	GACGCGCTCTGGGCTTAATC	<i>Micromonas</i> clade C	20	61	60	184	80	47
Mam391-f	ACGCGCTCTGGGCTTAA	<i>Mamiella</i>	18	58	61			
Mant417-f	GTGCACTGGCTGGTCTCA	<i>Mantoniella</i>	18	58	61	198	80	42
Crus3273-f	ACTGGCGGGCGTCATCTT	<i>Crustomastix</i>	18	58	61	196	80	47
Dol3274-f	CGCAGTTCTGGCCTTTGTT	<i>Dolichomastix</i>	19	58	53	167	80	46
Mono255-f	ACTGGTCGGCGCCGTCT	<i>Monomastix</i>	17	58	71	214	80	51
Mami-r	TACGAATGCCCCAACTGT	Mamiellophyceae	19	58	53			
OstreM-r	CTCCGCTCCGTGAGGA	<i>Ostreococcus mediterraneus</i>	17	60	70			
MicroA-r	GAATGCCCCAACTGTTCC	<i>Micromonas</i> clade A	19	59	58			
MicroB-r	GAAATACGAATGCCCCAACT	<i>Micromonas</i> clade B	21	58	48			

ii) Tests de spécificité

Les tests de spécificité ont été réalisés sur des souches en culture (Tableau 2). Afin de déterminer l'affinité des différents couples d'amorces pour leur cible, nous avons fait varier la température d'hybridation, la concentration d'amorces et la concentration d'ADN cible ou non cible. Les réactions d'amplification PCR ont été réalisées avec un Mastercycler (Eppendorf®). Le mix optimal utilisé (volume final : 25 μ L), inclue, 1,5 μ L $MgCl_2$ 25 mM ; 0,5 μ L dNTP (20 mM) ; 0,5 μ L de chaque amorce (10 mM) ; 5 μ L GreenTaq 5X ; 0,2 μ L GoTaq 5u. μ L⁻¹ ; eau ultra pure QSP au volume final. Le programme PCR consiste en 5 min à 95°C, puis 35 cycles avec 1 min à 95°C ; 20 s à 58°C et 40 s à 72°C, suivis d'une étape de 2 min à 72°C. Chaque amplification a été contrôlée sur gel 0,8% d'agarose contenant 0,05 % de BET.

Une fois la spécificité des amorces contrôlée, une étape d'optimisation a été réalisée afin de connaître la concentration optimale d'amorces pour atteindre une efficacité maximale. Pour ce protocole de qPCR, nous avons opté pour une quantification basée sur le SYBR Green, ainsi le mix qPCR de 10 μ L a été élaboré à partir du SYBR Green PCR core reagents (Applied Biosystems®) : 1 μ L SYBR Green 10X ; 0,8 μ L dNTP 12,5mM ; 0,1 μ L ampEraseUNG 1u. μ L ; 0,05 AmpliTaq Gold 5U/ μ L ; 1,2 $MgCl_2$ 25 mM ; 5,45 μ L d'eau ultra pure.

1. Application sur échantillons naturels

L'ADN extrait des communautés naturelles concentrées sans tri cellulaire a été utilisé en qPCR, chacun des échantillons préparés a été placé en plaque de 96 puits (MicroAmp EnduraPlate Optical 96-well ; Applied Biosystems®) en triplicat avec quatre conditions de dilution : 0, 10, 100, 1000. Pour chaque plaque, un témoin négatif composé d'ADN non cible, un blanc sans ADN et un témoin positif composé d'une gamme d'ADN cible ont été ajoutés afin de contrôler le bon fonctionnement de l'amplification. Le programme de qPCR consiste en une étape d'activation de l'AmpErase UNG, 2 min à 50°C (hydrolyse

des amplicons formés pendant la préparation de la plaque), une étape d'activation de la polymérase 10 à 95°C puis 40 cycles avec 1 min à 95°C ; 20 s à 58°C et 40 s à 72°C. L'obtention de courbe de dissociation suit le dernier cycle, et va de 60°C à 95°C, permettant de visualiser la qualité de la réaction de qPCR.

III Traitement des séquences

1) Les virus isolés

Les séquences obtenues par séquençage Sanger ont été traitées avec le logiciel Bioedit (<http://www.mbio.ncsu.edu/Bioedit/bioedit.html>). Les chromatogrammes ont d'abord été analysés afin d'éliminer les fragments de mauvaise qualité en 3' et 5'. Les fragments sens et anti-sens ont été alignés à l'aide du programme ClustalW. Les positions indéterminées ont été résolues grâce aux chromatogrammes correspondant aux séquences. Les séquences complètes ont ensuite été obtenues en combinant les deux fragments sens et anti-sens.

a. Les communautés environnementales

Le séquençage d'amplicon provenant de l'amplification des communautés virales et eucaryotes par la technologie MiSeq Illumina produit un nombre extrêmement élevé de séquences (environ 18 millions de *reads* pour les gènes viraux et 21 à 24 millions de *reads* pour les communautés eucaryotiques). Nous avons donc été dans l'obligation d'automatiser le nettoyage, la suppression des séquences de faible qualité ainsi que l'assignation des séquences.

b. Les communautés virales

Les *reads* provenant du séquençage des amplicons viraux ont été traités grâce au logiciel USEARCH 7 (<http://www.drive5.com/usearch/>). Succinctement, suite à la suppression des séquences de faible qualité en utilisant une fenêtre mobile, l'utilisation du programme AmpliconNoise a permis la suppression des chimères engendrées par l'amplification des gènes. Par la suite, les séquences sens et anti-sens ont été associées avec un chevauchement d'environ 75 à 100 bp. Ensuite, les séquences ont été regroupées au sein d'unités taxonomiques opérationnelles (OTUs : *Operational Taxonomic Units*) avec 99 % de similarité afin de garder une bonne échelle résolutive. Une nouvelle étape de détection des chimères a ensuite été réalisée en utilisant le programme UCHIME implémenté dans USEARCH. Enfin, la table des OTUs a été créée, comprenant le nombre de séquences relatif à l'OTU ainsi que la séquence centroïde. Le détail du traitement réalisé sur ces jeux de données est résumé dans les Annexes 1 et 2.

Avant l'assignement taxonomique des OTUs par *blast* (*blastn*) avec la base de référence, un blast de la base de référence contre elle-même a été réalisé pour connaître les seuils d'identité intra et inter-groupe avec une longueur d'alignement minimale de 95%. Ceci a permis d'ajuster le seuil de confiance pour l'assignation d'un OTU à un groupe particulier.

2) Les communautés eucaryotes

Les *reads* provenant du séquençage des amplicons eucaryotes (V4 et V9) ont été traités à l'aide du logiciel MOTHUR (<http://www.mothur.org/>). Succinctement, suite à la suppression des *reads* de faible qualité ou de taille incorrecte, nous avons formé des *contigs* en combinant les *reads* sens et anti-sens, puis les séquences uniques ont été identifiées et alignées. Après suppression des séquences uniques ne s'alignant pas correctement avec l'ensemble des séquences, l'utilisation de UCHIME a permis d'identifier et de supprimer les chimères. Le détail du traitement réalisé sur ces jeux de

données est résumé dans les Annexes 3 et 4. Une classification a ensuite été effectuée en utilisant la base de données d'ADNr 18S de la Station Biologique de Roscoff « PR2 ». Pour finir, les séquences ont été regroupées à un seuil de 99% en OTUs et comptabilisées par station d'échantillonnage.

IV Mesures des variables environnementales

La majorité des variables environnementales (Tableau 6), à l'exception de l'ammonium, des variables satellitaires et de l'anomalie de densité, a été mesurée en suivant les protocoles mis en place par le réseau SOMLIT (<http://somlit.epoc.u-bordeaux1.fr/fr/>). La différence entre les mesures en mer et dans les lagunes se situe au niveau de la température, de la salinité et du pH. Les données provenant de SOLA et MOLA ont été acquises grâce à une sonde CTD alors qu'elles ont été mesurées à l'aide d'un thermomètre, un salinomètre (Figure 5) et un pHmètre pour les sites lagunaires. La calibration des appareils a été vérifiée régulièrement afin de contrôler la différence de mesure entre ces deux techniques. Les concentrations en nitrate, nitrite, phosphate et silicate ont été mesurées par dosage colorimétrique, et les concentrations en chlorophylle *a* et phéophytine *a* ont été dosées par fluorimétrie. Les concentrations en carbone et en azote organique particulaire ont été mesurées par spectromètre de masse.

Tableau 6 : Ensemble des descripteurs mesurés

Descripteurs	Unité
Profondeur	m
Temperature	°c
Salinité	μM
Ammonium	μM
Nitrate	μM
Nitrite	μM
Phosphate	μM
Silicate	μM
Oxygène	mL.L ⁻¹
Chlorophylle a	μg.L ⁻¹
Phéophytine a	μg.L ⁻¹
Carbone organique particulaire (POC)	μg.L ⁻¹
Azote organique particulaire (PON)	μg.L ⁻¹
Matière en suspension (MES)	mg.L ⁻¹
Turbidité	m
pH	
Abondance des Picoeucaryotes (PPE)	cell.mL ⁻¹
Abondance des Nanoeucaryotes (PNE)	cell.mL ⁻¹
Abondance des Cyanobacteries	cell.mL ⁻¹
Abondance des cryptophyceae	cell.mL ⁻¹
Vitesse du vent	m.s ⁻¹
Vitesse du vent moyen sur les 7 jours précédents	m.s ⁻¹
Vitesse du vent sur le vecteur latitudinal	m.s ⁻¹
Vitesse du vent sur le vecteur latitudinal moyen sur les 7 jours précédents	m.s ⁻¹
Vitesse du vent sur le vecteur longitudinal	m.s ⁻¹
Vitesse du vent sur le vecteur longitudinal moyen sur les 7 jours précédents	m.s ⁻¹
Radiation photosynthétique active (PAR)	W.m ²
Radiation photosynthétique active (PAR) moyen sur les 7 jours précédents	W.m ²
Pression atmosphérique	Pa
Pression atmosphérique moyen sur les 7 jours précédents	Pa
Anomalie de densité	

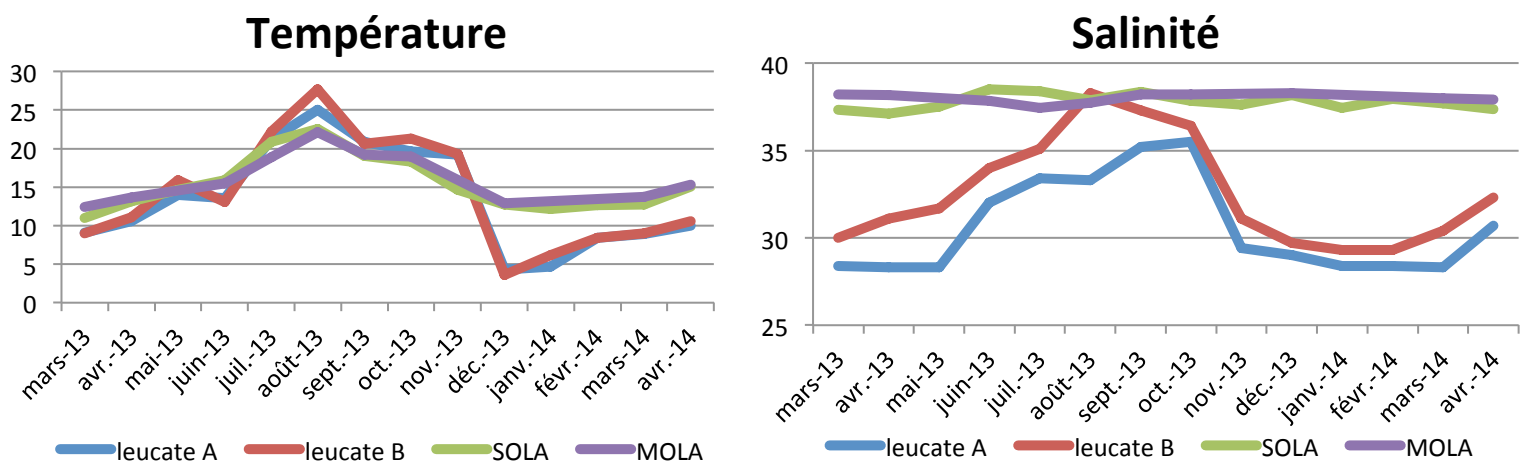


Figure 5 : Evolution temporelle de la salinité et de la température dans les 4 sites échantillonnés

1) Mesure de l'ammonium

La mesure de l'ammonium est basée sur le principe de mesure de la fluorescence après l'ajout d'un réactif (orthophtaldialdéhyde). Le résultat obtenu est ensuite comparé à une gamme étalon de salinité fixe, elle-même adaptée à l'échantillonnage réalisé. Pour tenir compte de la salinité variable dans les environnements lagunaires, lors de chaque prélèvement une gamme étalon a été réalisée par ajout de concentrations connues de NH_4 dans des échantillons prélevés. La concentration totale d'ammonium a ensuite été mesurée en tenant compte de cette gamme étalon.

2) Variables satellitaires

Ces données ont été extraites du modèle ERA-interim, un modèle océanique et atmosphérique qui assimile en permanence des données d'observation afin d'être ajusté. Il fait partie de l'*European Centre for Medium-Range Weather Forecasts* (ECMWF, <http://www.ecmwf.int/>). Nous avons extrait de ce modèle les données de pression atmosphérique (Pa), température de surface (SST), direction et force du vent (vecteur U : longitudinal ; vecteur V : latitudinal ; wind speed m.s^{-1}).

3) Anomalie de densité

Les anomalies de densité ont été calculées en suivant le model ESOS 80 (Fofonoff et Millard, 1983) permettant de calculer la densité de l'eau de mer en tenant compte de la pression atmosphérique, de la température ainsi que de la salinité.

V Analyses statistiques et multidimensionnelles

Les analyses statistiques et multidimensionnelles ont été réalisées grâce au logiciel R v3.2 (R Development Core Team. 2015. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria (<https://www.R-project.org/>)).

L'ensemble des sorties graphiques de R a été exporté en Portable Document Format (PDF) et modifié grâce au logiciel de traitement vectoriel Inkscape® (<https://inkscape.org/fr/>) et au logiciel de traitement d'image GIMP® et Adobe Illustrator® (<http://www.gimp.org/> ; <http://www.adobe.com/fr/products/illustrator.html>)

1) Analyses statistiques

La normalité des données a été testée grâce au test de Shapiro-Wilk (package `{stats}`, fonction `shapiro.test`). Lors du rejet de l'hypothèse nulle de normalité ou de l'absence de significativité après transformations, des tests non paramétriques ont été effectués afin d'estimer la corrélation entre les variables, principalement grâce au coefficient de corrélation de Spearman (package `{stat}`, fonction `cor.test`). Les analyses de différence entre deux moyennes ou plusieurs moyennes pour les données non-normales ont été réalisées respectivement par un test de Wilcoxon (package `{stat}`, fonction `wilcox.test`) et un test de Kruskal-Wallis (package `{stat}`, fonction `kruskal.test`).

a. Méthodes de groupement

Afin d'estimer la ressemblance entre les différentes stations d'échantillonnage, nous nous sommes servis des différentes variables (environnementales ou d'OTU) afin de

calculer les distances entre les objets. Dans le cas des variables environnementales qui étaient de nature continue, une normalisation a été faite sur ces variables en utilisant la fonction *decostand* (package *{vegan}*), puis les distances Euclidiennes ont été calculées grâce à la fonction *vegdist* (package *{vegan}*). Dans le cas des variables « espèce » représentant les abondances normalisées des OTUs viraux ou eucaryotes, la dissimilarité de Bray-Curtis a été employée (fonction *vegdist*, package *{vegan}*). Les groupements ont ensuite été effectués en utilisant la méthode « *Unweighted pair group method using arithmetic average* » (UPGMA), permettant un groupement par association moyenne (package *{stats}*, fonction *hclust*).

2) Analyses multidimensionnelles

Très brièvement, ces analyses permettent de résumer, en 2 voire exceptionnellement en 3 dimensions d'un plan factoriel maximisant la variance, la corrélation entre les différentes variables et la position relative des différents objets du jeu de données.

Des analyses en composantes principales (ACP) ont été réalisées pour les variables quantitatives normalisées par la fonction *decostand* (package *{vegan}*, grâce à la fonction *PCA* (package *{FactoMineR}*). Ces analyses ont permis de comprendre les différences entre les stations d'échantillonnage sur la base des variables environnementales.

Des analyses factorielles de correspondances (AFC) ont été réalisées sur les tableaux de contingence (fréquences des OTUs) après normalisation de ce nombre de *reads* par la somme des *reads* observés à chaque point. Ces analyses ont été effectuées avec la fonction *cca* (package *{vegan}*) et ont permis de comprendre les relations existantes entre les différents objets du jeu de données en les projetant dans un espace factoriel commun. Cependant, dans ce cas, les variables environnementales ne contraignent pas le plan factoriel, elles sont utilisées uniquement en tant que variables descriptives.

Afin de comprendre comment ces variables contraignent les relations dans le jeu de données, nous avons utilisé des analyses canoniques des correspondances (CCA) permettant de sélectionner une combinaison de variables maximisant la dispersion des objets ou variables, c'est-à-dire participant à la mise en place du plan factoriel. Un

modèle simplifié en terme de variable a donc été mis en place ne prenant en compte que les variables les plus explicatives par rapport au modèle complet utilisant une régression multiple comprenant l'intégralité de nos variables. La sélection des variables participant au plan factoriel a été fait grâce à la fonction *ordistep* (package *{vegan}*). Une fois le modèle simplifié, une analyse de variance, testée par 999 permutations, a été utilisée pour vérifier l'hypothèse que notre modèle explique une part significative de la variation (fonction *anova*, package *{vegan}*). Nous avons réalisé la CCA grâce à la fonction *cca* (package *{vegan}*).

Une analyse de la co-structure entre les variables environnementales et les variables espèces (OTU) a également été réalisée dans le but de connaître les relations existantes entre les différentes stations dans les différents environnements. Pour résumer brièvement la méthode appelée STATICO (Chessel *et al.*, 2004), il s'agit d'une analyse triadique partielle (PTA, Thioulouse et Chessel, 1987) d'une série de k tableaux provenant d'une analyse de co-inertie effectué sur chaque tableau. Elle permet de résumer dans un même espace de compromis la structure de chaque jeu de données et permet ainsi la comparaison inter et intra-site. Elle a été réalisée grâce aux fonctions *Withinpca*, *dudi.pca*, *wca*, *katab.within* et *statico* (package *{ade4}*).

Chapitre III – Interplay between *Micromonas* species and their viruses in the Western English Channel

Résumé



Les virus de *Micromonas* possèdent des propriétés en lien avec les caractéristiques des hôtes



Leurs dynamiques sont congruentes avec les variations d'abondances des différents clades de *Micromonas*



Certaines souches virales sont inféodées au clade C alors que les virus du clade A et B possèdent une gamme d'infectivité plus grande, autorisant potentiellement un changement d'hôte entre les différents clades

*Cet article présente une partie de mon travail effectué lors de mon stage de Master 2. Bien que je ne sois pas premier auteur de cette publication, nous avons jugé pertinent de l'ajouter à ce travail de thèse car il fournit des informations complémentaires sur les dynamiques des prasinovirus en lien avec les communautés phytoplanctoniques de *Micromonas*.*

Participation du doctorant

- ✓ Recherche bibliographique
- ✓ Test de spécificité
- ✓ Cinétique de lyse
- ✓ Extraction d'ADN
- ✓ Séquençage et analyses génétiques

Ce chapitre est présenté sous la forme d'un article publié dans *Environmental Microbiology Reports*. A-C. Baudoux, H. Lebretonchel, H. Dehmer, M. Latimier, R. Edern, F. Rigaut-Jalabert, P. Gei, L. Guillou, E. Foulon, Y. Bozec, T. Cariou, Y. Desdevises, E. Derelle, N. Grimsley, H. Moreau, and N. Simon. 2015. Interplay between *Micromonas* species and their viruses in the Western English Channel. *Environmental Microbiology Reports* 7(5) : 765-773

Interplay between the genetic clades of *Micromonas* and their viruses in the Western English Channel

A.-C. Baudoux,^{1*} H. Lebretonchel,² H. Dehmer,¹ M. Latimier,¹ R. Edern,¹ F. Rigaut-Jalabert,³ P. Ge,¹ L. Guillou,¹ E. Foulon,¹ Y. Bozec,¹ T. Cariou,³ Y. Desdevises,² E. Derelle,² N. Grimsley,² H. Moreau² and N. Simon¹

¹Sorbonne Universités, UPMC Univ Paris 06, CNRS, Adaptation et Diversité en Milieu Marin (AD2M UMR7144), Station Biologique de Roscoff, 29680 Roscoff, France

²Sorbonne Universités, UPMC Univ Paris 06, CNRS, Biologie Intégrative des Organismes Marins (BIOM), Observatoire Océanologique, 66650 Banyuls sur Mer, France

³Sorbonne Universités, UPMC Univ Paris 06, CNRS, Fédération de Recherche (FR2424), Station Biologique de Roscoff, 29680 Roscoff, France

Summary

The genus *Micromonas* comprises distinct genetic clades that commonly dominate eukaryotic phytoplankton community from polar to tropical waters. This phytoplankton is also recurrently infected by abundant and genetically diverse prasinoviruses. Here we report on the interplay between prasinoviruses and *Micromonas* with regard to the genetic diversity of this host. For 1 year, we monitored the abundance of three clades of *Micromonas* and their viruses in the Western English Channel, both in the environment using clade-specific probes and flow cytometry, and in the laboratory using clonal strains of *Micromonas* clades to assay for their viruses by plaque-forming units. We showed that the seasonal fluctuations of *Micromonas* clades were closely mirrored by the abundance of their corresponding viruses, indicating that the members of *Micromonas* genus are susceptible to viral infection, regardless of their genetic affiliation. The characterization of 45 viral isolates revealed that *Micromonas* clades are attacked by specific virus populations, which exhibit distinctive clade specificity, life strategies and genetic diversity. However, some viruses can also cross-infect different host clades, suggest-

ing a mechanism of horizontal gene transfer within the *Micromonas* genus. This study provides novel insights into the impact of viral infection for the ecology and evolution of the prominent phytoplankton *Micromonas*.

Introduction

Viruses are undoubtedly the most abundant biological entities in the ocean, and they intimately interact with every facet of the marine biosphere. Through these interactions, viruses profoundly influence the global biogeochemical cycles by altering the structure and the function of marine communities and by contributing to the cycling of major elements (for review, see Suttle, 2005; 2007; Sime-Ngando, 2014). In spite of their global-scale implications, the nature and the dynamics of virus–host interactions in marine environments are far from understood, and many basic, yet fundamental, questions remain unsolved: Who are the hosts infected by marine viruses? What are the infection strategies evolved by viruses? How do virus–host interactions vary in time and space? In this respect, characterizing the interplay of ecologically relevant virus–host model systems is a prerequisite for advancing our understanding of virus impacts in nature.

Viruses that infect the cosmopolitan green picoalga *Micromonas* were first reported by (Pienaar, 1976), and they were found to co-occur with their hosts in a wide range of marine environments, from temperate to cold waters (e.g. Cottrell and Suttle, 1991; 1995a; Sahlsten, 1998; Zingone *et al.*, 1999). The majority of known *Micromonas* viruses belong to the genus *Prasinovirus* in the Phycodnaviridae family. Virions are tailless; they exhibit icosahedral capsids of 100–125 nm in diameter and dsDNA genome of approximately 200 kb (Mayer and Taylor, 1979; Cottrell and Suttle, 1991; Zingone *et al.*, 1999). Yet these prasinoviruses have variable tolerance to chloroform (Martinez Martinez *et al.*, 2015), suggesting ultrastructural divergence, and they exhibit significant level of genetic variation (Cottrell and Suttle, 1995b). These viruses display a marked seasonal dynamics (Sahlsten, 1998; Zingone *et al.*, 1999), and they were shown to cause considerable amount of mortality in their host population (Cottrell and Suttle, 1995a; Evans *et al.*, 2003). *Micromonas* strains show a great variability with respect to their susceptibility to infection, indicating that

Received 13 February, 2015; revised 13 May, 2015; accepted 4 June, 2015. *For correspondence. E-mail acbaudoux@sb-roscoff.fr; Tel. (+33)2 9829 2537; Fax (+33)2 9829 2324.

© 2015 Society for Applied Microbiology and John Wiley & Sons Ltd

766 A.-C. Baudoux et al.

viruses do not only quantitatively but also qualitatively regulate their host populations (Sahlsten, 1998; Zingone *et al.*, 1999; 2006). So far, the infection patterns and processes underlying these complex interactions are poorly understood. Of particular interest, the relation between the susceptibility to infection and the genetic diversity of *Micromonas* has seldom been investigated (Zingone *et al.*, 2006; Martinez Martinez *et al.*, 2015).

The ubiquitous genus *Micromonas* is genetically diverse and comprises an assemblage of three (Guillou *et al.*, 2004) to five (Šlapeta *et al.*, 2006; Worden, 2006) discrete phylogenetic clades (or lineages) of flagellated cells that correspond to at least three distinct species (N. Simon, unpublished). These clades are often sympatric in marine ecosystems (Šlapeta *et al.*, 2006; Foulon *et al.*, 2008). Yet their relative contribution to total *Micromonas* abundance varies in time and space, suggesting that they occupy specific niches (Foulon *et al.*, 2008). The factors that regulate this clade dynamics are not clearly understood. In all likelihood, *Micromonas* clades exhibit differential responses to abiotic factors, but they might also respond differently to predation risks, including those imposed by viruses.

To test whether *Micromonas* clades display differential susceptibility to virus infection, we combined field and laboratory experiments on 45 novel viral isolates in order to characterize the interactions between the genetic clades of *Micromonas* and their viruses in the Western English Channel (WEC) throughout the year 2009. The WEC constitutes an ideal study site where *Micromonas* is known to dominate the picophytoplankton community (Not *et al.*, 2004). The three main *Micromonas* genetic types, designated as clades [A.ABC.12], [B.E.3] and [C.D.5] (Worden, 2006) and here referred to as clades A, B and C, respectively, are recorded year-round in this ecosystem (Foulon *et al.*, 2008). Our study revealed that *Micromonas* clades interact with specific viral populations that display distinctive dynamics and life history traits.

Results and discussion

Physicochemistry at the sampling station

Strong tidal mixing produces a permanently mixed water column at the long-term monitoring coastal station SOMLIT-Astan (48°46'N, 3°58'W, Marrec *et al.*, 2013). During the sampling period (February–December 2009), the water temperature progressively increased from 8.8°C in February to a maximum of 16.9°C in September, while salinity varied between 34.8 (March) and 35.3 (September). Nutrient dynamics varied according to the classical pattern observed in this area (Not *et al.*, 2004). Phosphate and nitrate concentrations showed comparable dynamics with minima recorded in summer (0.16 and 0.7 µM,

respectively) and maxima during the winter period (0.53 and 11 µM, respectively). Detailed dynamics of biogeochemical variables at SOMLIT-Astan are publicly available on the observatory website (<http://somalit-db.epoc.u-bordeaux1.fr/bdd.php?serie=ST>).

Dynamics of Micromonas clades in the WEC

The phytoplankton community at SOMLIT-Astan during this period was numerically dominated by picophytoplanktonic cells, which comprised, on average, 90% of the total community in 2009 (data not shown). Among these, the abundance of the picoeukaryotes varied between 2.1 and 16.5 × 10³ cells ml⁻¹ with *Micromonas* spp. accounting for 20–80% of the total counts as shown by a combination of flow cytometry (FCM) and TSA-FISH analyses (Fig. 1A). As reported previously, this prominent genus displayed marked seasonal dynamics, with major peaks of abundance recorded in late spring (6.6 × 10³ cells ml⁻¹), early summer (6.3 × 10³ cells ml⁻¹) and mid-autumn (4.4 × 10³ cells ml⁻¹) (Not *et al.*, 2004; Foulon *et al.*, 2008).

The three main *Micromonas* genetic clades (hereafter referred to as clades A, B and C) exhibit recurring seasonal dynamics in the WEC according to previous time series (Foulon *et al.*, 2008). We confirmed this pattern of succession in 2009 (Fig. 1B). *Micromonas* clade A numerically dominated the bloom from February to mid-May, comprising 63–77% of the total counts. From June to December, both genetic types A and B equally contributed to *Micromonas* counts, although clade B cells showed a sudden, yet unexplained, drop in abundance on June 16. Cells that belonged to clade C were the least abundant, accounting, on average, for 7 ± 6% of the total counts with somewhat higher contribution during the winter period (14%) than the remainder of the year (5%).

The ecological processes that influence the shifts in the relative abundance of *Micromonas* clades are unclear. It is traditionally believed that global-scale distribution and diversity of species are mostly driven by the ability to withstand abiotic controls (Wiens, 2011). Previous study on the global distribution of *Micromonas* clades suggested that clade B tends to thrive in warmer coastal waters, whereas clade C mostly occurs during low-light conditions (Foulon *et al.*, 2008). At local spatial extents, it is however well accepted that species distribution is also influenced by biotic controls (Wiens, 2011). It is thus likely that besides ecophysiological adaptations, biotic interactions also regulate the observed *Micromonas* clade dynamics in the WEC.

Dynamics of Micromonas viruses in the WEC

Viruses that infect *Micromonas* (hereafter referred to as MicV), mainly prasinoviruses, have been observed and

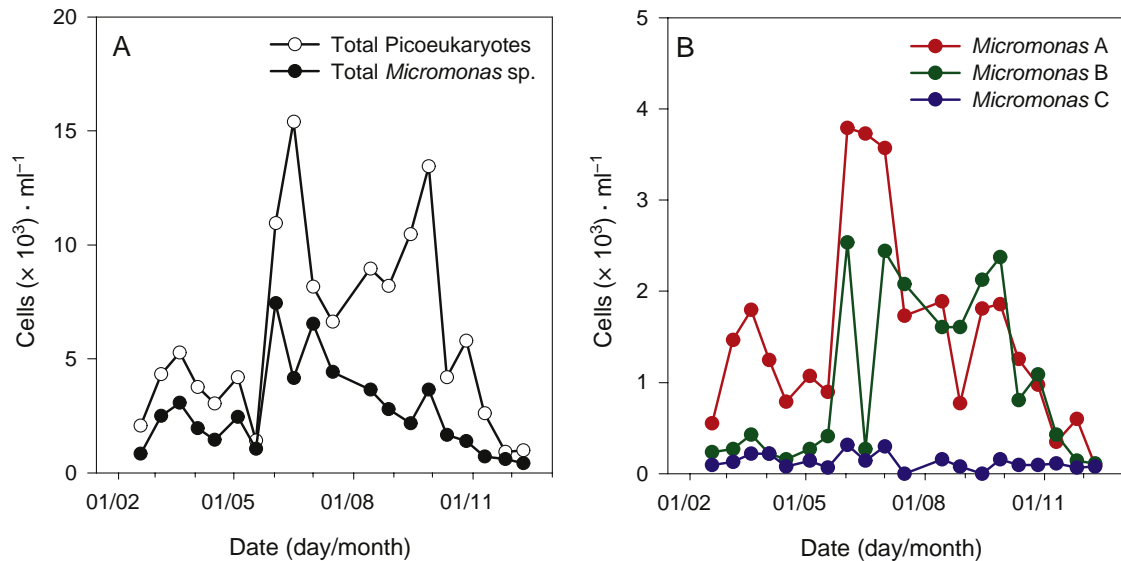


Fig. 1. Dynamics of (A) total picoeukaryotes and *Micromonas* cells and (B) *Micromonas* clades at the station SOMLIT-Astan. *Micromonas* clade names A, B and C correspond to those of Guillou and colleagues (2004). Picoeukaryote abundance was determined by flow cytometry on 3 μm prefiltered and fixed samples (formaldehyde 1% final concentration). *Micromonas* abundance (total and clade-specific) was assessed by FISH-TSA, as described in Foulon and colleagues (2008). *Micromonas* taxon-specific probes MICROA01, MICROB01 and MICROC01 and MICRO01 targeted *Micromonas* clade A.ABC.12, clade B.E.3, clade C.D.5 (Worden, 2006; Foulon *et al.*, 2008), and all clades respectively.

isolated from different oceanic regions (Cottrell and Suttle, 1991; 1995a; Sahlsten, 1998; Zingone *et al.*, 1999; Short and Short, 2008). Flow cytometry analyses of cultured isolates showed that these prasinoviruses consistently cluster within a well-defined population based on their nucleic acid fluorescence (upon Sybr green I staining) and side scatter properties (Fig. S1). In the WEC, this specific population accounted for a small but significant proportion (on average 10%) of the total virus counts and displayed a marked seasonal dynamics with concentration ranging from $3.4 \times 10^5 \text{ ml}^{-1}$ in February to $8.8 \times 10^5 \text{ ml}^{-1}$ in July 2009 (Fig. 2A). This FCM population probably does not exclusively comprise *Micromonas* viruses, but counts of this viral population covary strikingly well with *Micromonas* abundance ($r = 0.675$, $n = 20$, $P < 0.005$), suggesting that viruses might perhaps control abundance of this genus during the sampling period.

To further investigate this hypothesis, we quantified infectious MicV in the WEC by plaque assay using cultures of each of the three *Micromonas* clades. Significant differences in the number of plaque-forming units (PFU) were detected depending on the phylotype of the host culture (Fig. 2B). As also recorded for their hosts, PFU obtained on *Micromonas* clades A and B (hereafter, PFU-A and PFU-B, respectively) were substantially more abundant than PFU formed on clade C hosts (hereafter PFU-C). During the sampling period, PFU-A abundance increased concomitantly with the development of *Micromonas* clade A. Their maximum abundance ($8 \times 10^2 \text{ virus ml}^{-1}$) was observed 2 weeks after the peak of host abundance (Figs 1B and 2B).

The monitoring of PFU-B showed low concentration during the winter–spring period, but their dynamics was tightly coupled to their host abundance from the summer period until the end of the year, reaching a peak of $11 \times 10^2 \text{ virus ml}^{-1}$ in autumn. Regarding PFU-C, infectious particles were detected year-round, but their abundance was 10–100-fold lower than PFU-A and PFU-B with no clear relation to their specific host dynamics. Because of variable strain susceptibilities to viral infection, it is difficult to compare our results with former PFU monitoring on *Micromonas* host cell clades, although maximal abundance was within the same range ($1.0\text{--}4.6 \times 10^3 \text{ PFU ml}^{-1}$, Cottrell and Suttle, 1991; Sahlsten, 1998; Zingone *et al.*, 1999). For the same reason, this assay most likely underestimates the global abundance of viruses that infect each *Micromonas* phylotype. Nonetheless, it provides unequivocal evidence that *Micromonas*, regardless of their genetic clade, were subject to viral attack year-round in the WEC.

Micromonas virus–host interactions

The above results led us to question whether specific virus populations are associated to *Micromonas* clades, and if so how these virus populations interact with their hosts. To tentatively address this question, 45 viruses (14–16 per host clade) were isolated throughout the sampling period and they were characterized. These viruses were selected randomly from a collection of 176 viral isolates obtained from plaques on *Micromonas* clades A, B and C along the year 2009. Regardless of the host on

768 A.-C. Baudoux et al.

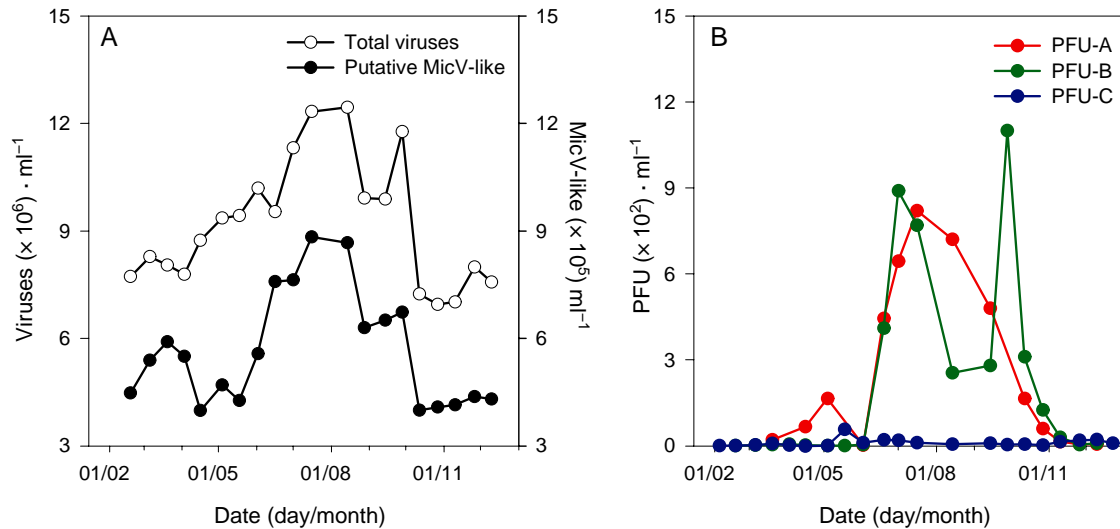


Fig. 2. Dynamics of (A) putative *Micromonas* virus (MicV-like) and (B) plaque-forming units (PFU) formed on *Micromonas* clades at the station SOMLIT-Astan throughout the year 2009. The abundance of total virus and MicV-like was determined by flow cytometry on glutaraldehyde-fixed samples (0.5% final concentration) as described in Brussaard (2004). MicV-like population was determined based on the flow cytometric signature of *Micromonas* virus isolates maintained in the laboratory (Fig. S1). The PFU dynamics was monitored by combining natural seawater with individual cultures of the three *Micromonas* clades (10–20% v/v), namely RCC451 (clade A), RCC828 and RCC829 (clade B), and RCC497 and RCC834 (clade C). Each mixture was complemented with a commercial antibiotic mixture (Sigma-Aldrich P4083) and incubated at room temperature for 30 min. The sample was mixed with cooled molten agarose (0.1% w/v final concentration) and spread in a sterile Petri dish. The plaque assay was incubated under the host culture conditions for 10 days.

which they were isolated, MicV isolates exhibited sizeable icosahedral capsids (110–130 nm in diameter) as determined by negative staining using transmission electron microscopy, and large dsDNA genome, as suggested by their FCM signature upon Sybr green I staining (Fig. S1). We therefore assigned our isolates to the virus family Phycodnaviridae and the genus *Prasinovirus*.

Host specificities. The host range of these 45 prasinoviruses was then determined by pairwise infection using 14 strains belonging to the three *Micromonas* clades (Table 1, Table S1). The large majority of our isolates (35 MicVs) were strictly clade-specific. They infected host strains that belonged only to the clade of their initial host. The remaining 10 viruses could propagate at comparable yield on hosts belonging to clades A and B. Six viruses (out of 15) isolated on a clade A strain could also infect clade B strains, and four viruses (out of 14) isolated on a clade B strain could also infect clade A strains. Within each main pattern of specificity (A, B, AB and C), a remarkable clonal diversity was observed with 26 unique specificity profiles that included variable virus infection types ranging from generalist to specialist.

In order to test the structure of this infection network, we applied a network-based analysis according to Flores and colleagues (2011) and Weitz and colleagues (2013). As observed previously for phage–bacteria infection networks (Flores *et al.*, 2011), the structure of *Micromonas*–virus interactions was statistically modular. This statistical

analysis applied at the whole matrix scale and intra-module scale discriminated three modules, which comprised the viruses isolated on hosts from clade A, clade B and clade C respectively (Fig. S2). Altogether, these results suggest that *Micromonas* clades A, B and C interact with distinct, yet diversified, viral populations. The observed variability in strain specificity indicates that viruses influence the intraspecific diversity within each *Micromonas* clade. We also showed that host switches can occur between the more closely related host clades (those that belong to clades A and B), suggesting that viruses could also promote *Micromonas* diversity through horizontal gene transfers within this genus.

Virus infection strategies. One-step growth experiments conducted on four to five representative viruses per host clade indicated a similar clade-specific grouping of our isolates. This assay provided estimates of the virus burst size and latent period, which can be related to the virus life strategy. By analogy to the theory of the r- and K-selection (MacArthur, 1967), opportunistic (the most r-selected) viruses are those that exhibit short generation times and high burst sizes, whereas less virulent (the most K-selected) viruses are those that induce low mortality in their host population (Suttle, 2007) and they propagate through latent or chronic infection as observed for specific *Ostreococcus* viruses (Thomas *et al.*, 2011). Based on these parameters, we detected divergent life strategies in *Micromonas* viruses depending on the host

Table 1. Host range of 45 *Micromonas* viruses isolated from SOMLIT–Astan throughout the year 2009.

host	Virus	MicV-A	MicV-B	MicV-C
A	451	○	○	○
A	803	○	○	○
A	658	○	○	○
A	299	○	○	○
B	828	○	○	○
B	829	○	○	○
B	418	○	○	○
B	1109	○	○	○
C	497	○	○	○
C	834	○	○	○
C	629	○	○	○
C	465	○	○	○
C	373	○	○	○
C	114	○	○	○
	4240	○	○	○
	4241	○	○	○
	4242	○	○	○
	4243	○	○	○
	4244	○	○	○
	4245	○	○	○
	4246	○	○	○
	4247	○	○	○
	4248	○	○	○
	4249	○	○	○
	4250	○	○	○
	4251	○	○	○
	4252	○	○	○
	4253	○	○	○
	4254	○	○	○
	4255	○	○	○
	4256	○	○	○
	4257	○	○	○
	4258	○	○	○
	4259	○	○	○
	4260	○	○	○
	4261	○	○	○
	4262	○	○	○
	4263	○	○	○
	4264	○	○	○
	4265	○	○	○
	4266	○	○	○
	4267	○	○	○
	4268	○	○	○
	4225	○	○	○
	4224	○	○	○
	4226	○	○	○
	4227	○	○	○
	4228	○	○	○
	4229	○	○	○
	4230	○	○	○
	4231	○	○	○
	4232	○	○	○
	4233	○	○	○
	4234	○	○	○
	4235	○	○	○
	4236	○	○	○
	4237	○	○	○
	4238	○	○	○
	4239	○	○	○

Micromonas viruses are ordered chronologically according to the host clade on which they were isolated. *Micromonas* strains are listed according to their genetic types based on 18S rRNA and ITS sequences (Table S1). Viruses were isolated from plaque assays and they were made clonal by repeated end-point dilution. For the pairwise infection, freshly produced MicV isolate was added to exponentially growing algal cultures (MOI 10). Cultures that underwent lysis were inspected for viral proliferation using flow cytometry (Brussaard, 2004). ■, infection; ○, no infection; – : not determined.

clades that they attack (Fig. 3). The viruses isolated on clade B hosts tended to be the most virulent with short latent period and relatively high progeny production, regardless of their infection range types (i.e. specialist or generalist), whereas the viruses isolated on clade C strains tended to be the least virulent with generally extended latent period (up to 30 h) and moderate to high burst size (Fig. 3). Viruses isolated on clade A hosts exhibited an intermediate phenotype with moderate latent period and low burst sizes.

In contrast to past theories [trade-offs hypothesis (Poulin, 1998)], generalism (i.e. viruses that propagate on several host clades) was not associated with any compensatory effect on the burst size or on the latent period. The observed differences in viral infection kinetics did not correlate to host growth rates (0.6–0.9 d⁻¹, Fig. S3). However, viral infection strategies appeared to be related to the dynamics of their respective host clade in the studied area. For example, *Micromonas* clade C is present year-round at low abundance. It is thus tempting to speculate that viruses specific to this host clade had to evolve long latent periods to be maintained within the host population and fairly high progeny production to increase the probability of encounter with their host (Abedon *et al.*, 2001). By contrast, *Micromonas* clade B displays a marked seasonal dynamics in the WEC, with high abundance recorded between mid-June and end-September. We could thus expect that corresponding viruses had to evolve an opportunist lifestyle to efficiently exploit their transient resources.

Phylogenetic relationships between Micromonas viruses

To investigate the genetic diversity among our viral isolates and to determine how this diversity relates to the virus biological properties, we amplified a fragment of the gene encoding their DNA polymerase (*polB*, Chen and Suttle, 1996; Table S2). This sequence is well conserved among Phycodnaviridae, yet there is sufficient sequence variability to build phylogenetic relationships between viruses from the same genus (Bellec *et al.*, 2009; Clerissi *et al.*, 2014). In spite of their affiliation to the Phycodnaviridae family, the amplification of the *polB* gene using the AVS1/AVS2 primers failed for 8 of the 45 selected isolates, including B-, C- and AB-specific viruses. The reason for this failure is unclear. We cannot exclude that the sequence targeted by the degenerated primers might be too divergent. It is also possible that these polymerases carry an intein that makes the polymerase chain reaction product too long to amplify easily under the tested conditions (Clerissi *et al.*, 2013).

The partial *polB* sequences of the remaining 37 isolates fell into 28 haplotypes (Fig. 4). In most cases, viral isolates that belonged to the same haplotype exhibited

770 A.-C. Baudoux et al.

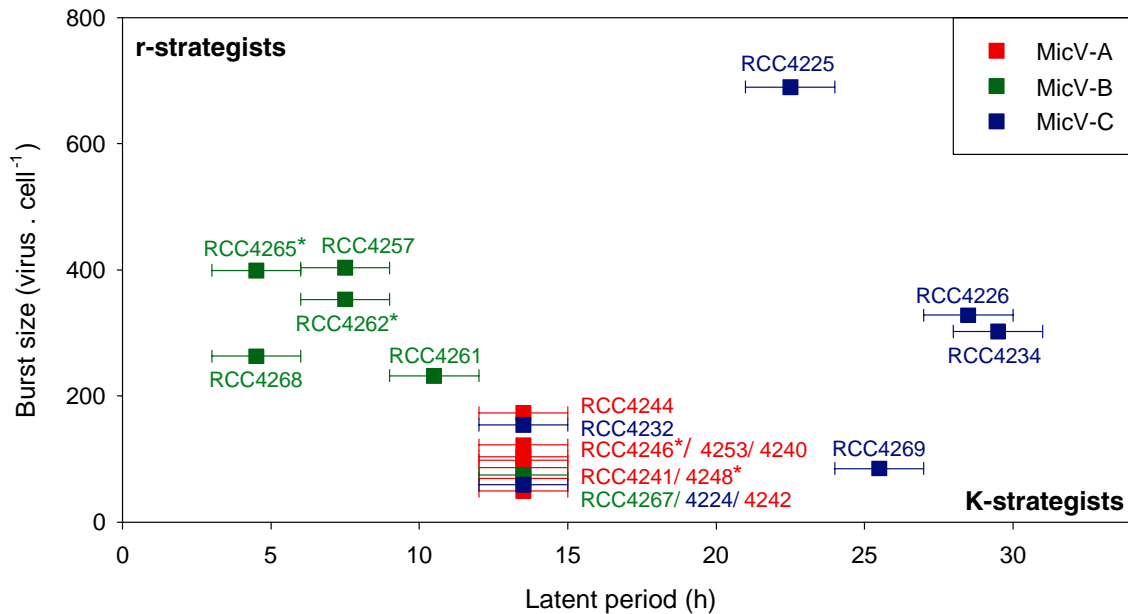


Fig. 3. Growth characteristics of representative *Micromonas* viruses isolated at SOMLIT-Astan in 2009. The asterisk indicates generalist virus isolate (those able to cross-infect different host clades). For each experiment, *Micromonas* host cultures were growing exponentially at μ_{max} ranging from 0.6 to 0.9 d⁻¹. Host cultures (1×10^6 ml⁻¹) were infected with a fresh MicV lysate (1×10^7 ml⁻¹). The samples were incubated under host culture conditions and sampled every 4 h for 90 h. *Micromonas* host and virus abundance was monitored by flow cytometry. Algal samples were analysed directly upon sampling, whereas virus samples were fixed with glutaraldehyde (0.5% final concentration) and stored at -80°C until analysis according to Brussaard (2004). The latent period was defined as the time period until extracellular release of the virions. The virus burst size was determined by calculating the ratio of the net increase in viral abundance to the maximum net decline in host abundance.

different specificity patterns, suggesting that *polB* sequencing underestimates the functional diversity of phycodnaviruses as reported previously (Baudoux and Brussaard, 2005; Bellec *et al.*, 2009; Clerissi *et al.*, 2012). In the phylogenetic trees, *polB* sequences of viruses infecting clade C hosts formed a well-supported cluster, while viruses that infect clade A and clade B hosts distributed into multiple branches, suggesting a higher level of genetic diversity. Interestingly, regardless of the host strain they infect, viruses that were highly specific (that infect one single strain among all strains tested) fell apart from other viruses (i.e. RCC4232 and RCC4240, 4242, 4243, Fig. 4 and Table 1). The prasinoviruses that infect *Ostreococcus tauri* showed a similar phylogenetic pattern and tended to cluster with counterparts that exhibit similar infection range types (Clerissi *et al.*, 2012).

This recurrent correlation between *polB* phylogeny and the number of hosts infected by a given prasinovirus is somewhat puzzling. The *polB* gene is indeed used as a neutral marker of evolution in Phycodnaviridae (Chen and Suttle, 1996; Dunigan *et al.*, 2006), that is this gene is well conserved among this virus family, whereas specificity markers are expected to be more variable genes (Clerissi *et al.*, 2012). Furthermore, virus–host specificity is typically determined during events upstream the virus DNA replication. Several studies reported that the virus attach-

ment on host cell surface determines the susceptibility of marine plankton to infection (e.g. Tarutani *et al.*, 2006; Mizumoto *et al.*, 2007; Stoddard *et al.*, 2007). In other words, the evolution of resistance to viral infection is mostly due to changes in host surface properties that limit viral attachment. Recently, Thomas and colleagues (2011) evidenced an intriguing mode of resistance in the prasinophyte *O. tauri*. This study demonstrated that prasinoviruses could still attach on resistant strains, but they did not produce viral progeny. It is thus likely that the resistance mechanism was not due to a change in the host receptor but rather occurred at the DNA entry or DNA replication stage, in which the *polB* gene product would be involved. Alternatively, viral attachment and internalization might be a two-stage process as in reoviruses (Reiter *et al.*, 2011). The notion that DNA replication machinery might be involved in host specificity is of course highly speculative, but the repeatedly observed link between *polB* clustering and prasinovirus infectivity range certainly deserves appropriate investigation.

Concluding remarks

In the WEC, a complex assemblage of viruses is associated with the dominating picophytoplankton species *Micromonas* spp. Our study revealed, for the first time,

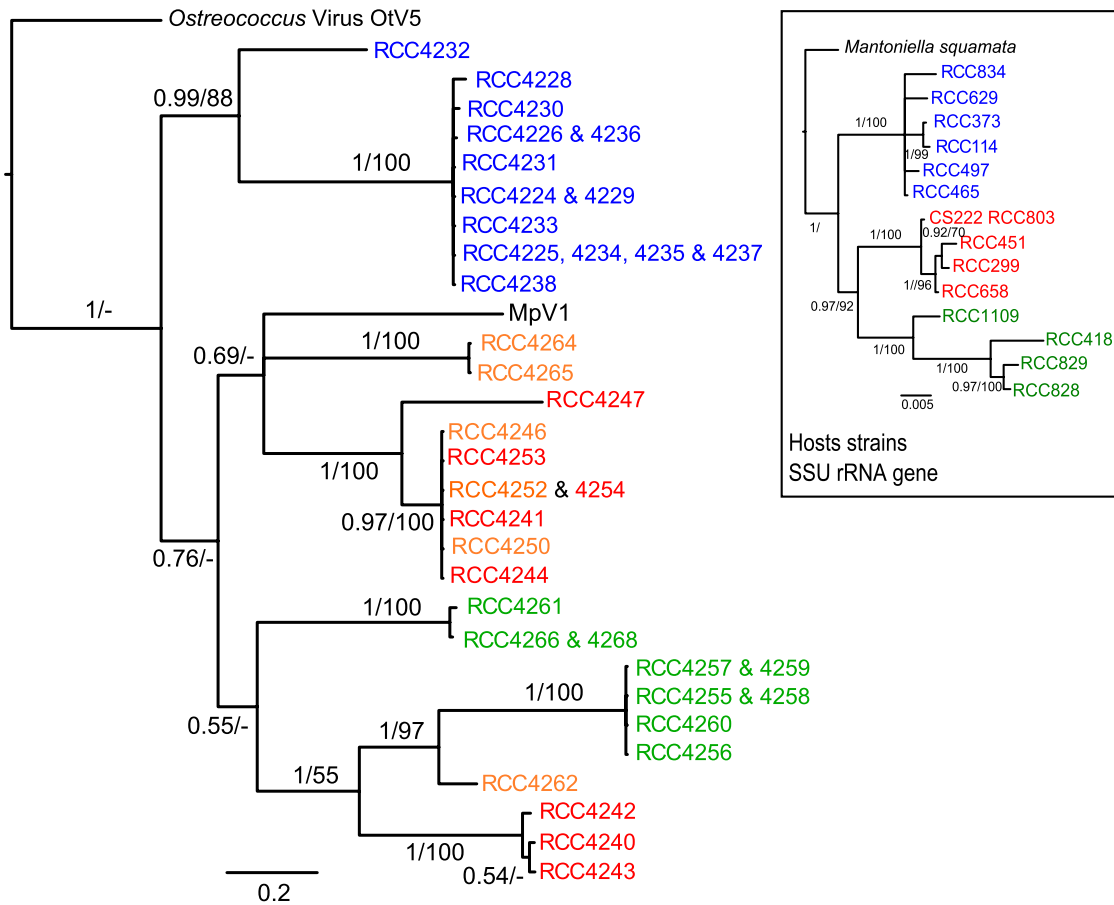


Fig. 4. Phylogenetic relationships between *Micromonas* viruses (unique sequences) inferred from a partial sequence of DNA polymerase B (*polB*, 478 nt) and host strain phylogeny based upon partial 18s rDNA sequences (1769 nt). Colour codes for host strains correspond to clades A, B and C. For viral strains, colour codes correspond to their host specificities: red, clade A-specific virus; green, clade B-specific virus; orange, clade A- and B-specific virus; blue, clade C-specific virus. *polB* gene fragment was amplified from a fresh lysate using the set of degenerated primers AVS1 and AVS2 based on Chen and Suttle (1995). DNA fragments were sequenced in reverse and forward directions at the Genomic Core Facility (GENOMER) of the Station Biologique de Roscoff, France. DNA sequences were aligned using MUSCLE (Edgar, 2004), taking into account the codon structure for virus sequences. Evolutionary model was chosen with MODELTEST (Posada and Crandall, 1998). The phylogenetic tree was reconstructed using Bayesian inference (BI) and maximum likelihood (ML), respectively, with MRBAYES 3.2 (Ronquist *et al.*, 2012) and PAUP 4.0b10 (Swofford, 2003), using a GTR + G model with parameters relaxed across codon positions for virus sequences, and a GTR + I + G for host sequences. Bayesian analysis was carried out with four chains of 1 000 000 generations, trees sampled every 1000 generations, and burn-in 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. The ML tree was validated using 100 bootstrap replicates. *Ostreococcus* virus 5 (OtV5) was used as an outgroup for the virus tree, and *Mantoniella squamata* for the host tree, and the trees built using BI are shown (numbers on branches are posterior probabilities (BI)/bootstrap proportions in % (ML – values below 50 not shown). The accession numbers of MicV *polB* sequences were deposited in GenBank. Accession numbers are listed in Table S2.

that the main *Micromonas* genetic clades (clade A, clade B and clade C) are attacked by specific virus populations, which displayed distinctive dynamics and life history traits. The apparently high variability in *Micromonas* virus specificities suggests that viruses maintain a high genetic diversity within each of these clades. However, host switching can occur (particularly for hosts that belong to clades A and B), suggesting that virus could also influence *Micromonas* diversity through horizontal gene transfers. The complete genome sequencing of representative viruses and hosts should provide insightful information on the extent of these putative genetic exchanges. It is of

course critical to keep in mind that our model-based approach only reflects a partial picture of the actual diversity of virus–host interactions. Still, these approaches provide invaluable information to improve our understanding of the mechanisms of virus–host interactions, as well as the evolution and outcomes of these interactions in natural environments.

Acknowledgements

The Captain and the crew of R.V. Neomysis are acknowledged for their excellent assistance onboard. We thank the

772 A.-C. Baudoux et al.

'Service d'Observation en Milieu Littoral (SOMLIT), INSU-CNRS' from Station Biologique de Roscoff for the data on the biogeochemistry and biology at ASTAN. We are grateful to Sophie Lepage from the Microscopy core facilities, Morgan Perennou from the GENOMER core facilities at the Station Biologique de Roscoff and David Demory for stimulating discussions. We thank the two anonymous reviewers for their constructive comments on a previous version of this manuscript. This research was funded by the ANR Funding Agency (PICOVIR, Grant No. BLAN07-1_200218; and REVIREC, Grant No. 12-BSV7-0006-01).

References

- Abedon, S.T., Herschler, T.D., and Stopar, D. (2001) Bacteriophage latent-period evolution as a response to resource availability. *Appl Environ Microbiol* **67**: 4233–4241.
- Baudoux, A.C., and Brussaard, C.P.D. (2005) Characterization of different viruses infecting the marine harmful algal bloom species *Phaeocystis globosa*. *Virol* **341**: 80–90.
- Bellec, L., Grimsley, N., Moreau, H., and Desdevises, Y. (2009) Phylogenetic analysis of new Prasinoviruses (Phycodnaviridae) that infect the green unicellular algae *Ostreococcus*, *Bathycoccus* and *Micromonas*. *Environ Microbiol Reports* **1**: 114–123.
- Brussaard, C.P.D. (2004) Optimization of procedures for counting viruses by flow cytometry. *Appl Environ Microbiol* **70**: 1506–1513.
- Chen, F., and Suttle, C.A. (1995) Amplification of DNA polymerase gene fragments from viruses infecting microalgae. *Appl Environ Microbiol* **61**: 1274–1278.
- Chen, F., and Suttle, C.A. (1996) Evolutionary relationships among large double-stranded DNA viruses that infect microalgae and other organisms as inferred from DNA polymerase genes. *Virol* **219**: 170–178.
- Clerissi, C., Desdevises, Y., and Grimsley, N. (2012) Prasinoviruses of the marine green alga *Ostreococcus tauri* are mainly species specific. *J Virol* **86**: 4611–4619.
- Clerissi, C., Grimsley, N., and Desdevises, Y. (2013) Genetic exchanges of inteins between prasinoviruses (Phycodnaviridae). *Evol* **67**: 18–33.
- Clerissi, C., Grimsley, N., Ogata, H., Hingamp, P., Poulain, J., and 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.
- Cottrell, M.T., and Suttle, C.A. (1991) Wide-spread occurrence and clonal variation in viruses which cause lysis of a cosmopolitan, eukaryotic marine phytoplankton, *Micromonas pusilla*. *Mar Ecol Prog Ser* **78**: 1–9.
- Cottrell, M.T., and Suttle, C.A. (1995a) Dynamics of a lytic virus infecting the photosynthetic marine picoflagellate *Micromonas pusilla*. *Limnol Oceanogr* **40**: 730–739.
- Cottrell, M.T., and Suttle, C.A. (1995b) Genetic diversity of algal viruses which lyse the photosynthetic picoflagellate *Micromonas pusilla* (Prasinophyceae). *Appl Environ Microbiol* **61**: 3088–3091.
- Dunigan, D.D., Fitzgerald, L.A., and Van Etten, J.L. (2006) Phycodnaviruses: a peek at genetic diversity. *Virus Res* **117**: 119–132.
- Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**: 1792–1797.
- Evans, C., Archer, S.D., Jacquet, S., and Wilson, W.H. (2003) Direct estimates of the contribution of viral lysis and microzooplankton grazing to the decline of a *Micromonas* spp. population. *Aquat Microbial Ecol* **30**: 207–219.
- Flores, C.O., Meyer, J.R., Valverde, S., Farr, L., and Weitz, J.S. (2011) Statistical structure of host–phage interactions. *Proc Natl Acad Sci USA* **108**: E288–E297.
- Foulon, E., Not, F., Jalabert, F., Cariou, T., Massana, R., and Simon, N. (2008) Ecological niche partitioning in the picoplanktonic green alga *Micromonas pusilla*: evidence from environmental surveys using phylogenetic probes. *Environ Microbiol* **10**: 2433–2443.
- Guillou, L., Eikrem, W., Chrétiennot-Dinet, M.-J., Le Gall, F., Massana, R., Romari, K., et al. (2004) Diversity of picoplanktonic prasinophytes assessed by direct nuclear SSU rDNA sequencing of environmental samples and novel isolates retrieved from oceanic and coastal marine ecosystems. *Protist* **155**: 193–214.
- MacArthur, R.H. (1967) *The Theory of Island Biogeography*, Vol. 1. Princeton, New Jersey, USA: Princeton University Press.
- Marrec, P., Cariou, T., Collin, E., Durand, A., Latimier, M., Macé, E., et al. (2013) Seasonal and latitudinal variability of the CO₂ system in the Western English Channel based on voluntary observing ship (VOS) measurements. *Mar Chem* **155**: 29–41.
- Martinez Martinez, J., Boere, A., Gilg, I., van Lent, J.W.M., Witte, H.J., van Bleijswijk, J.D.L., and Brussaard, C.P.D. (2015) New lipid envelope-containing dsDNA virus isolates infecting *Micromonas pusilla* reveal a separate phylogenetic group. *Aquat Microbial Ecol* **74**: 17–28.
- Mayer, J.A., and Taylor, F.J.R. (1979) A virus which lyses the marine nanoflagellate *Micromonas pusilla*. *Nature* **281**: 299–301.
- Mizumoto, H., Tomaru, Y., Takao, Y., Shirai, Y., and Nagasaki, K. (2007) Intraspecies host specificity of a single-stranded RNA virus infecting a marine photosynthetic protist is determined at the early steps of infection. *J Virol* **81**: 1372–1378.
- Not, F., Latasa, M., Marie, D., Cariou, T., Vaulot, D., and Simon, N. (2004) A single species, *Micromonas pusilla* (Prasinophyceae), dominates the eukaryotic picoplankton in the Western English Channel. *Appl Environ Microbiol* **70**: 4064–4072.
- Pienaar, R.N. (1976) Virus-like particles in three species of phytoplankton from San Juan Island, Washington. *Phycologia* **15**: 185–190.
- Posada, D., and Crandall, K.A. (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**: 817–818.
- Poulin, R. (1998) Large-scale patterns of host use by parasites of freshwater fishes. *Ecol Letters* **1**: 118–128.
- Reiter, D.M., Frierson, J.M., Halvorson, E.E., Kobayashi, T., Dermody, T.S., and Stehle, T. (2011) Crystal structure of reovirus attachment protein $\sigma 1$ in complex with sialylated oligosaccharides. *PLoS Pathog* **7**: e1002166.

- Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., *et al.* (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* **61**: 539–542.
- Sahlsten, E. (1998) Seasonal abundance in Skagerrak-Kattegat coastal waters and host specificity of viruses infecting the marine photosynthetic flagellate *Micromonas pusilla*. *Aquat Microbial Ecol* **16**: 103–108.
- Short, S.M., and Short, C.M. (2008) Diversity of algal viruses in various North American freshwater environments. *Aquat Microbial Ecol* **51**: 13–21.
- Sime-Ngando, T. (2014) Environmental bacteriophages: viruses of microbes in aquatic ecosystems. *Front Microbiol* **5**: 355. doi: 10.3389/fmicb.2014.00355
- Šlapeta, J., López-García, P., and Moreira, D. (2006) Global dispersal and ancient cryptic species in the smallest marine eukaryotes. *Mol Biol Evol* **23**: 23–29.
- Stoddard, L.I., Martiny, J.B., and Marston, M.F. (2007) Selection and characterization of cyanophage resistance in marine *Synechococcus* strains. *Appl Environ Microbiol* **73**: 5516–5522.
- Suttle, C.A. (2005) Viruses in the sea. *Nature* **437**: 356–361.
- Suttle, C.A. (2007) Marine viruses – major players in the global ecosystem. *Nat Rev Microbiol* **5**: 801–812.
- Swofford, D.L. (2003) *PAUP: Phylogenetic Analysis Using Parsimony. Version 4*. Sunderland, MA, USA: Sinauer Associates.
- Tarutani, K., Nagasaki, K., and Yamaguchi, M. (2006) Virus adsorption process determines virus susceptibility in *Heterosigma akashiwo* (Raphidophyceae). *Aquat Microbial Ecol* **42**: 209–213.
- Thomas, R., Grimsley, N., Escande, M.L., Subirana, L., Derelle, E., and Moreau, H. (2011) Acquisition and maintenance of resistance to viruses in eukaryotic phytoplankton populations. *Environ Microbiol* **13**: 1412–1420.
- Weitz, J.S., Poisot, T., Meyer, J.R., Flores, C.O., Valverde, S., Sullivan, M.B., *et al.* (2013) Phage–bacteria infection networks. *Trends Microbiol* **21**: 82–91.
- Wiens, J.J. (2011) The niche, biogeography and species interactions. *Philos Trans R Soc Lond B Biol Sci* **366**: 2336–2350.
- Worden, A.Z. (2006) Picoeukaryote diversity in coastal waters of the Pacific Ocean. *Aquat Microbial Ecol* **43**: 165–175.
- Zingone, A., Sarno, D., and Forlani, G. (1999) Seasonal dynamics in the abundance of *Micromonas pusilla* (Prasinophyceae) and its viruses in the Gulf of Naples (Mediterranean Sea). *J Plank Res* **21**: 2143–2159.
- Zingone, A., Natale, F., Biffali, E., Borra, M., Forlani, G., and Sarno, D. (2006) Diversity in morphology, infectivity, molecular characteristics and induced host resistance between two viruses infecting *Micromonas pusilla*. *Aquat Microbial Ecol* **45**: 1–14.

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Fig. S1. Flow cytograms of (A) *Micromonas* virus RCC4265 and (B) natural virus community at SOMLIT-Astan in August 2009 using Sybr green I staining according to Brussaard (2004).

Fig. S2. Modularity analysis of *Micromonas* virus infection networks using (A) the complete virus infection matrix (45 viruses \times 18 hosts) and (B) infection matrix only including viruses isolated on host from clades A and B. HA, HB and HC correspond to host from clades A, B and C respectively. VA, VB and VC correspond to viruses isolated on host from clades A, B and C respectively.

Fig. S3. Relationship between viral infection kinetics and host growth rates during the one-step growth experiments.

Table S1. List of *Micromonas* cultures used for the determination of MicV host range.

Table S2. AVS1 and AVS2 primer sequences (Chen and Suttle, 1996).

Table S3. Accession numbers of MicV isolates *polB*.

Supplementary InformationTable S1: List of *Micromonas* cultures used for the determination of MicV host range

RCC N°	Other strain names	Clade	Isolation site	SSU rRNA accession number
451	CCMP492, IIE1	A	North Atlantic	AY955006
803	CS-222	A	Southern Ocean	AY955010
658	CS-170	A	Pacific Ocean	AY955009
299	NOUM17, CCMP2709	A	South Pacific Ocean	NC_013045
828	CCMP 1646, Mp1	B	Mediterranean Sea	KF501021.1
829	MpPart	B	Mediterranean Sea	KF501022.1
418	RD010614-71-1, DOURDUFF_71-1	B	English Channel	KF501026
1109	54.4.A, BCC23000	B	Mediterranean Sea	KF501031
497	BL_105-7, Blanès-105-7	C	Mediterranean Sea	KF501023.1
834	PLY27, CCMP 1545, UTEX LB991, CCAP LB 1965/4, MpUK	C	English Channel	AY954994
629	He010117-D1-D5, Klaus 2a	C	North Sea	KF501018
465	RA010613-65-8, ASTAN_65-8	C	English Channel	KF501028
373		C	Baltic Sea	KF501025
114	DW8, CCMP490	C	North Atlantic	AY955003

Table S2: AVS1 and AVS2 primer sequences (Chen and Suttle, 1996)

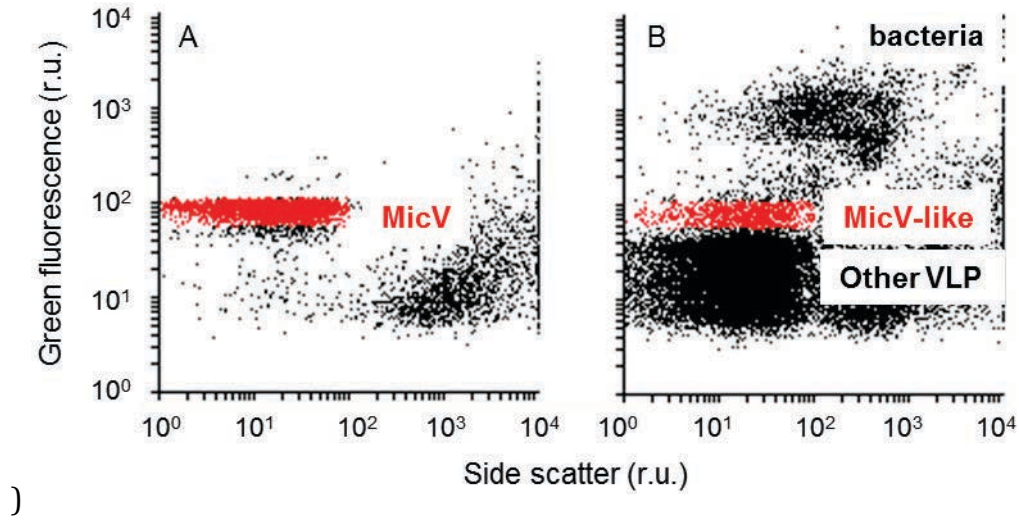
Primer	Sequence
AVS1	5'-GA(A/G) GGI GCI ACI GTI (T/C)TI GA(T/C) GC-3'
AVS2	5'-GCI GC(A/T) TAI C(G/T)(T/C) TT(T/C) TTI (G/C)(A/T)(A/G) TA-3'

Table S3: Accession numbers of MicV isolates *polB*

RCC N°	Host specificities	Isolation date	<i>polB</i> accession number
4224	C	5/02/2009	KP734132
4225	C	5/02/2009	KP734141
4226	C	5/02/2009	KP734135
4228	C	18/02/2009	KP734130
4229	C	02/03/2009	KP734133
4230	C	3/04/2009	KP734131
4231	C	3/04/2009	KP734137
4232	C	18/05/2009	KP734129
4233	C	04/05/2009	KP734134
4234	C	18/05/2009	KP734142
4235	C	02/06/2009	KP734139
4236	C	02/06/2009	KP734136
4237	C	02/06/2009	KP734138
4238	C	01/07/2009	KP734140
4240	A	24/03/2009	KP734164
4241	A	24/03/2009	KP734155
4242	A	07/04/2009	KP734163
4243	A	07/04/2009	KP734165

4244	A	16/04/2009	KP734152
4246	A	16/04/2009	KP734149
4247	AB	16/04/2009	KP734148
4250	AB	16/04/2009	KP734151
4252	AB	04/05/2009	KP734150
4253	AB	04/05/2009	KP734154
4254	A	04/05/2009	KP734153
4255	B	20/03/2009	KP734159
4256	B	20/03/2009	KP734160
4257	B	20/03/2009	KP734156
4258	B	20/03/2009	KP734158
4259	B	03/04/2009	KP734157
4260	B	03/04/2009	KP734161
4261	B	16/04/2009	KP734145
4262	AB	18/05/2009	KP734162
4264	AB	28/09/2009	KP734143
4265	AB	28/09/2009	KP734144
4266	B	12/10/2009	KP734147
4268	B	10/11/2009	KP734146

Figure Captions Supplementary Information



)
Fig S1: Flow cytograms of (A) *Micromonas* virus RCC4265 and (B) natural virus community at SOMLIT-Astan on August 2009 using Sybr GreenI staining according to Brussaard (2004

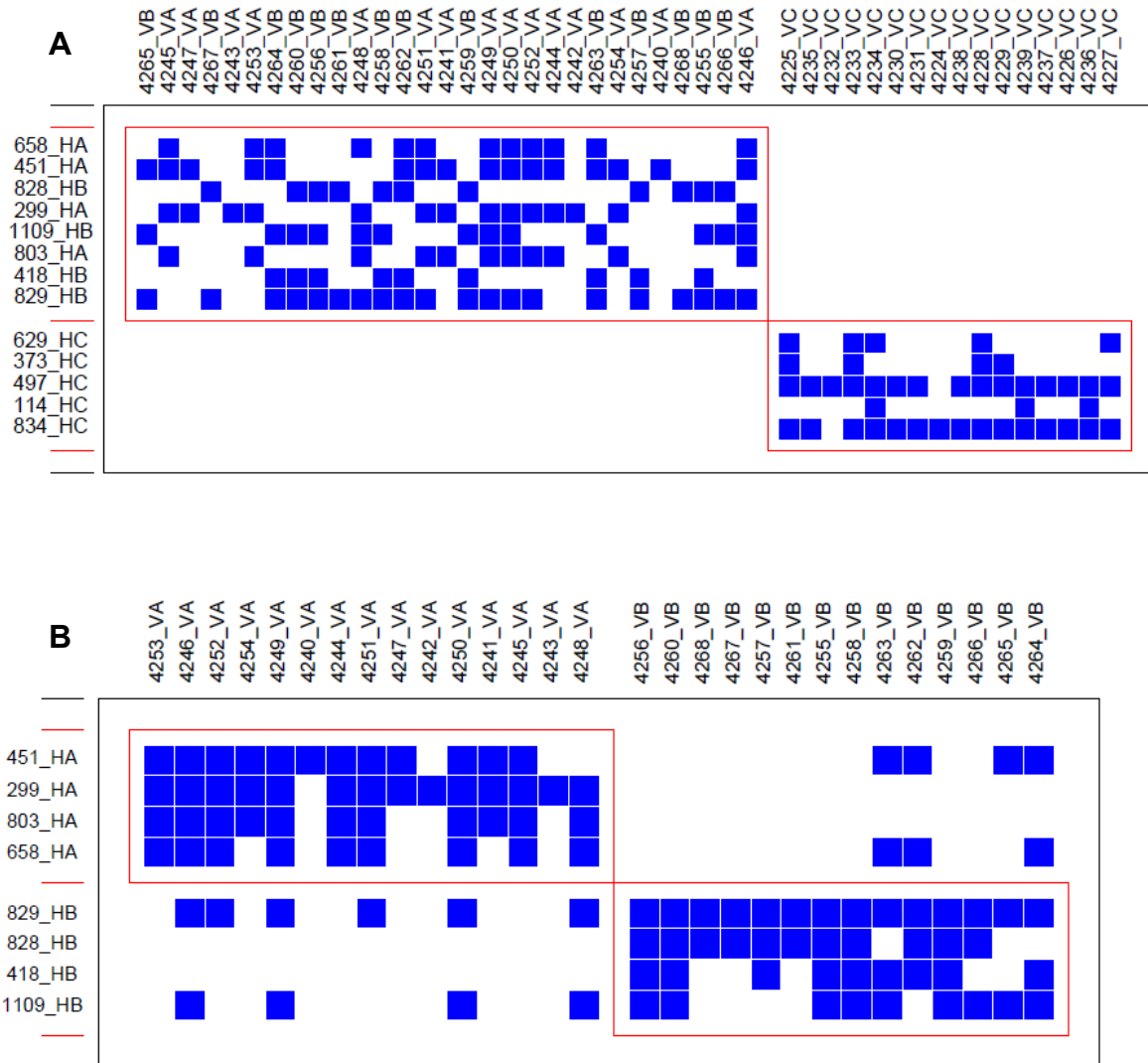


Figure S2: Modularity analysis of *Micromonas* virus infection networks using (A) the complete virus infection matrix (45 viruses x 18 hosts) (B) infection matrix only including viruses isolated on host from clades A and B. HA, HB, and HC correspond to host from clade A, B, and C, respectively. VA, VB, and VC correspond to viruses isolated on host from clade A, B, and C, respectively

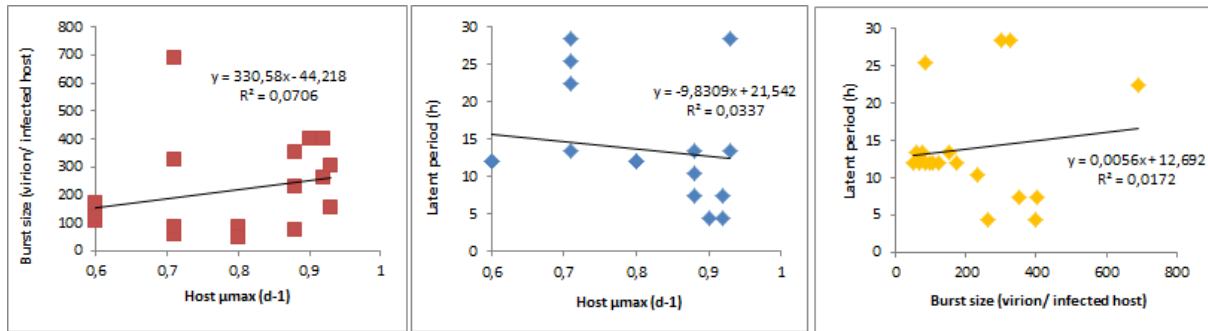


Figure S3: Relationship between viral infection kinetics and host growth rates during the one step growth experiments



Chapitre IV – Mise en valeur de l’abondance annuelle des Mamiellophyceae dans les communautés picoeucaryotiques au sein des lagunes par PCR quantitative

Chapitre IV - Green microalgae of the Mamiellophyceae dominate picoeukaryotic communities throughout the year in a coastal lagoon, as revealed by qPCR

Résumé



Les amorces conçues permettent de suivre 7 genres des Mamiellophyceae en qPCR



Les Mamiellophyceae sont dominants au sein de la fraction picoeucaryotique phytosynthétique



Ostreococcus tauri possède une dynamique forte au cours du temps dans les lagunes

Participation du doctorant

- ✓ Recherche bibliographique
- ✓ Mise en place d'un protocole d'échantillonnage
- ✓ Échantillonnage et préparation des échantillons
- ✓ Analyses de séquence
- ✓ Conception et mise au point du protocole de qPCR
- ✓ Rédaction

Ce chapitre est présenté sous la forme d'un article (soumis à Applied and Environmental Microbiology) Lebredonchel H, Ros M, Grimsley N, Desdevises Y. *Green microalgae of the Mamiellophyceae dominate picoeukaryotic communities throughout the year in a coastal lagoon, as revealed by qPCR.*

ABSTRACT

Photosynthetic picoeukaryotes generally represent a large proportion of phytoplanktonic eukaryotic communities in marine environments. Using a qPCR technique on samples from a Mediterranean lagoon and an oligotrophic offshore area (in the Gulf of Lion, France) from April 2013 to April 2014, we investigated the abundance and composition of photosynthetic eukaryotes (<20 µm), focusing on the Mamiellophyceae (Chlorophyta). To estimate abundances, primers were specifically designed to discriminate 7 genera of the Mamiellophyceae. We found that the Mamiellophyceae, mainly the genera *Micromonas*, *Ostreococcus* and *Bathycoccus*, accounted for 75% to 83% of the picoeukaryotic community. The Mamiellophyceae exhibited a higher abundance in the eutrophic lagoon than in the offshore station. Within the seasonal dynamic variations, *Micromonas* dominated the Mamiellophyceae population all around the year (57 to 87%) except during summer when *Ostreococcus* were extremely abundant and accounted for up to 77% of the picoeukaryotic community. In contrast, only a small part of nano-phytoplanktonic population was composed of Mamiellophyceae (3–20 µm). The absence of significant environmental differences between lagoon stations during this period suggests that the exchanges between lagoon and seawater play a significant role on picoeukaryote communities.

Key words: Photosynthetic picoeukaryotes, Mamiellophyceae, Quantitative PCR, 18S rDNA, community structure, Mediterranean eutrophic lagoon

INTRODUCTION

Picoeukaryotes from fresh and marine water ecosystems contribute significantly to picophytoplankton primary production and biomass (1,2) and they can contribute up to 80% in open oligotrophic water (3) and 87% in the English Channel offshore near Roscoff (4). Photosynthetic picoplankton are composed of prokaryotic and eukaryotic organisms less than 2-3 μm , are widely distributed and they show high genetic diversity (5,6). The prokaryotic fraction comprises mainly *Synechococcus* and *Prochlorococcus*. In contrast, since their discovery, many eukaryotic prototrophic organisms have been described, but their actual diversity is still a subject of ongoing investigations (7,8). In this particular group, the Mamiellophyceae (Chlorophyta) algae form a significant part of picoeukaryotes in marine water (6). They are classified in 8 genera, *Micromonas*, *Bathycoccus*, *Ostreococcus*, *Mantoniella*, *Mamiella*, *Crustomastix*, *Dolichomastix* and *Monomastix*, wherein only *Monomastix* contains freshwater species, (9). In marine coastal water, Mamiellophyceae are composed generally by three main picoeukaryotic groups, *Micromonas pusilla* (10), *Bathycoccus prasinos* (11) and *Ostreococcus* sp. (12), as shown by culture studies, optical and electronic microscopy observations and/or 18S rDNA for fluorescent in situ hybridization or gene sequencing analyses (4,13IN). These genera might each be considered to be complexes of cryptic species, because their small size hinders definition of morphological differences (9,16,17). However, in this class, the other known genera are nanoeukaryotic (3-20 μm , *Mantoniella*, *Mamiella*, *Crustomastix*, *Dolichomastix* and *Monomastix*) and are not dominant in this size fraction (9). Diatoms, dinoflagellates and haptophytes appear to dominate this 3-20 μm size fraction on continental shelves and are responsible for seasonal blooms in temperate and polar water (18,19).

Some authors have used Fluorescent in situ hybridization (FISH) to identify and estimate eukaryotic picoplankton abundance (20,21), but this method is technically demanding and should preferably include automation of microscopical analyses (22). Using this technique, these authors (21) showed that *Ostreococcus*, *Micromonas* and *Bathycoccus* accounted for 28% of the eukaryotes off the Spanish coast between 2001 and 2002.

Understanding the community composition is paramount for understanding ecosystem responses to short- or long-term perturbations (23). Quantitative PCR (qPCR) is an attractive approach because it permits quantitative estimation of the numbers of specific species or groups by estimating the number of DNA copies. The V4 18S rDNA region is particularly useful for screening diversity in this group because it contains enough variation to discriminate between most of the known species (24,25). Regarding phytoplanktonic communities, there are important differences between eutrophic and oligotrophic environments. One of them is the ratio between prokaryotic and eukaryotic abundances, which is higher in oligotrophic than eutrophic water (3,4). Picoeukaryotes, and particularly Mamiellophyceae, were found to be more abundant in nutrient-rich environments (26). Furthermore, *Ostreococcus tauri* was found to be very abundant in the Thau lagoon (27). The Gulf of Lion is an advantageous region for sampling, because it provides contrasting environments in a restricted area.

In this study, we developed a new set of primers (based on the V4 18S rDNA region) to monitor the abundance of genera in the Mamiellophyceae by qPCR in a Mediterranean lagoon and an oligotrophic offshore area (in the Gulf of Lion, France). Using reference strains for each of the main genera, we estimated genomic DNA content, determined specific primers and optimized qPCR assays. We monitored the distribution and abundance of Mamiellophyceae monthly during one year. Cytometric measurements provided an independent approximation for pico- and nano-eukaryote community abundances. Secondly, we monitored 13 environmental variables to better understand the ecological processes taking place inside this highly dynamic lagoon area. We present an overview of the phytoplanktonic community in this region with emphasis on the Mamiellophyceae, and ecological dynamics.

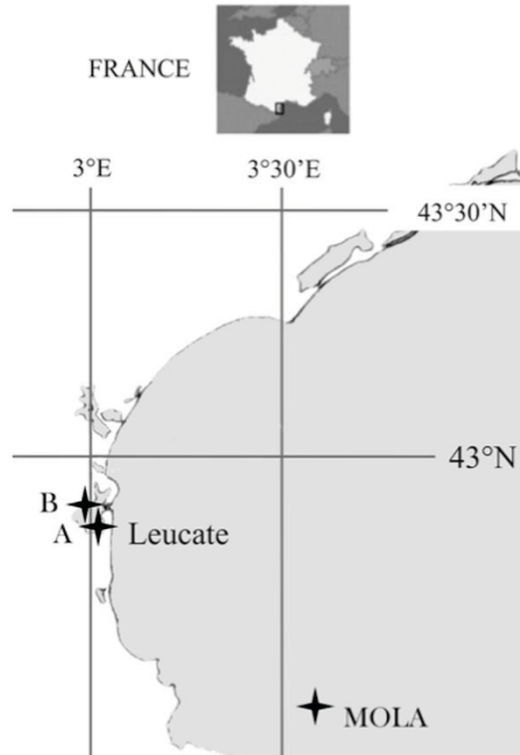


Figure 1. Locations of sampling sites.

Two locations are in the Leucate lagoon (A: 42° 8' 08" N, 3° 02' 26" E, B: 42° 8' 72" N, 2° 99' 8" E) and one is a marine offshore station (MOLA: 42° 27' N, 3° 32' E)

MATERIALS AND METHODS

2.1 Cultures

All cultures were grown in culture flasks (Starstedt) at 20°C under light-dark cycle of 12-12h, at 200 $\mu\text{E m}^{-2} \text{s}^{-1}$ light intensity. To maximize growth rate, strains were grown in three different media (K, L1 and C for freshwater alga *Monomastix minuta*) (Table 1) and subcultured regularly (every 7 days). Cell densities were measured by flow cytometry using a FACS Calibur (Becton Dickinson Biosciences) based on natural fluorescence. Cells were harvested for DNA extractions during the exponential growth phase. We

chose representative strains, to test primer specificity, to serve as templates and internal standards for the optimization and qPCR assays, and to estimate the quantity of cellular genomic DNA for each strain tested. The *Micromonas* genus is genetically diverse and includes three (14) to five (28) phylogenetic clades. In this study, we used three *Micromonas* clades (A, B, C) corresponding to clades [A.ABC. 12], [B.E.3] and [C.D.5] (28). When it was possible, we chose strains isolated from the Mediterranean Sea with known phylogenetic positions. All *Ostreococcus* strains came from this area, except *Ostreococcus* clade B (deep strain) which came from the Tropical Atlantic.

Table 1. Algal strains used in this study.

RCC: Roscoff Culture Collection, NCMA (formerly CCMP) - National Center for Marine Algae and Microbiota, USA; NIES: National Institute for Environmental Studies, Japan.

Order	Family	Genus	Species	Origin	Medium used	Name	Cell size (um)	size class
Mamiellales	Bathycoccaceae	<i>Ostreococcus</i>	sp. (Clade B)	Tropical Atlantic (+21° 2', -31° 8')	K	RCC 809	1	Picoeukaryote
		<i>Ostreococcus</i>	<i>tauri</i>	Mediterranean Sea (+42° 29', +3° 8')	K	RCC 1108	1	Picoeukaryote
		<i>Ostreococcus</i>	<i>mediterraneus</i>	Mediterranean Sea (+43° 3', +3° 0')	K	RCC 2572	1	Picoeukaryote
		<i>Ostreococcus</i>	<i>lucimarinus</i>	Mediterranean Sea (+43° 3', +3° 0')	K	CCMP2490	1	Picoeukaryote
		<i>Bathycoccus</i>	<i>prasinus</i>	North Sea (+54° 11', +7° 54')	K	RCC 682	1	Picoeukaryote
	Mamiellaceae	<i>Micromonas</i>	sp. Clade A	Pacific ocean (-22° 20', +166° 20')	K	RCC 299	1-2	Picoeukaryote
		<i>Micromonas</i>	sp. Clade B	Mediterranean Sea (+41° 40', +2° 48')	K	RCC 434	1-2	Picoeukaryote
		<i>Micromonas</i>	sp. Clade C	Baltic sea (+58° 11', +9° 6')	K	RCC 373	1-2	Picoeukaryote
		<i>Mantoniella</i>	<i>squamata</i>	North sea (+52° 58', +0° 39')	K	RCC 417	4	Nanoeukaryote
		<i>Mamiella</i>	sp.	Mediterranean Sea (+30° 8', +10° 3')	L1	RCC 391	4	Nanoeukaryote
Dolichomastigales	Crustomastigaceae	<i>Crustomastix</i>	<i>stigmata</i>	Mediterranean Sea (+40° 54', +14° 9')	L1	CCMP 3273	5	Nanoeukaryote
	Dolichomastigaceae	<i>Dolichomastix</i>	<i>tenuilepis</i>	Mediterranean Sea (+40° 30', +14° 15')	L1	CCMP 3274	5	Nanoeukaryote
Monomastigales	Monomastigaceae	<i>Monomastix</i>	<i>minuta</i>	Tschiura, Ibaraki (Japan)	C	NIES 255	8-10	Nanoeukaryote

2.2 Marine samples

Environmental samples were collected in three different locations (Figure 1) in the Gulf of Lion (Mediterranean Sea): two (A and B) from the Leucate lagoon (A: 42° 808'N, 3° 026'E, B: 42°872'N, 2°998'E), and one from a marine offshore station (MOLA: 42°27'N,

3°32'E). Samples were taken monthly from April 2013 to April 2014. Five litres of water were collected from the surface and strained through a 200 µm nylon mesh into black bottles. Samples were then pre-filtered through a 20 µm nylon mesh and concentrated down to 50 ml using a Vivaflow 200 (Sartorius) unit with a peristaltic pump. Glycinebetaine was added (6.2% final concentration) to protect the cell before freezing and storage at -80°C until DNA extraction. During sampling temperature, salinity, pH, oxygen, turbidity, quantity of suspended matter, particulate nitrogen and carbon and nutrients were measured.

2.3 Calibration cultures

To use qPCR for estimating cell abundance, an internal genomic DNA standard is required. Using the 5 members of Mamiellophyceae, we examined the relationship between cell numbers measured either by flow cytometry or by qPCR over a wide range of cell densities.

Based on cytometer measurements, we prepared in duplicate 100 ml of cultures of 10,000, 1000 and 100 cells per ml of 5 strains of Mamiellophyceae (RCC 1108, 417, CCMP 3273, 3274 and NIES 255) focusing on nano-eukaryotic cell which can have an higher 18S copy number. We centrifuged each sample 15 min at 8000 g, the supernatant was removed and pellets were stored at -80°C until DNA extraction.

2.4 DNA extraction

DNA from cultures and natural samples was extracted using a CTAB protocol (29). After cell counting by flow cytometry, cells were harvested by centrifugation for 15 min at 7500 g. 5 ml of CTAB buffer (preheated to 60°C) was added to the pellet with 100 µl of DTT (10mM) and gently mixed to disperse cells. After DNA re-suspension overnight at 4°C in DNA-free water, DNA sample concentrations were measured with a Nanodrop (Eppendorf) instrument and biophotometer (Eppendorf) and stored at -80°C.

2.5 Primer design

We designed clade- or species-specific primers to amplify the V4 region of the 18S rRNA gene using a recommended SYBR Green protocol. Sequence alignments were performed using Clustal software and examined visually using Bioedit (<http://www.mbio.ncsu.edu/bioedit/bioedit.html>). We identified three variable regions in the V4 region. For all targeted groups, we used the same reverse primer except for *Ostreococcus mediterraneus* and *Micromonas* clade A and B, where an internal sequence in the V4 region of 18S rDNA gene was used. For the forward primer, we selected a variable region at the beginning of the V4 region. We manually designed forward primers in this region with four constraints, (i) an amplicon size close to 150-200 bp, (ii) a 3' tail rich in A and T to avoid non-specific hybridization, (iii) a specific primer size close to 20 bases and (iv) a melting temperature (T_m) of 58°C. The theoretical specificity of the primers was checked using the NCBI tool Primer-Blast (<http://www.ncbi.nlm.nih.gov/tools/primer-blast/>).

2.6 qPCR assays

The final reaction volume was 10 μ l using SYBR Green PCR core reagents (Applied Biosystems) and containing 1 μ l SYBR Green 10X, 0.8 μ l dNTP 12.5 mM, 0.1 μ l ampErase UNG 1 U. μ l, 0,05ul AmpliTaq Gold 5 U. μ l, 1.2 μ l MgCl₂ 25 mM, 5.85 μ l ultrapure water and 2 μ l DNA (extracted from environment or culture and diluted 10, 100 and 1000-fold). Amplifications were done in MicroAmp EnduraPlate Optical 96-well (Applied Biosystems) plates and sealed with Optical Adhesive Covers (Applied Biosystems). All reactions were performed with StepOne Plus (Applied Biosystems) programmed with one step of 2 min at 50°C, allowing AmpErase UNG to hydrolyze PCR amplicon possibly carried over from previous reactions. An enzyme activation step (10 min at 95°C) followed the initial step. Forty cycles of 1 min denaturation at 95°C, 20 sec hybridization at 58°C and 40 sec elongation at 72°C were carried out. The melting curve analysis followed the last qPCR cycle. The dissociation curves were performed from 60 to 95°C. This melting curve analysis was necessary because SYBR Green binds all double-

stranded DNA without specificity. This step is required to check the reaction specificity. qPCR amplifications were done in triplicate for each concentration.

2.7 Primer specificity

The specificity was first checked by PCR using Gotaq Flexi DNA Polymerase (Promega) with Mastercycler Gradient (Eppendorf). Reaction mixes were prepared for a final volume of 25 μ l, containing 1.5 μ l 25mM MgCl₂, 0.5 μ l 20mM dNTP, 0.5 μ l 10 mM of primers, 5 μ l GreenTaq 5X, 0.2 μ l GoTaq 5 u. μ l⁻¹. Two μ l of genomic DNA (2ng/ μ l) extracted from culture strain was added as template. The PCR steps were performed with 5 min at 95°C, 35 cycles of 1 min at 95°C of denaturation, 20 sec of hybridization with gradient temperature from 55 to 61°C and 40 sec of elongation. Five μ l of the reaction product were migrated on 0.8% (w/v) agarose gel stained with ethidium bromide. The reaction product of the qPCR reaction was also checked on agarose gels following the previous protocol.

2.8 qPCR analysis of natural samples

After DNA extraction from 2 ml of natural concentrated seawater, DNA pellets were suspended in 50 μ l of ultrapure water. DNA was diluted by 10, 100 and 1000-fold to avoid amplification inhibition from non-target DNA sequences. qPCR assays were performed with the protocol described above with primer concentration appropriately adjusted (200 nM for *Ostreococcus*, *Micromonas*, *Bathycoccus* and 400 nM for the others).

2.9 Statistical analysis

We tested differences between lagoon stations based on all measured variables using Student's t-test implemented in R (<http://www.R-project.org>, package {stats}). If variance homogeneity was not respected the Welch correction was applied. One objective was to find out whether phytoplanktonic composition differences observed between lagoon stations were due to these variables or to other ecological factors.

RESULTS

3.1 DNA content

In contrast to DNA cell quantification using cytometry measurements, DNA contents from extracted cell were lower than that reported in the literature (Table 2). DNA contents measured for the picoeukaryotic Mamiellophyceae (between 20.7 and 55.5 fg per cell for *Ostreococcus*, *Bathycoccus* and *Micromonas*) were lower than for nanoeukaryotic cells such as *Mantoniella*, *Crustomastix*, *Dolichomastix* and *Monomastix* (from 221,2 to 337,8 fg per cell).

Table 2: Cell genomic DNA content.

Strain	Name	DNA content (fg)	SD	Genome size (Mb)	extrapolated (fg)* ¹
<i>Ostreococcus</i> sp.	RCC 809	25,8	6,1	13,3	27,2
<i>Ostreococcus tauri</i>	RCC 1108	24,0	5,1	12,6	25,8
<i>Ostreococcus lucimarinus</i>	RCC 802	20,7	5,7	13,2	27,0
<i>Bathycoccus prasinos</i>	RCC 682	29,3	2,5	15,1	30,9
<i>Micromonas pusilla</i>	RCC 299	43,5	5,0	20,9	42,7
<i>Micromonas pusilla</i>	RCC 434	52,5	7,7	-	-
<i>Micromonas pusilla</i> * ²	RCC 373	55,5	10,1	21,9	44,8
<i>Mantoniella squamata</i>	RCC 417	266,7	71,2	-	-
<i>Mamiella</i> sp. nov.	RCC 391	275,6	31,9	-	-
<i>Crustomastix stigmata</i>	CCMP 3273	278,1	61,5	-	-
<i>Dolichomastix tenuilepis</i>	CCMP 3274	221,2	22,9	-	-
<i>Monomastix minuta</i>	NIES 255	337,8	46,3	-	-

SD: standard deviation; *¹genome extrapolated: DNA content (fg) = (2 x genome size (b))/(0,978 x 10⁷), *² genome of CCMP1545

3.2. QPCR optimization and validation

We designed and tested several forward and reverse primers. Since we could not design specific primers to distinguish *Ostreococcus tauri* and *Ostreococcus* clade B (deep strain), due to the lack of sequence differences between the strains in the targeted region, we designed an additional forward primer specifically for this clade B (Table 3).

Table 3: Primers designed for this study

Primer Name	Sequence (5'>3')	Targeted group	Primer size (b)	Tm primer (°C)	Primer GC%	Amplicon size (bp)	Amplicon Tm (°C)	Amplicon GC%	primer concentration (nM)	efficacité (%)
OstreI-f	TGTGCTTCAGGGCCACTTA	<i>Ostreococcus tauri</i>	20	58	53	162	79	45	200	98,5
OstreB-f	GGATTTTGGCTGAGAACGAA	<i>Ostreococcus</i> clade B	20	58	45	223	81	48	200	99,5
OstreM-f	CCCTTAACGAGGATCCATTG	<i>Ostreococcus mediterraneus</i>	21	58	52	196	83	56	200	76,6
OstreL-f	TGGTGAAGGAAGGTGTGCTT	<i>Ostreococcus lucimarinus</i>	20	58	50	183	79	45	200	97,6
Bath682-f	TCTGTACTGGCTGTCTTGA	<i>Bathycoccus</i>	20	58	53	197	79	48	200	99,4
MicroA-f	CGTGCTCTGGGTTTAAAGAC	<i>Micromonas</i> clade A	20	59	55	179	79	46	200	106
MicroB-f	ACGGGCTCTGGCTTCA	<i>Micromonas</i> clade B	16	58	62	185	80	46	200	83,4
MicroC-f	GACGGGCTCTGGGCTTAATC	<i>Micromonas</i> clade C	20	61	60	184	80	47	200	89,6
Mam391-f	ACGGGCTCTGGGCTTAA	<i>Mamiella</i>	18	58	61					
Mant417-f	GACGGGCTCTGGGCTTAAAG	<i>Mantoniella</i>	20	61	60	184	80	42	400	98
Crus3273-f	ACTGGGGGGCGTCATCTT	<i>Crustomastix</i>	18	58	61	196	80	47	400	101,1
Dol3274-f	CGCAGTTCTGGGCTTTGTT	<i>Dolichomastix</i>	19	58	53	167	80	46	400	103,1
Mono255-f	ACTGGTCCGGGGCTCT	<i>Monomastix</i>	17	58	71	214	80	51	400	99,9
Mami-r	TACGAATGCCCCCAACTGT	Mamiellophyceae	19	58	53					
OstreM-r	CTCCGGCTCCGTGAGGA	<i>Ostreococcus mediterraneus</i>	17	60	70					
MicroA-r	GAATGCCCCCAACTGTTC	<i>Micromonas</i> clade A	19	59	58					
MicroB-r	GAATACGAATGCCCCCAAC	<i>Micromonas</i> clade B	21	58	48					

- f : Forward primer, -r : Reverse primer

We first assessed its specificity *in silico* against NCBI non-redundant database using primer-BLAST (Table 4). Reaction conditions were then optimized for the 13 primer sets by varying primer concentrations and temperatures for fluorescence measurements. We chose to use 3 mM MgCl₂ based on similar studies using qPCR on bacterial abundance (30,31).

Table 4: *in silico* control of primer specificity

Primer Name	Minimum number of mismatches to non-target group inside Mamiellophyceae for amplicon size between 120 and 250bp	Non-target Mamiellophyceae related to minimum mismatches	Number of mismatches to non-target groups outside Mamiellophyceae for amplicon size between 120 and 250bp
OstreT-f	Two central mismatches	<i>Ostreococcus clade B</i>	> 9
OstreB-f	Two 3' mismatches or more	<i>Ostreococcus tauri</i>	> 9
OstreM-f	Perfect match	None	> 9
OstreL-f	Perfect match	None	> 9
Bath682-f	Perfect match	None	> 9
MicroA-f	Perfect match	None	> 9
MicroB-f	Perfect match	None	7
MicroC-f	Two 3' mismatches	<i>Mantoniella</i> CCMP 1436 (antartica)	4
Mam391-f	Three 5' mismatches	<i>Micromonas</i> sp.	1
Mant417-f	Two 3' mismatches	<i>Micromonas</i> sp.	3
Crus3273-f	Perfect match	None	> 9
Dol3274-f	Perfect Match	None	6
Mono255-f	Perfect Match	None	> 9

Specificity tests on cultured strains shown that our primers were specific to their target-group (Table 5) except for the *Ostreococcus tauri* primer, which as expected amplified also *Ostreococcus* clade B and Mam391-f primer, which was not specific to Mamiella RCC 391 and amplified each *Micromonas* clade. However, the forward primer designed to monitor *Ostreococcus* clade B amplified only this strain, allowing us to improve the discrimination of *Ostreococcus tauri*.

Table 5. Primer specificity.

	Forward primers	OstreT-f	OstreB-f	OstreM-f	Ostrel-f	Bath682-f	MicroA-f	MicroB-f	MicroC-f	Mam391-f	Mant417-f	Crus3273-f	Dol3274-f	Mono255-f
Strains	Reverse primers	Mami-r	Mami-r	OstreM-r	Mami-r	Mami-r	MicroA-r	MicroB-r	Mami-r	Mami-r	Mami-r	Mami-r	Mami-r	Mami-r
<i>Ostreococcus tauri</i>	RCC 1108	+	0	0	0	0	0	0	0	0	0	0	0	0
<i>Ostreococcus</i> sp. Clade B	RCC 809	+	+	0	0	0	0	0	0	0	0	0	0	0
<i>Ostreococcus mediterraneus</i>	RCC 2572	0	0	+	0	0	0	0	0	0	0	0	0	0
<i>Ostreococcus lucimarinus</i>	CCMP 2490	0	0	0	+	0	0	0	0	0	0	0	0	0
<i>Bathycoccus prasinos</i>	RCC 682	0	0	0	0	+	0	0	0	0	0	0	0	0
<i>Micromonas</i> Clade A	RCC 299	0	0	0	0	0	+	0	0	+	0	0	0	0
<i>Micromonas</i> Clade B	RCC 434	0	0	0	0	0	0	+	0	+	0	0	0	0
<i>Micromonas</i> Clade C	RCC 373	0	0	0	0	0	0	0	+	+	0	0	0	0
<i>Mamiella</i> sp.	RCC 391	0	0	0	0	0	0	0	0	+	0	0	0	0
<i>Mantoniella squamata</i>	RCC 417	0	0	0	0	0	0	0	0	0	+	0	0	0
<i>Crustomastix stigmata</i>	CCMP 3273	0	0	0	0	0	0	0	0	0	0	+	0	0
<i>Dolichomastix tenuilepsis</i>	CCMP3274	0	0	0	0	0	0	0	0	0	0	0	+	0
<i>Monomastix minuta</i>	NIES 255	0	0	0	0	0	0	0	0	0	0	0	0	+

+: amplification, 0: no amplification

These tests also showed that the best reaction temperature was 58 °C for all primer sets and the optimal primer concentration was 200 nM except for primers targeting the larger members of the Mamiellophyceae (*Mantoniella*, *Crustomastix*, *Dolichomastix* and *Monomastix*) which had a lower Ct using 400 nM of primer concentration and this for all DNA concentrations (Table 3). We compared qPCR estimation results to cytometry measurement of calibrated cultures and the linearity of the graph (Figure S1) supports the use of qPCR to estimate cell abundance in natural samples.

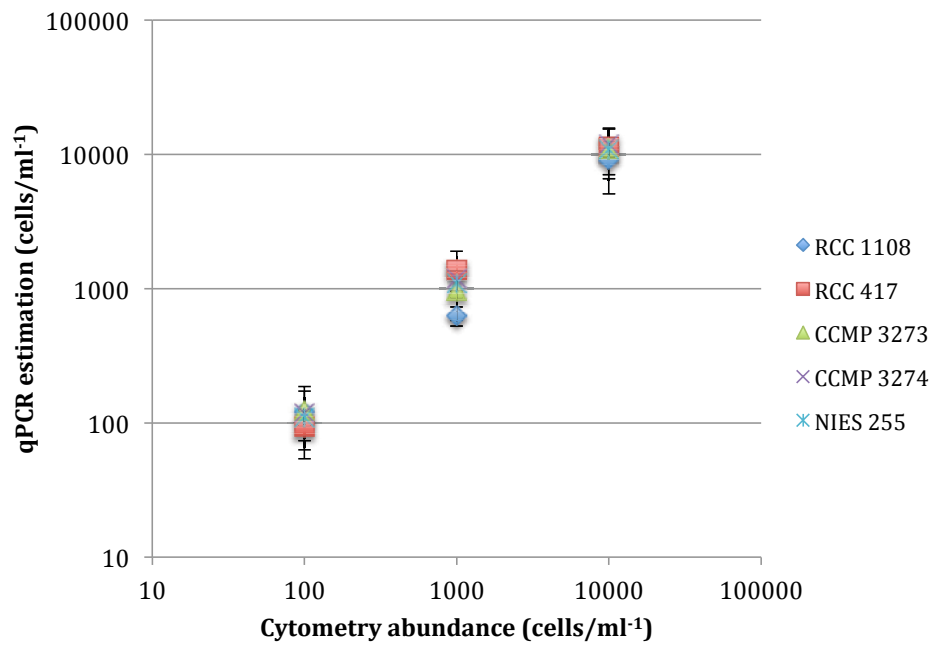


Figure S1: Correlation between cytometry measurements and qPCR estimations

3.3 Analysis of natural samples

The qPCR approach, using the nine primer sets previously designed to estimate Mamiellophyceae abundance, exhibited the same ranges as cytometry (Figure 2). The temporal series of samples from lagoon and off shore Mediterranean Sea taken between April 2013 and April 2014 showed different profiles of phytoplankton communities.

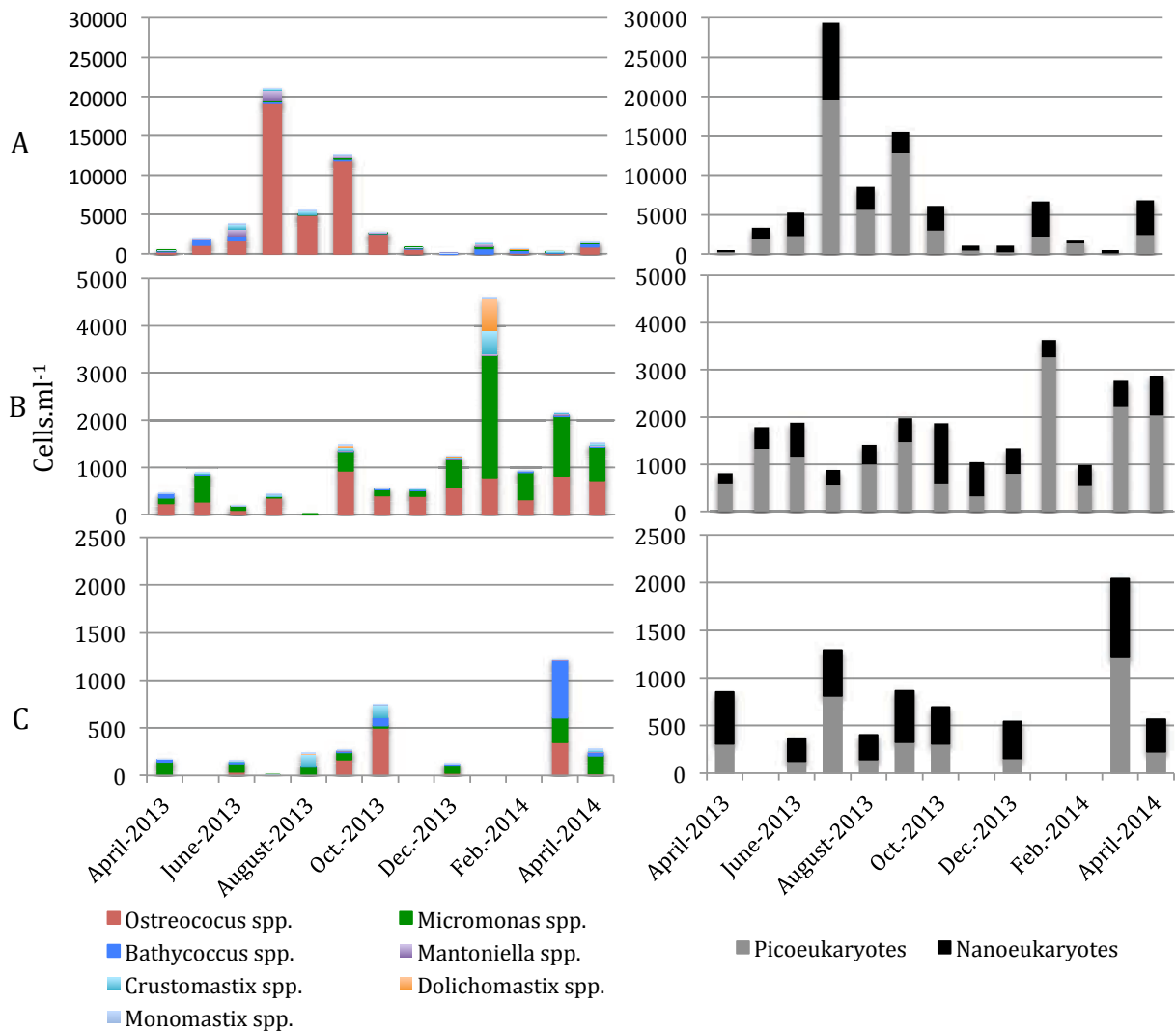


Figure 2: Phytoplankton community abundances. qPCR (left) and cytometry (right) abundance for the Mamiellophyceae and pico and nano phytoplankton populations (cells/ml) at Leucate A (A), Leucate B (B), and MOLA (C)

Cytometry revealed that higher abundance values were found in Leucate A and B (on average 4004 and 1223 picoeukaryotes per mL, respectively) (Figure 2, panels A; B). Lower abundance values were found in MOLA station samples (an average of 403 picoeukaryotes per mL, Figure 2-C). In contrast to the MOLA station where pico- and nano- eukaryote communities were found in the same proportions, picoeukaryote communities dominated at Leucate A and B, by 1.5 and 2.15-fold. In Leucate A and MOLA the higher abundances were observed at the end of spring. In contrast at Leucate B they were observed during winter.

Mamiellophyceae seemed to dominate the picoeukaryote community throughout the year; they represented 97% at Leucate A and around 87% for Leucate B and MOLA. *Micromonas* contributed significantly to this fraction, between 37.3% (Leucate A) and 40% (Leucate B). This contribution was changed only in Leucate A during the summer bloom where *Ostreococcus* dominated. *Micromonas* clade composition changed through the year with clade A in spring and summer, and clade B in autumn and winter (Figure S2). We also noticed a pattern in the dynamics in *Ostreococcus* species compositions: the *Ostreococcus* bloom at Leucate A was composed of *O. tauri*. *Ostreococcus mediterraneus* were found throughout the year and *O. lucimarinus* during autumn and winter (Figure S2). The proportion of *Bathycoccus* within the community of Mamiellophyceae was higher in the open ocean (14% of Picoeukaryote cytometry measurements) than in lagoon (A: 1% and B: 2%). In terms of abundance, *Bathycoccus* was found to be lower in the lagoon A and B stations than in the open-ocean (respectively 16, 42 and 98 cells/ml).

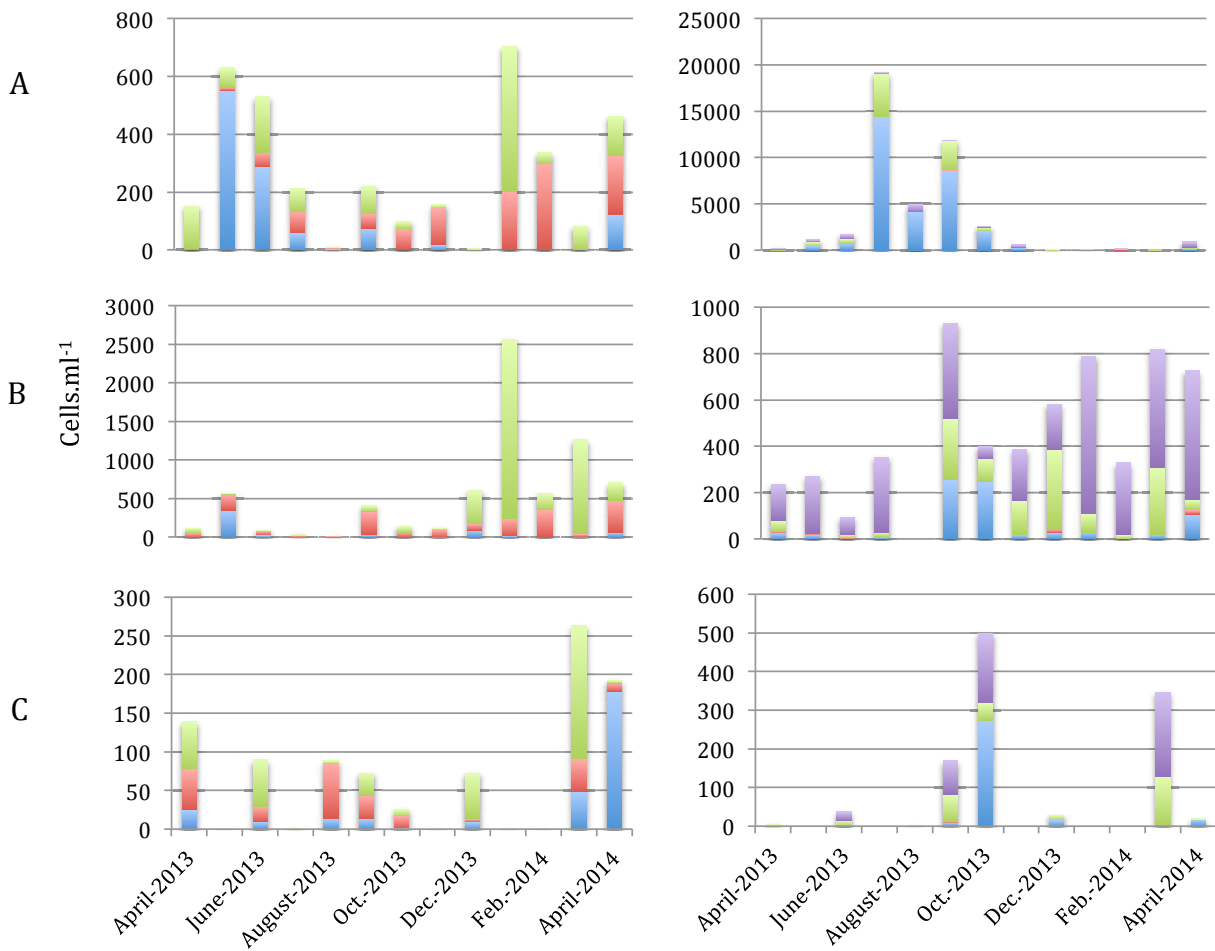


Figure S2: *Micromonas* and *Ostreococcus* clade distribution. Abundance (Cells.ml⁻¹) in Leucate A (A), Leucate B (B), MOLA (C); *Micromonas* (left): Clade A (blue bars), Clade B (red Bars), Clade C (green bars); *Ostreococcus* (right): *O. tauri* (blue bars), *O. mediterraneus* (purple bars), *O. lucimarinus* (green bars), *Ostreococcus* Clade B (red bars)

In contrast, the Mamiellophyceae were not represented in the nanoeukaryote community. The picoeukaryote abundance by cytometry and qPCR abundance were strongly correlated with a regression slope close to 1 (Figure 3).

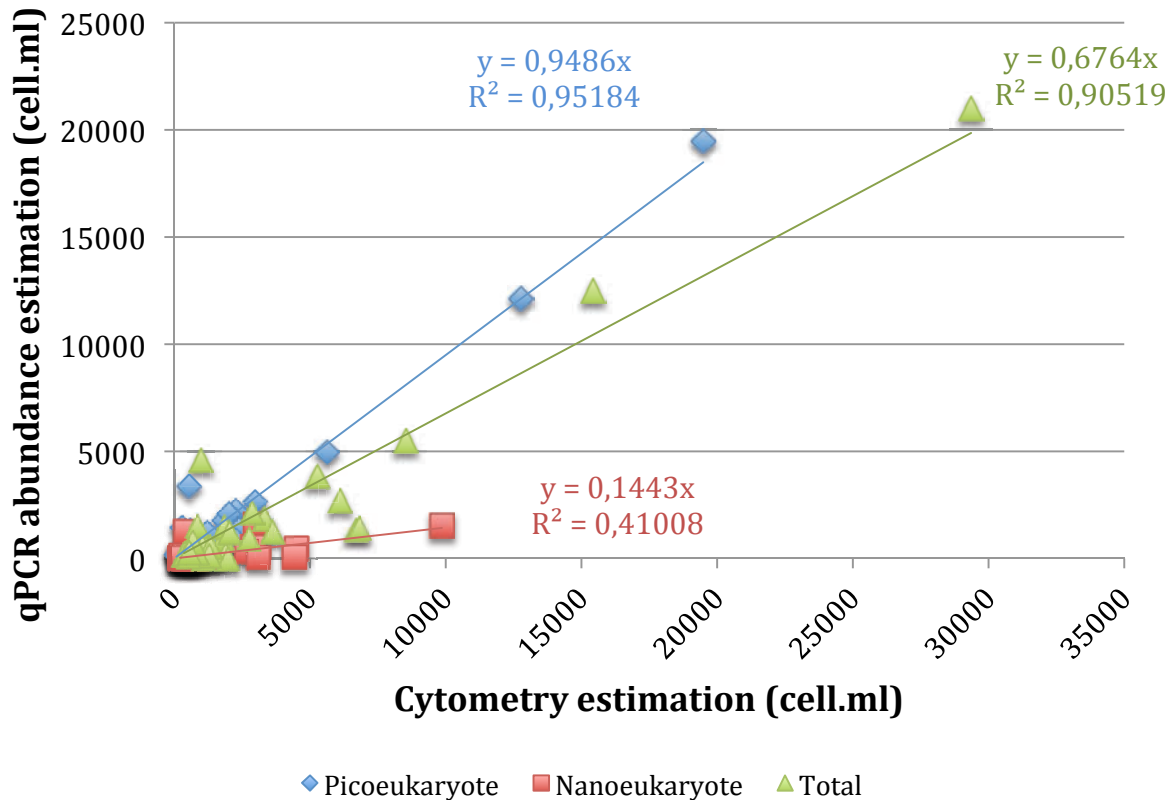


Figure 3. Linear regression between phytoplanktonic communities and qPCR estimations on environmental samples

3.4 Environmental variables

The environmental variables measured did not reveal any statistically significant differences between lagoon stations (p -value > 0.05), except for phosphate concentrations ($p = 0.02$). However, this difference was not significant during the summer (from June to October, data not shown).

DISCUSSION

To understand ecosystem dynamics and evolution in the short- or long-term, it is necessary to estimate the dynamics and contributions of taxonomic groups like the Mamiellophyceae. Previous studies (4,32,33) suggested using different approaches that the Mamiellophyceae dominate picoeukaryotes in nutrient-rich environments.

The majority of qPCR studies (34) were based on gene copy number, using artificial plasmids as internal standards. We opted to use genomic DNA as an internal standard to take into account DNA extraction efficiency (35). Genomic DNA content estimations were lower than values previously reported. According to our measurements and available information already reported (21), the extraction protocol led to a loss of around 30%, even if we centrifuged and resuspended the cell pellet directly in lysis buffer, which is a protocol known to reduce cell loss (36,37). However, we consider that this loss is uniform and the correlation between cytometry and qPCR estimations confirmed that our protocol gave reliable estimates (Figure S1), showing as expected that Mamiellophyceae were dominant in environmental samples. Our quantification of DNA in members of the Mamiellophyceae tends to confirm that DNA content is related to cell size (38), as suggested for phytoplankton (39). Furthermore, there is a correlation between genome size and rDNA copy number in eukaryotes (40). The higher DNA content values found in *Mantoniella*, *Crustomastix*, *Dolichomastix* and *Monomastix* could explain why higher concentrations of primers used for this strain were necessary to lower Ct values (more 18S rDNA copies or more non-specific primer binding).

Nano-eukaryotic genera of Mamiellophyceae (*Mantoniella*, *Crustomastix*, *Dolichomastix* and *Monomastix*) represented only between 4 and 7% of the phytoplanktonic community. This observation is in agreement with the strong correlation between flow cytometry and qPCR estimations (Figure 4). However, there was no correlation between abundances of nanoplanktonic Mamiellophyceae qPCR and nanophytoplankton, which suggests that other groups such as dinoflagellates or prymnesiophytes dominate these communities. Furthermore, the estimated abundances using qPCR were lower than cytometric estimations, generally during spring and summer (52% on average compared to cytometric phototrophic eukaryote abundance). This probably arises from

an increase in the diversity of non-targeted groups such as Heterokontophyta or Haptophyta.

Similar seasonal dynamics along the time series of sampling were observed at the three stations using qPCR and cytometry (Figure 2). In contrast to the offshore station, which showed low variability in terms of pico- and nano-eukaryote abundance, Leucate stations exhibited different patterns, which might be explained by community behaviors and origins. Leucate A is closer to the sea than Leucate B. This proximity implies that seawater entering the lagoon will be at a higher concentration in Leucate A than in Leucate B. Leucate B is on this other side of the lagoon, reducing possibilities of population exchanges with the sea but it is closer to the influx of fresh water. Our observations probably reflect these effects. At Leucate B populations are probably better adapted to the lower salinity provided by fresh water inputs during winter, which would explain the higher abundances found during low temperature and salinity periods. Temperature and salinity increases, together with lower exchanges between adapted populations, might also explain why during summer lower abundances were found at Leucate B. The proximity of Leucate A to the sea might explain why higher abundances were found when temperature and salinity values were closer to those found in the adjacent sea (deleterious to the autochthonous population of Mamiellophyceae) and which may give an advantage to sea communities, particularly in summer when increased evaporation from the lagoon leads to greater influx from the sea. During winter, environmental conditions in the lagoon may have been less advantageous for the adjacent marine community and abundances became similar to those found in the other lagoon station (autochthonous Mamiellophyceae populations).

The class Mamiellophyceae appears to dominate the picoeukaryote community (Figure 4), confirming previous studies based on microscopy (13,15,41) and 18S rDNA analyses (4,14), where Mamiellophyceae and particularly *Micromonas* dominated (accounting for around 40% of the picoeukaryotes). This latter genus is known to be more abundant in nutrient-rich water (4,42). In accordance with this, at Leucate station *Micromonas* is 2 to 5-fold more abundant than the off-shore station. This genus exhibited a clear seasonal variation in clade composition, especially a Leucate A where clade A dominated during spring and clade B during autumn and winter. In contrast clade C seems to be stable throughout the year (Figure S2). Although we can also see these patterns at the other

station, this trend is less clear. Different seasonal variations in terms of clade distribution were observed in the English Channel (43) where clade A et B had a similar dynamics, using FISH to determine abundances.

The bloom of *Ostreococcus* during summer at Leucate A was consistent with the previous observations that this genus seems to form blooms in Summer (44) and can contribute up to 70% of phototrophic picoeukaryotes (45), 89% in the present study. This high abundance of *Ostreococcus* occurred when the salinity was higher than 33 (during the summer period), which is close to lab growth conditions, where high growth rates are measured. The species composition within the *Ostreococcus* genus seems to be dominated by *O. mediterraneus* except during *O. tauri* blooms. Low abundance of *Ostreococcus* clade B was expected, consistent with clade B being a low-light adapted strain (46) and in these shallow lagoons, light intensity is not limited (Figure S2). The observed bloom of *Ostreococcus tauri* also supports the hypothesis that phytoplanktonic communities from the sea entered the lagoon where they found conditions that promoted their development, since we found this community of *O. tauri* at Leucate B, farther from the sea than Leucate A.

Little information about the distribution of *Bathycoccus* is available. Johnson and Sieburth (41) described a phytoplankter with a similar morphology before *Bathycoccus prasinos* was described as a species (11). Using electronic microscopy (MET) on North-Atlantic samples, they found a small scaly picoeukaryote in many samples. Their description now suggests that it was *Bathycoccus*, and they estimated the cell concentration to be 10³-10⁴ cell.mL⁻¹. Zhu *et al.* (21) found that *Bathycoccus* represented 3.4 to 4.2% of total eukaryotes along the Mediterranean Sea coast and Wu *et al.* (34) observed that *Bathycoccus* contributed up to 6.5% (max. 1628 cell.mL⁻¹) in the South China Sea. In the latter study, *Bathycoccus* exhibited higher population densities (10.7% of picoeukaryotes) in oligotrophic water, which is consistent with our observations (14%).

CONCLUSION

Using the qPCR technique, we observed that the Mamiellophyceae, particularly the picoeukaryotes. *Micromonas*, *Bathycoccus* and *Ostreococcus* were clearly dominant in the eutrophic lagoon. However, the Mamiellophyceae account for only a small proportion of the nanoeukaryotic community. Connections between environments probably provide opportunities for strains to occupy free ecological niches. When environmental factors such as temperature and salinity become deleterious for autochthonous strains, connection by currents may bring in adapted strains that can become abundant. Compared to *Micromonas*, which represents a high proportion of picoeukaryotes throughout the year, *Bathycoccus* seems to be poorly adapted to eutrophic environments, but in oligotrophic environments it seems more competitive than other Mamiellophyceae. *Ostreococcus tauri* in the lagoon environment tends to be an opportunistic species, and exhibits high blooms when the environment becomes advantageous for them.

This study provides an overview of the abundances of genera over the course of a year in these particular environments, to help understand processes that influence phytoplanktonic populations. Lagoon ecosystems are highly dynamic and may be extremely variable between years depending on climate, but in this Mamiellophyceae-rich environment, qPCR provides a useful tool for addressing the impact environmental changes, such as water exchanges between lagoon and sea, on picoeukaryotic community dynamics.

This study provides an overview of the abundances of genera over the course of a year in these particular environments, to help understand processes that influence phytoplanktonic populations. Lagoon ecosystems are highly dynamic and may be extremely variable between years depending on climate, but in this Mamiellophyceae-rich environment, qPCR provides a useful tool for addressing the impact of environmental changes, such as water exchanges between lagoon and sea, on picoeukaryotic community dynamics.

Acknowledgements

This work was supported by a “Agence Nationale de la Recherche” grant “DECOVIR” N° ANR-12-BSV7-0009. Hugo Lebretonchel benefited from a doctoral fellowship from the Université Pierre et Marie Curie.

References

1. Grob, C., Ulloa, O., Claustre, H., Huot, Y., Alarcón, G., Marie, D. Contribution of picoplankton to the total particulate organic carbon concentration in the eastern South Pacific. *Biogeosciences* **4**, 837–852 (2007).
2. Cuvelier, M. L., Allen, A.E., Monier, A., McCrow, J.P., Messié, M., Tringe, S.G., Woyke, T., Welsh, R.M., Ishoey, T., Lee, J.-H., Binder, B.J., DuPont, C.L., Latasa, M., Guigand, C., Buck, K.R., Hilton, J., Thiagarajan, M., Caler, E., Read, B., Lasken, R.S., Chavez, F.P., Worden, A.Z. Targeted metagenomics and ecology of globally important uncultured eukaryotic phytoplankton. *Proc. Natl. Acad. Sci.* **107**, 14679–14684 (2010).
3. Ishizaka, J., Harada, K., Ishikawa, K., Kiyosawa, H., Furusawa, H., Watanabe, Y., Ishida, H., Suzuki, K., Handa, N., Takahashi, M. Size and taxonomic plankton community structure and carbon flow at the equator, 175°E during 1990–1994. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **44**, 1927–1949 (1997).
4. Not, F., Latasa, M., Marie, D., Cariou, T., Vaultot, D., Simon, N. A single species, *Micromonas pusilla* (Prasinophyceae), dominates the eukaryotic picoplankton in the western English Channel. *Appl. Environ. Microbiol.* **70**, 4064–4072 (2004).
5. Worden, A. Z. & Not, F. in *Microbial Ecology of the Oceans* (ed. Kirchman, D. L.) 159–205 (John Wiley & Sons, Inc., 2008).
6. Díez, B., Pedrós-Alió, C. & Massana, R. Study of Genetic Diversity of Eukaryotic Picoplankton in Different Oceanic Regions by Small-Subunit rRNA Gene Cloning and Sequencing. *Appl. Environ. Microbiol.* **67**, 2932–2941 (2001).
7. Moon-van der Staay, S. Y., De Wachter, R. & Vaultot, D. Oceanic 18S rDNA sequences from picoplankton reveal unsuspected eukaryotic diversity. *Nature* **409**, 607–610 (2001).

8. Vargas, C. de, Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., Lara, E., Berney, C., Bescot, N.L., Probert, I., Carmichael, M., Poulain, J., Romac, S., Colin, S., Aury, J.-M., Bittner, L., Chaffron, S., Dunthorn, M., Engelen, S., Flegontova, O., Guidi, L., Horák, A., Jaillon, O., Lima-Mendez, G., Lukeš, J., Malviya, S., Morard, R., Mulo, M., Scalco, E., Siano, R., Vincent, F., Zingone, A., Dimier, C., Picheral, M., Searson, S., Kandels-Lewis, S., Coordinators, T.O., Acinas, S.G., Bork, P., Bowler, C., Gorsky, G., Grimsley, N., Hingamp, P., Iudicone, D., Not, F., Ogata, H., Pesant, S., Raes, J., Sieracki, M.E., Speich, S., Stemmann, L., Sunagawa, S., Weissenbach, J., Wincker, P., Karsenti, E., Boss, E., Follows, M., Karp-Boss, L., Krzic, U., Reynaud, E.G., Sardet, C., Sullivan, M.B., Velayoudon, D. Eukaryotic plankton diversity in the sunlit ocean. *Science* **348**, 1261605 (2015).
9. Marin, B. & Melkonian, M. 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 (2010).
10. Bucher, R. W. Contribution to our knowledge of the smaller marine algae. (1952).
11. Eikrem, W. & Throndsen, J. Phycologia Online - The ultrastructure of *Bathycoccus* gen. nov. and *B. prasinos* sp. nov., a non-motile picoplanktonic alga (Chlorophyta, Prasinophyceae) from the Mediterranean and Atlantic. (1990).
12. Chrétiennot-Dinet, M.-J. , Courties, C., Vaquer, A., Neveux, J., Claustre, H., Lautier, J., Machado, M.C. A new marine picoeucaryote: *Ostreococcus tauri* gen. et sp. nov. (Chlorophyta, Prasinophyceae). *Phycologia* **34**, 285–292 (1995).
13. Throndsen, J. & Kristiansen, S. *Micromonas pusilla* (Prasinophyceae) as part of pico-and nanoplankton communities of the Barents Sea - THRONSDEN - 2007 - Polar Research - Wiley Online Library. (1991).
14. Guillou, L., Eikrem, W., Chrétiennot-Dinet, M.-J., Le Gall, F., Massana, R., Romari, K., Pedrós-Alió, C., Vaulot, D. Diversity of Picoplanktonic Prasinophytes Assessed by Direct Nuclear SSU rDNA Sequencing of Environmental Samples and Novel Isolates Retrieved from Oceanic and Coastal Marine Ecosystems. *Protist* **155**, 193–214 (2004).
15. Zingone, A., Sarno, D. & Forlani, G. Seasonal dynamics in the abundance of *Micromonas pusilla* (Prasinophyceae) and its viruses in the Gulf of Naples (Mediterranean Sea). *J. Plankton Res.* **21**, 2143–2159 (1999).
16. Subirana, L., Péquin, B., Michely, S., Escande, M.-L., Meilland, J., Derelle, E., Marin, B., Piganeau, G., Desdevises, Y., Moreau, H., Grimsley, N.H. Morphology, genome plasticity, and phylogeny in the genus *ostreococcus* reveal a cryptic species, *O. mediterraneus* sp. nov. (Mamiellales, Mamiellophyceae). *Protist* **164**, 643–659 (2013).
17. Šlapeta, J., López-García, P. & Moreira, D. Global Dispersal and Ancient Cryptic Species in the Smallest Marine Eukaryotes. *Mol. Biol. Evol.* **23**, 23–29 (2006).
18. Simon, N., Cras, A.-L., Foulon, E. & Lemée, R. Diversity and evolution of marine phytoplankton. *C. R. Biol.* **332**, 159–170 (2009).

19. Dasilva, C. R., Li, W. K. W. & Lovejoy, C. Phylogenetic diversity of eukaryotic marine microbial plankton on the Scotian Shelf Northwestern Atlantic Ocean. *J. Plankton Res.* fbt123 (2013). doi:10.1093/plankt/fbt123
20. Speel, E. J. M., Hopman, A. H. N. & Komminoth, P. Amplification Methods to Increase the Sensitivity of In Situ Hybridization: Play CARD(S). *J. Histochem. Cytochem.* **47**, 281–288 (1999).
21. Zhu, F., Massana, R., Not, F., Marie, D. & Vaulot, D. Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. *FEMS Microbiol. Ecol.* **52**, 79–92 (2005).
22. Pernthler, J., Pernthaler, annelie & Rudolf, amann. Automated Enumeration of Groups of Marine Picoplankton after Fluorescence In Situ Hybridization. (2003).
23. Weithoff, G. The concepts of ‘plant functional types’ and ‘functional diversity’ in lake phytoplankton – a new understanding of phytoplankton ecology? *Freshw. Biol.* **48**, 1669–1675 (2003).
24. Zimmermann, J., Jahn, R. & Gemeinholzer, B. Barcoding diatoms: evaluation of the V4 subregion on the 18S rRNA gene, including new primers and protocols. *Org. Divers. Evol.* **11**, 173–192 (2011).
25. Gonzalez, P. & Labarère, J. Sequence and Secondary Structure of the Mitochondrial Small-Subunit rRNA V4, V6, and V9 Domains Reveal Highly Species-Specific Variations within the Genus *Agrocybe*. (1998).
26. Bec, B., Collos, Y., Souchu, P., Vaquer, A., Lautier, J., Fiandrino, A., Benau, L., Orsoni, V., Laugier, T. Distribution of picophytoplankton and nanophytoplankton along an anthropogenic eutrophication gradient in French Mediterranean coastal lagoons. *Aquat. Microb. Ecol.* **63**, 29–45 (2011).
27. Courties, C. , Vaquer, A., Troussellier, M., Lautier, J., Chrétiennot-Dinet, M.J., Neveux, J., Machado, C., Claustre, H. Smallest eukaryotic organism. *Nature* **370**, 255–255 (1994).
28. Worden, A. Z. Picoeukaryote diversity in coastal waters of the Pacific Ocean. *Aquat. Microb. Ecol.* **43**, 165–175 (2006).
29. Doyle, J. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull* **19**, 11–15 (1987).
30. Suzuki, M. T., Taylor, L. T. & DeLong, E. F. Quantitative Analysis of Small-Subunit rRNA Genes in Mixed Microbial Populations via 5′ -Nuclease Assays. *Appl. Environ. Microbiol.* **66**, 4605–4614 (2000).
31. Campbell, M. S. & Wright, A. C. Real-Time PCR Analysis of *Vibrio vulnificus* from Oysters. *Appl. Environ. Microbiol.* **69**, 7137–7144 (2003).
32. Throndsen, J. Occurrence and productivity of small marine flagellates. *Nor. J. Bot.* (1976).

33. Balzano, S., Marie, D., Gourvil, P. & Vaultot, D. Composition of the summer photosynthetic pico and nanoplankton communities in the Beaufort Sea assessed by T-RFLP and sequences of the 18S rRNA gene from flow cytometry sorted samples. *ISME J.* (2012).
34. Wu, W., Huang, B. & Zhong, C. Photosynthetic picoeukaryote assemblages in the South China Sea from the Pearl River estuary to the SEATS station. *Aquat. Microb. Ecol.* **71**, 271–284 (2013).
35. Brittain, S. M., Wang, J., Babcock-Jackson, L., Carmichael, W.W., Rinehart, K.L., Culver, D.A. Isolation and Characterization of Microcystins, Cyclic Heptapeptide Hepatotoxins from a Lake Erie Strain of *Microcystis aeruginosa*. *J. Gt. Lakes Res.* **26**, 241–249 (2000).
36. Galluzzi, L. , Penna, A., Bertozzini, E., Vila, M., Garcés, E., Magnani, M. Development of a Real-Time PCR Assay for Rapid Detection and Quantification of *Alexandrium minutum* (a Dinoflagellate). *Appl. Environ. Microbiol.* **70**, 1199–1206 (2004).
37. Biegala, I. C., Not, F., Vaultot, D. & Simon, N. Quantitative Assessment of Picoeukaryotes in the Natural Environment by Using Taxon-Specific Oligonucleotide Probes in Association with Tyramide Signal Amplification-Fluorescence In Situ Hybridization and Flow Cytometry. *Appl. Environ. Microbiol.* **69**, 5519–5529 (2003).
38. Cavalier-Smith, T. Selfish DNA and the origin of introns. *Nature* **315**, 283–284 (1985).
39. Boucher, N., Vaultot, D. & Partensky, F. Flow cytometric determination of phytoplankton DNA in cultures and oceanic populations. *Mar. Ecol. Prog. Ser.* **71**, 75–84 (1991).
40. Prokopowich, C. D., Gregory, T. R. & Crease, T. J. The correlation between rDNA copy number and genome size in eukaryotes - Genome. *Genome* **46**, 48–50 (2003).
41. Johnson, P. W. & Sieburth, J. M. In-Situ Morphology and Occurrence of Eucaryotic Phototrophs of Bacterial Size in the Picoplankton of Estuarine and Oceanic Waters1. *J. Phycol.* **18**, 318–327 (1982).
42. Viprey, M., Guillou, L., Ferréol, M. & Vaultot, D. Wide genetic diversity of picoplanktonic green algae (Chloroplastida) in the Mediterranean Sea uncovered by a phylum-biased PCR approach. *Environ. Microbiol.* **10**, 1804–1822 (2008).
43. Baudoux, A.-C. , Lebretonchel, H., Dehmer, H., Latimier, M., Edern, R., Rigaut-Jalabert, F., Ge, P., Guillou, L., Foulon, E., Bozec, Y., Simon, N. Interplay between the genetic clades of *Micromonas* and their viruses in the Western English Channel. *Environ. Microbiol. Rep.* **7**, 765–773 (2015).
44. O’Kelly, C. J., Sieracki, M. E., Thier, E. C. & Hobson, I. C. A Transient Bloom of *Ostreococcus* (chlorophyta, Prasinophyceae) in West Neck Bay, Long Island, New York. *J. Phycol.* **39**, 850–854 (2003).

45. Countway, P. D. & Caron, D. A. Abundance and Distribution of *Ostreococcus* sp. in the San Pedro Channel, California, as Revealed by Quantitative PCR. *Appl. Environ. Microbiol.* **72**, 2496–2506 (2006).
46. Rodríguez, F. , Derelle, E., Guillou, L., Le Gall, F., Vaultot, D., Moreau, H. Ecotype diversity in the marine picoeukaryote *Ostreococcus* (Chlorophyta, Prasinophyceae). *Environ. Microbiol.* **7**, 853–859 (2005).

Chapitre V – Impact of environmental factors on Prasinovirus communities

Résumé



La comparaison spatiale de 4 sites à fort contraste environnemental échantillonnés entre mars 2013 et avril 2014 a montré un lien entre les communautés de prasinovirus et les conditions environnementales



Bien que les virus possèdent une distribution étendue, leur abondance est limitée spatialement, en lien avec les communautés phytoplanctoniques



Les conditions environnementales semblent jouer un rôle direct dans la sélection des communautés virales

Participation du doctorant

- ✓ Recherche bibliographique
- ✓ Mise en place de la stratégie d'échantillonnage
- ✓ Echantillonnage et préparation des échantillons
- ✓ Extraction d'ADN et préparation des banques pour le séquençage Illumina
- ✓ Analyses génétiques et bioinformatiques
- ✓ Analyses statistiques
- ✓ Rédaction

Ce chapitre est présenté sous la forme d'un article en préparation :

Lebretonchel H, Alric B, Derelle E, Grimsley N, Dray S, Desdevises Y. Impact of environmental factors on Prasinovirus communities.

INTRODUCTION

Viruses are the most abundant biological entities in oceanic ecosystems (Wommack and Colwell, 2000; Suttle, 2005). Beyond an impressive abundance, viruses play an ecological key role, by influencing the composition and function of microbial communities through the conversion of biomass to dissolved and particulate organic matter via host lysis (Suttle, 2005). Likewise, they regulate the evolutionary processes of their host populations through horizontal gene transfer, which can enhance host diversification (Nagasaki *et al.*, 1999; Tomaru *et al.*, 2004). The extent to which viruses drive the microbial processes can be influenced by both abiotic and biotic factors of the environment in which they occur.

Most of our knowledge in virus ecology concerns the dynamics of their abundance, which tend to be related to trophic conditions of marine aquatic systems, with higher values in eutrophic than oligotrophic systems (Boehmen *et al.* 1993; Weinbauer *et al.*, 1993; Noble and Furhman, 2000; Wommack et Colwell, 2000). While the impact of environmental changes on viruses has been considered mainly through cultured-based approaches (Danovaro, 2010; Clasen *et al.*, 2008; Harvey et Ryan, 2004; Gerba, 1984, Jin *et al.*, 2005), few studies focused on the natural communities. Environmental conditions such as temperature (Rowe *et al.*, 2012), oxygen (Taylor *et al.*, 2003), pH (Weinbauer, 2004) or salinity (Clasen and Suttle, 2009) have however been reported to influence natural viral communities. So far, and in contrast to viruses that infect prokaryotes (e.g. Weinbauer, 2004; Sime-Ngando, 2014), there are little data on viruses that infect eukaryotic phytoplankton (Short, 2012). However, among marine phytoplankton, photosynthetic picoeukaryotes are the major players in CO₂ fixation and primary production (Worden *et al.*, 2004; Jardillier *et al.*, 2010). Consequently, with the ongoing global climatic change, it is important to unravel how environmental factors affect algal virus dynamics.

This study focuses therefore on Prasinovirus, which are eukaryote marine viruses from the family Phycodnaviridae. They infect photosynthetic eukaryotes within the class Mamiellophyceae (Marin and Melkonian, 2010), dominated by three main

picoeukaryotic genera (*Micromonas*, *Ostreococcus* and *Bathycoccus*). Several viruses infecting these microalgae are described, including *Ostreococcus* viruses [OtVs (Derelle *et al.*, 2008; Weynberg *et al.*, 2009, 2011), OIVs and OmV (Moreau *et al.*, 2010; Derelle *et al.*, 2015)] infecting respectively *Ostreococcus tauri* (Courties *et al.* 1994), *Ostreococcus lucimarinus* (Palenik *et al.*, 2007) and *Ostreococcus mediterraneus* (Subirana *et al.*, 2013); *Micromonas* viruses (MpVs) infecting *Micromonas spp.* (Moreau *et al.*, 2010; Baudoux *et al.*, 2015; Mayer and Taylor, 1979), and *Bathycoccus* viruses (BpVs; Moreau *et al.*, 2010) infecting *Bathycoccus spp.* Prasinoviruses are present in all the oceans (Weynberg *et al.*, 2009; Bellec *et al.*, 2010a; Manrique *et al.*, 2012, Hingamp *et al.*, 2013) but are commonly found on coastal and lagoon of NW Mediterranean Sea (Bellec *et al.*, 2010b; Clerissi *et al.*, 2012).

In order to assess the impact of environmental conditions on Prasinovirus communities, we compared the virus communities in contrasted environments. Four sites in the Mediterranean Sea were sampled and biotic and abiotic variables were measured.

In this study, we describe both spatial and temporal variations of Prasinovirus communities throughout the year, and relate this to their dispersion and to the dynamics of their host communities. We propose a hypothesis to explain how environmental factors influence Prasinovirus distributions and communities.

MATERIALS AND METHODS

2.1 Sampling sites and environmental variables

Water samples were collected monthly between March 2013 and April 2014 at four sites along a trophic gradient located in the northwest Mediterranean Sea (Figure 1). Sampling, extraction, amplification, and sequencing protocols and environmental parameter analysis were performed on discrete water samples. Prasinovirus diversity

was examined for each site according to a broad range of abiotic and biotic variables. These included depth (m), temperature (°C), salinity, ammonium (μM), nitrites (μM), nitrates (μM), phosphate (μM), silicates (μM), oxygen (mL/L), chlorophyll a ($\mu\text{g/L}$), pheophytine a ($\mu\text{g/L}$), particulate organic carbon (POC; $\mu\text{g/L}$), particulate organic nitrogen (PON; $\mu\text{g/L}$), suspended matter (MES; mg/L) and pH. The concentration (cell.ml⁻¹) of photosynthetic picoeukaryotes (PPE), nanoeukaryotes (PNE), cyanobacteria, and cryptophyceae were estimated using flow cytometry. The meteorological variables (wind speed (m/s), photosynthetic active radiation (PAR, W/m²) and atmospheric pressure (Pa)), were extracted from the ERA-interim model (European Center for Medium-Range Weather Forecasts, <http://www.ecmwf.int/>).

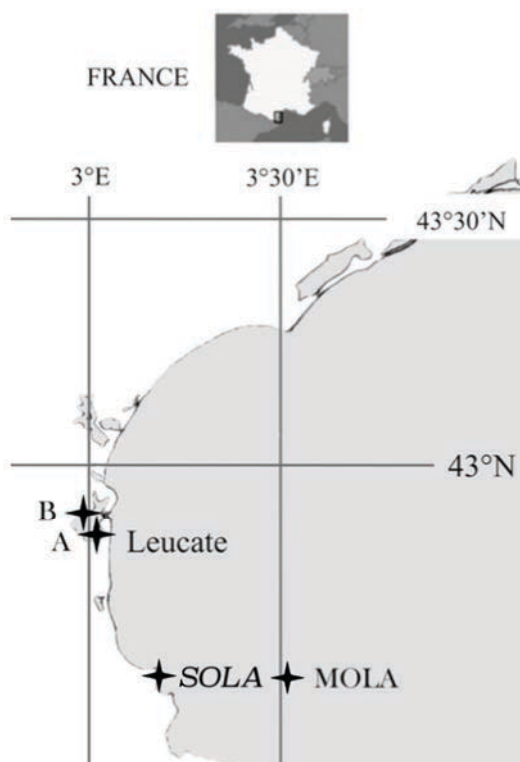


Figure 1: Sampling sites.

GPS coordinates as longitude/latitude (WGS84); Leucate (A: 3.021/42.805; B: 2,998/42.873); SOLA: 3.145/42.489; MOLA: 3.543/42.453.

2.2 Sampling of picoeukaryotes and prasinoviruses, and DNA extraction

Forty litres of seawater collected at the surface were first passed through 200- μm mesh size filters to remove larger particles directly during sampling process and split into 3 parts: 20 l and 5 l were used respectively to recovery virus and host, and 15 l for the physical factors.

At the lab, the 20L samples were filtered successively through 5, 3, 1, 0.45 and 0.22- μm mesh size filters (5- μm mesh filters; 3 μm PC Membrane, Millipore, other filter PES filters) using a peristaltic pump. After the 0.22 μm filtration, viruses were flocculated with iron chloride, filtered through 1 μm mesh size (or pore size) filters (PC Membrane) as described in John *et al.* (2010) and stored at 4°C, for up to 4 months. Viruses were resuspended by adding 20 mL of resuspension buffer (John *et al.*, 2011) for 24 hours at 4°C under agitation and concentrated in 2 mL by ultracentrifugation, during 2 hours at 100000 g. The pellet was then resuspended using SM buffer and frozen at -80°C before DNA extractions.

For the host recovery, 5 litres of water were filtered through 10 μm mesh filters, concentrated in 50 ml using VivaFlow 200 (50 kDa) and stored at -80°C with glycine betaine (8% final concentration). Based on SSC and FL3 fluorescence, cell sorting was performed on concentrated samples to extract in triplicate the phytoeukaryotic cells stored at -80°C before DNA extractions.

DNA extraction was performed following a modified CTAB protocol (Winnepenninckx *et al.*, 1994), (i) sorted eukaryotes and virus-containing buffer were incubated at 60°C for 1 hour with 10 mL of CTAB buffer (2% CTAB; 100 mM TrisHCl pH=8; 20mM EDTA; 1.4 M NaCl; 0.2 % β -mecaptoethanol; 0.1 mg proteinase K; 10 mM DTT), (ii) DNA was purified using an equal volume of chloroform:isomyalcohol (24:1) followed by 1 hour of RNase digestion at 37°C, (iii) DNA was precipitated with 2/3 volume of isopropanol and washed with 76% EtOH - 10 mM AcNH₄, (iv) DNA was resuspended using 50 μL of ultra-pure free DNA water and stored at -20°C until amplification and sequencing.

2.3 Amplification, sequencing and data analysis

The viral DNA polymerase B (PolB) and Major Capsid protein genes (MCP), and the eukaryotic 18S RNA (V4 and V9 regions) were partially amplified using the following primer sets: VpolAS4-VpolAAS1, VmcpAS3-VmcpAAS1, TAREuk454FWD1-TAREukREV3, 1389f-1510r; (Table 1). Adapters were added in order to multiplex the different amplicons for illumina sequencing (MiSeq).

Table 1: primer characteristics

Name	Sequence (5'-3')	Size (bp)	Sense	Gene/targeted group
VpolAS4	GARGGIGCIACIGTIYTNGA	20	Forward	PolB/Prasinovirus
VpolAAS1	CCIGTRAAICCRTAIACISWRITTCAT	26	Reverse	
VmcpAS3	GGIGGICARMGIRTIGAYAA	20	Forward	MCP6/Prasinovirus
VmcpAAS1	TGIACYTGYTCDATIARRTAYTCRTG	26	Reverse	
TAREuk454FWD1	CAGCASCYGC GGTAATTCC	20	Forward	18SV4/Eukaryotes
TAREukREV3	CTTTCGTTCTTGATYRA	18	Reverse	
1389f	TGTACACACCGCCC	15	Forward	18SV9/Eukaryotes
1510r	CTTCYGCAGGTTACCTAC	20	Reverse	

Polymerase Chain Reactions (PCR) amplification of PolB and MCP fragments were done in triplicate and set up as follows: 2 µL of environmental DNA (100ng) and 3 µL of DNA-free water were added to a 45 µL mix-PCR containing 0.25 mM of each desoxyribonucleoside triphosphate, 200 nM of each primer, 1/10 of Advantage 2 PCR reaction buffer and 1 U of Advantage 2 Polymerase. For viral DNA, the thermal cycling conditions consisted of 3 min initial denaturation of the DNA template at 95°C, followed by 35 cycles of 30 s at 95°C, 60 s at 50°C for PolB and 48°C for MCP6, 90 s at 72°C and a

final extension step of 4 min at 72°C. Amplification of V4 region of eukaryotic 18S RNA (18SV4) was performed with an initial denaturation of DNA template of 5 min at 95°C, followed by 10 cycles at 94°C (30 s), 57°C (45 s), 72°C (1 min), by 25 author cycles at 94°C (30 s), 48°C (45s), 72°C (1 min) and a final extension at 72°C (10 min). For the V9 region of eukaryotic 18S RNA (18SV9), amplification was performed with 5 min initial denaturation of the DNA template at 95°C, followed by 35 cycles at 94°C (30 s), 57°C (45s), 72°C (1 min) and a final extension at 72°C (2 min). PCR products were controlled by a 0.8% agarose gel electrophoresis containing 0.05% of Gel Red in 0.5% TAE buffer, 100V, 25 min. Bands were visualized under UV light. PCR products were sequenced using the MiSeq technology (2 x 250bp for PolB, MCP and 18sV4; 1x150bp for 18SV9).

Nest Generation Sequencing (NGS) data analyses were done using Mothur software (Schloss *et al.*, 2009) for 18S sequences and with USEARCH 7 program for viral sequences. Basically, all NGS data analyses were conducted as follows: (i) extraction of unique sequences and alignment using Muscle software (Edgar, 2004) for viral sequences and directly with Mothur for 18S (Schloss *et al.*, 2009), (ii) deletion of chimera sequences using UCHIME program implemented in both softwares, (iii) design of Operational Taxonomic Units (OTUs) based on 99% similarity cutoff, (iiii) assignment of centroid sequence for each OTU to a reference database (viral database formed with viral sequences available on NCBI and cleaned after alignment, and Protist Ribosomal Reference database (pr2; <http://ssu-rrna.org/pr2>). We used a similarity matrix for viruses to assign sequences to the closest taxonomic level.

2.4 Statistical and multivariate analysis

All statistical analyses were performed using R 3.2 software (R Core Team 2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (<https://www.R-project.org/>). Sampling stations were discriminated according to their environmental characteristics by a Principal Component Analysis (PCA) of environmental parameters after a standardization using vegan package (Oksanen, 2013). Relative abundance values of OTUs (host and virus) were obtained by dividing number of reads by number of all reads at each sampling

point. Samples were clustered according to their composition and content in Prasinovirus and eukaryotes (normalized reads numbers) using the unweighted pair group method with arithmetic mean (UPGMA) based on Bray-Curtis dissimilarity using the `hclust` function. Distance matrix using Bray-Curtis method is seen to be one of the most reliable performer to observe group forming based on abundance (Clarke, 1993). Venn diagrams showing shared Prasinovirus OTUs percentage between sampling sites were drawn using `venn` function in the `gplot` package, only with OTUs found twice at least in one station. A correspondence analysis (CA) was done to describe the distribution of Prasinovirus and Mamiellophyceae in the four sites during sampling year using the `cca` function in the `vegan` package. A canonical correspondence analysis (CCA) was used to investigate how relative abundance dynamics of the different Prasinovirus were constrained by the environmental variables. We selected the best variables that explain the variations of Prasinovirus abundance using a CA and environment vector fitting, using a forward selection implemented in the `ordisep` function from `vegan` package. We tested the significance of the model by ANOVA using 999 permutations with the `anova` function in the `vegan` package.

In order to investigate the common structure and stability of this structure of the relationships between relative abundance of virus (or host) OTUs and the environmental parameters, the STATICO method (Simier *et al.*, 1999; Thioulouse *et al.*, 2004) was carried out on two series of tables (i.e. one containing the values of p OTUs-columns measured at n sampling date-rows and another containing the values of q environmental variable-columns measured at n sampling date-rows). From a statistical point of view, STATICO is a multi-table analysis performed on the series of cross-tables resulting from the co-inertia analysis of each pair of tables. This method proceeds in three stages: (1) each table is first analyzed by a one-table method (here PCA); (2) the co-structure of each pair of table is estimated by a co-inertia analysis producing cross-table for each site; (3) since a cross-table is produced for each site, a partial triadic analysis (PTA) is applied on these leading to three main results. The first stage of the PTA correspond to interstructure analysis consisting to compare the four site structures and identify the proximity between each pair of table. The second step, the compromise analysis, gives an ordination of the environment variables and the OTU relative abundances on shared axes, and represents the average OUT-environment relationships across the sites showing the stable part of the relationships. The four subsets are then

projected separately onto the compromise in order to compare the trajectories between sites. Calculations and graphs were done using ade4 package. Regarding eukaryote and virus OTUs variables, Leucate A could not be represented in the same factorial space as the other stations because in terms of relative abundance this station displays a much higher variability. We standardized the data with a log function to clarify the temporal variation of other stations (Leucate B, SOLA, MOLA).

In order to investigate the spatial and temporal variability in the Prasinovirus and host communities, relative abundance of virus (or host) OTUs and environmental parameters for each site were combined to generate two series of tables (i.e. one containing the values of p OTUs-columns measured at n sampling date-rows and another containing the values of q environmental variable-columns measured at n sampling date-rows). The STATICO method (Simier *et al.* 1999; Thioulouse *et al.*, 2004) was carried out to analyse the two series of tables, and aims to find the common structure and stability of this structure of the relationships between two sets of variables. From a statistical point of view, STATICO is a symmetric coupling method that joins the STATIS method to operators of co-inertia analysis. In other words, it is a multi-table analysis performed on the series of cross-tables resulting from the co-inertia analysis of each pair of tables. This method proceeds in three stages: (1) each table is first analyzed by a one-table method (here PCA); (2) the co-structure of each pair of table is estimated by a co-inertia analysis producing cross-table for each site; (3) since a cross-table is produced for each site, a partial triadic analysis is used to determine the stable part of the structure of the series of cross-table. The STATICO method also allows to plot the projection of coordinates of each point for each station, in the same statistical compromised space, which allowed the comparison between stations.

RESULTS

- 3.1 Contrasted environments

-

The two first axes of PCA accounted for 63.3 % of the total canonical variation showing a high variability between the four sites (Figure 2). Stations were distributed along a trophic gradient: an oligotrophic offshore station, a mesotrophic coastal station and a eutrophic lagoon. The eutrophic state is principally explained by nitrate and oxygen concentrations. The distribution of the data from stations was also related to seasons, with winter samples characterised by higher concentration of particulate organic carbon (POC) and particulate organic nitrogen (PON), pheophytine and suspended matter.

The PCA points to higher seasonal variations in lagoons than in marine stations. Supplemental descriptive variables, which represent relative abundance of *Ostreococcus tauri* and *O. lucimarinus* viruses (OtV and OIV), *Bathycoccus virus* (BpV), and *Micromonas virus* (MpV), were not well represented in this analysis. The abundance of OtV was the best represented among virus variables and showed a positive correlation to both seasonal and trophic gradients.

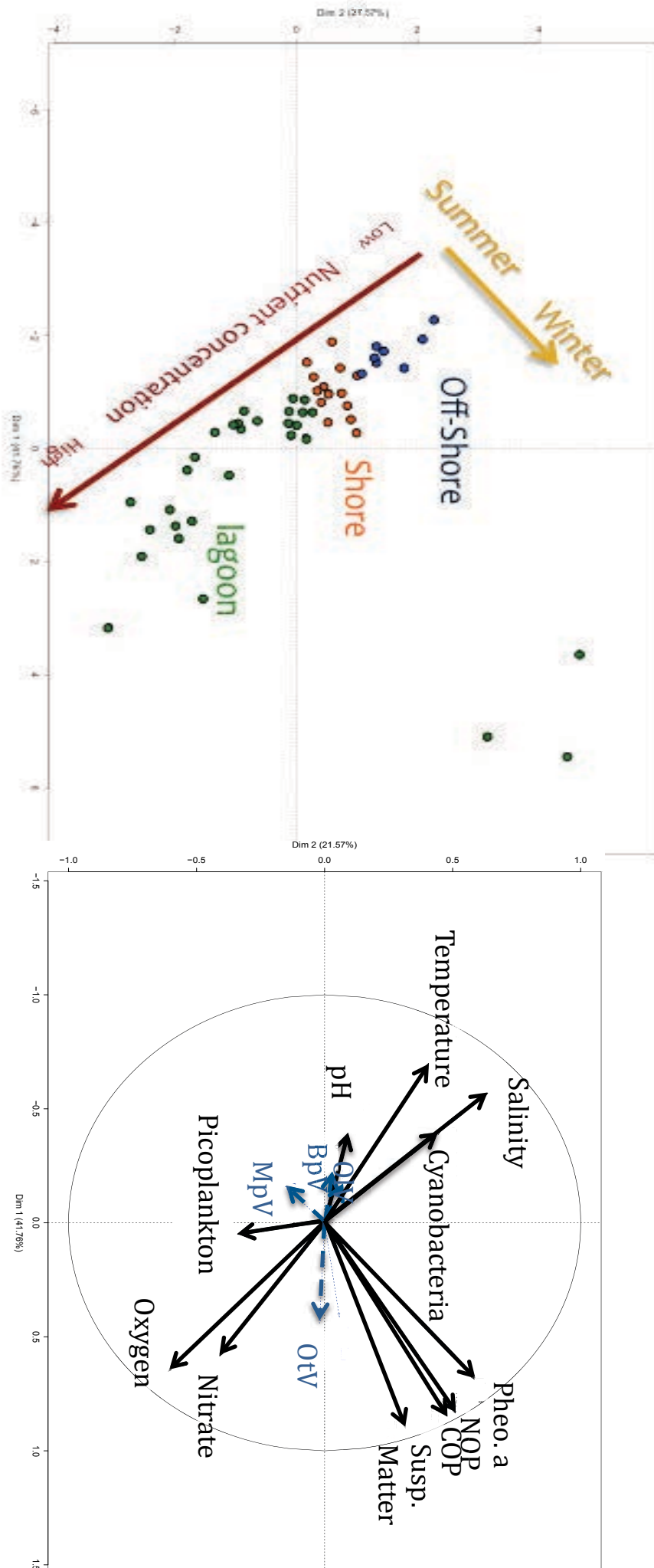


Figure 2: PCA based on environmental factors. Axis1 represents 41.76% of the total variance, and Axis 2 21.57%

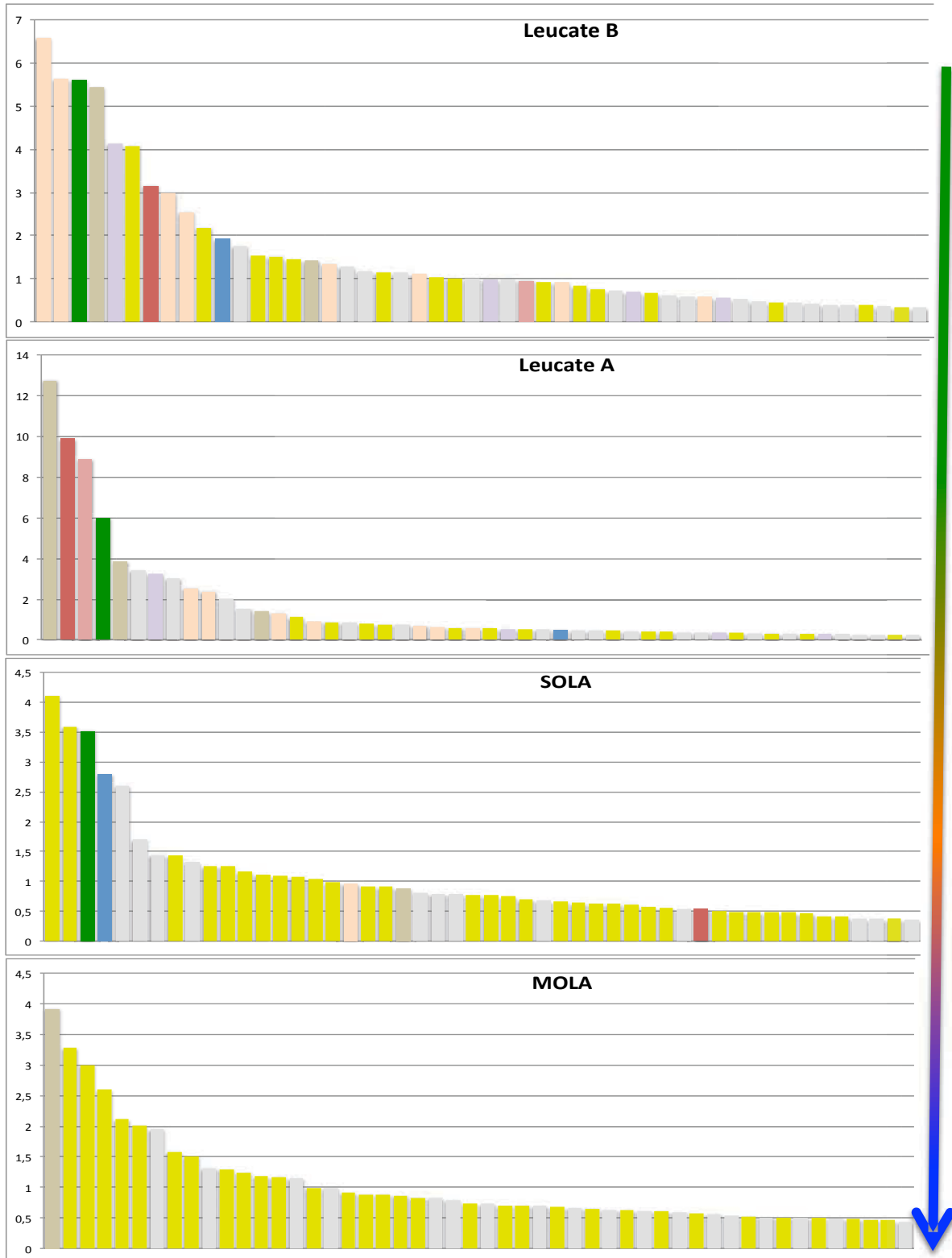
3. 4 Virus and eukaryote communities

The rank-abundance diagram of putative eukaryotic and virus sequences based on PolB and 18SV4 genetic markers displayed a clear change in communities, related to the trophic gradient (Figure 3). In the lagoon (Leucate A and B) and SOLA stations, *Micromonas spp.*, *Ostreococcus (tauri and mediterraneus)* and *Bathycoccus spp.* were represented in the 50 first eukaryotic OTUs, while in MOLA, these genera were almost absent (*Micromonas*: 86th, *Bathycoccus*: 110th and *Ostreococcus*: 421st OTU).

The change in eukaryotic community seemed to be linked to the trophic gradient, except for *Micromonas*, which was always represented in the four first OTUs. Leucate A and B showed the highest changes between *Ostreococcus* and *Bathycoccus* populations whereas, *Bathycoccus* was less well represented, in comparison to *Ostreococcus tauri* and *O. mediterraneus*, which were more abundant within the lagoon. At SOLA, *Bathycoccus* was better represented than in the other stations. This switch in host communities along the trophic gradient was also observed in virus communities. *Ostreococcus* viruses were mainly found in the lagoon and became less and less well represented along the trophic gradient. This observation was inverted for *Bathycoccus* viruses, whereas *Micromonas* viruses were represented in all four stations.

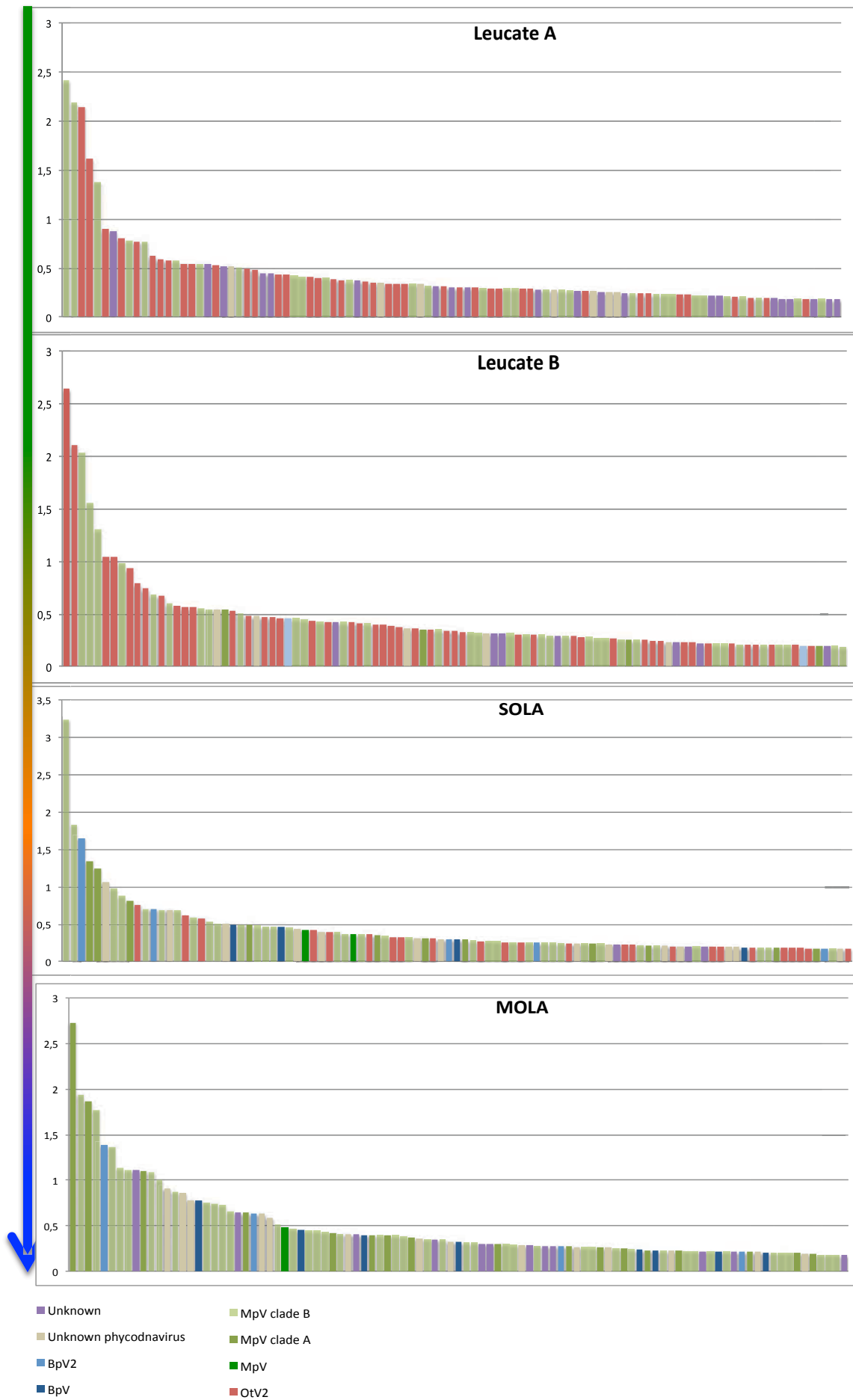
Besides, although analyses with MCP and 18SV9 genetic markers provided congruent results, they were less informative because a lower taxonomic resolution was reached with these markers (results not shown).

A



- *Ostreococcus mediterraneus*
- *Ostreococcus tauri*
- *Micromonas*
- *Bathycoccus*
- *Trebouxiophyceae*
- *Chaetoceros*
- *Cryptophyceae*
- *Dinophyceae*
- Other

Figure 3: Abundance OTUs percentage for host (V4; 50 first) (A) and Prasinovirus (PoLB; 100 first) (B) in the different sampling stations. Trophic gradient is represented by colored arrow (green = eutrophic to blue = oligotrophic), Mamiellophyceae putative hosts and viruses are represented (red = *Ostreococcus*; green = *Micromonas*; blue = *Bathycoccus*)



3.2 Station clustering

UPGMA dendrogram based on relative abundance of Mamiellophyceae OTUs (V4 and V9 regions, Figure 4) showed congruent patterns for both 18S regions, clustering OTUs in three distinct group that segregate marine and lagoon sites. More precisely, most marine sites were clustered in one group, while lagoon sites were separated in two clusters according to seasonality (spring-summer group and autumn-winter group) with some marine sites distributed over these two clusters.

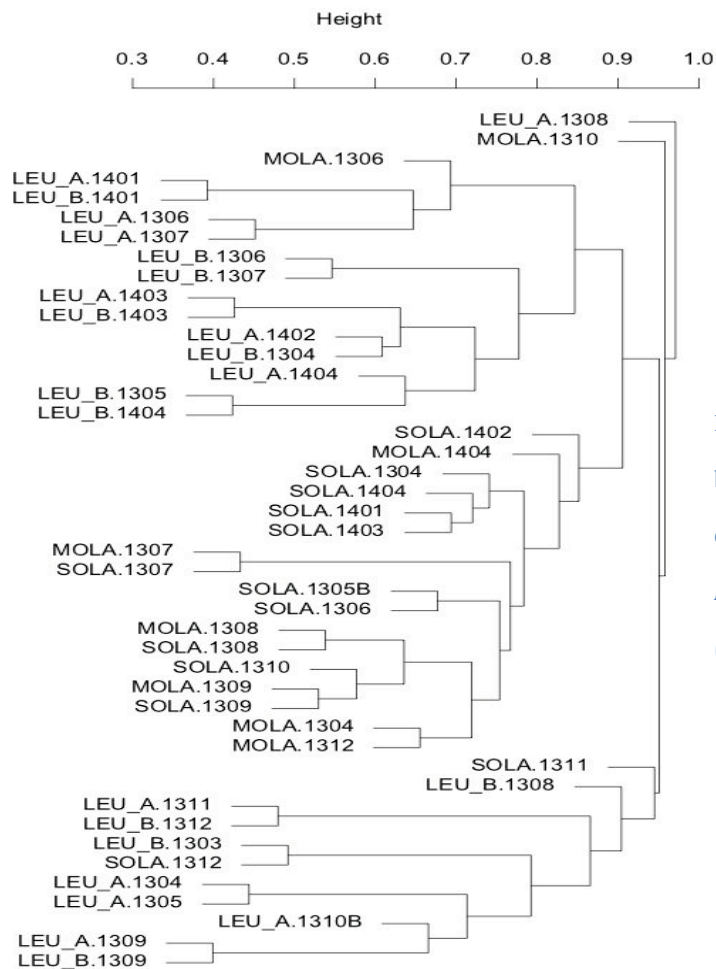
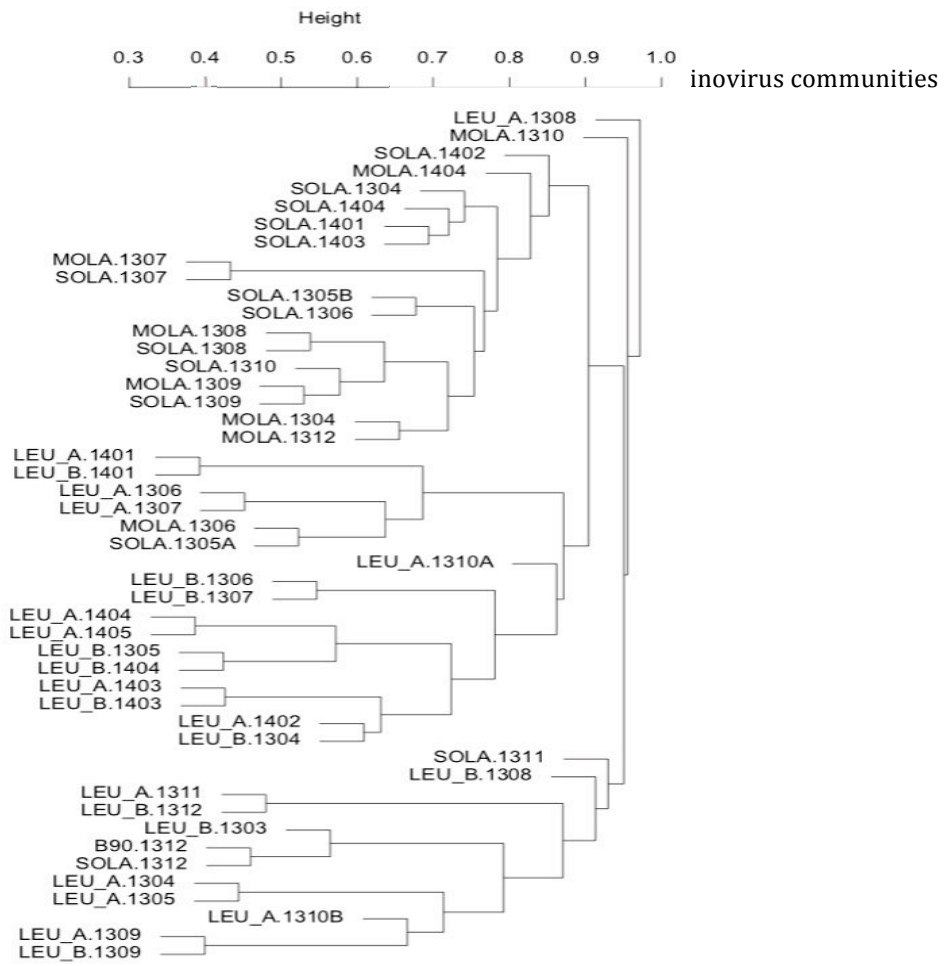


Figure 4: UPGMA clustering based on Bray-Curtis dissimilarity indices on *Mamiellophyceae* OTUs. V9 (up), V4 (down)

For Prasinovirus OTUs, UPGMA dendrogram showed similar results between PolB and MCP, with two distinct clusters which separate marine and lagoon OTUs (Figure 5). The lagoon cluster was comprised of two subgroups distinguishing winter-spring and summer-autumn groups.

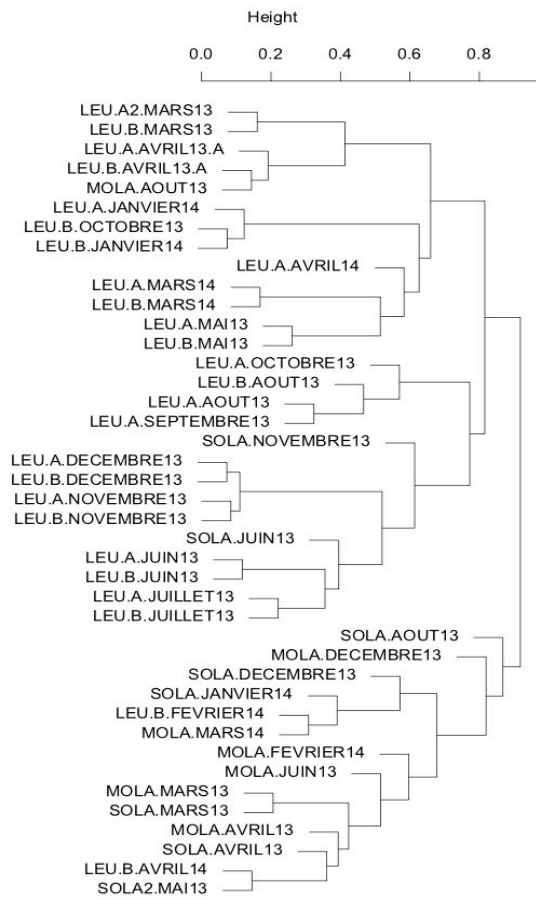
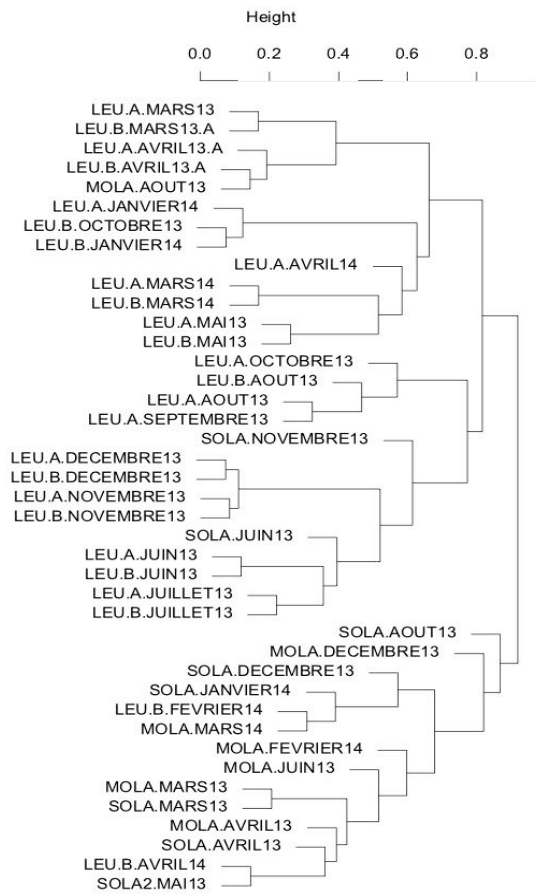


Figure 5: UPGMA clustering based on Bray-Curtis dissimilarity indices on *Prasinovirus* OTUs relative abundance. PolB (up), MCP (down)

3.3 Spatial segregation of prasinoviruses

In the Venn diagrams (Figure 6), representing the percentages of OTUs shared between the different sampling locations, *Micromonas* and *Bathycoccus* viruses OTUs were mainly shared by all environments, and this common fraction contained respectively 43.4% and 40.9% of MpV and BpV OTUs. Around 3% of these OTUs were specific to a single environment. Furthermore, even though MpV OTUs were found everywhere, an important proportion was shared only within the lagoon (6.1%) and between the lagoon and the coastal area (9.6%). Surprisingly, 30.6% of BpV OTUs were shared between the sea sampling sites and Leucate B, whereas only 0.4% were shared between the sea and Leucate A.

In comparison, the distribution of *Ostreococcus* viruses is different. These viruses are mainly shared between lagoon and coastal stations, this common fraction representing 47.1% for OtV and 28.9% for OIV. Moreover, 22.9% of the OIV OTUs are shared between SOLA and Leucate A, suggesting that OIVs are preferentially present in coastal areas. No OTUs related to OIV or OtV were specifically found at MOLA. No OTUs related to OIV or OtV were specifically found at MOLA.

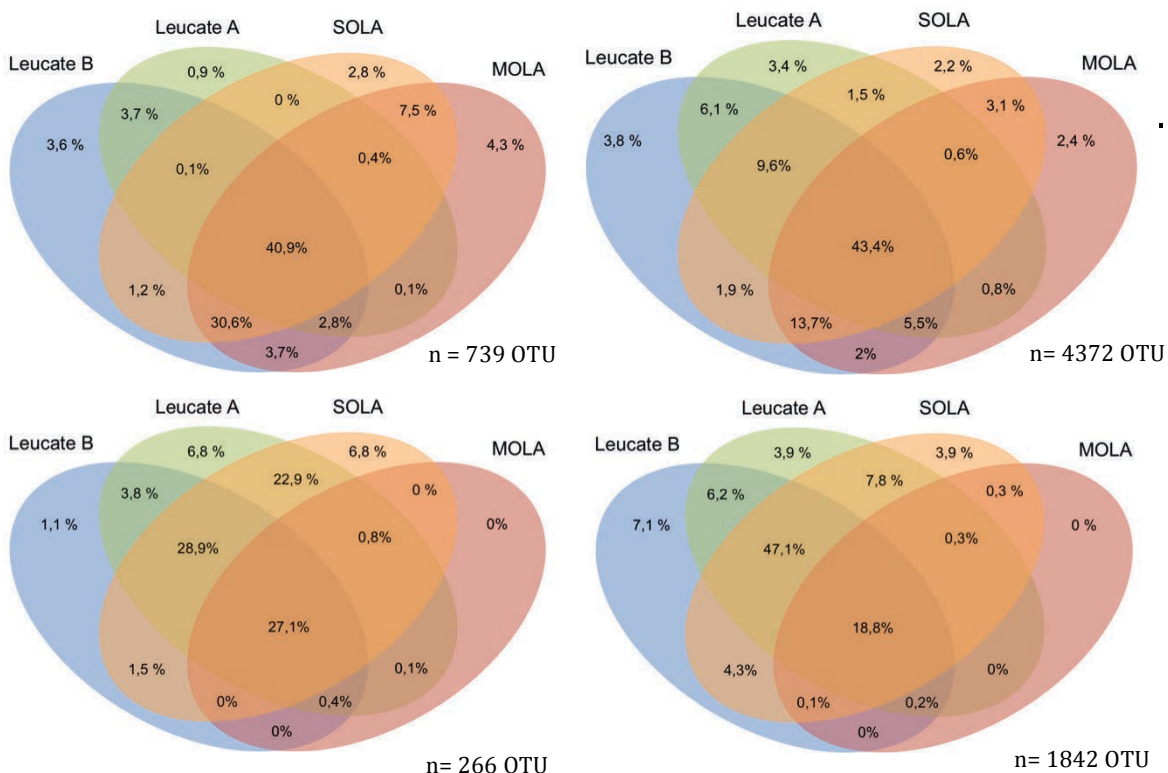


Figure 6: Venn diagrams representing Prasinovirus OTU percentages shared between stations. Upper-left = BpV, Upper-right = MpV, Bottom-left = OIV, Bottom-right = OtV,

3.5 The relationship between communities and environment

Temporal projections revealed annual variation in environmental variables (Figures 7-8). Lagoon stations showed higher environmental variations than marine stations, and variation was higher near the coast than offshore. Moreover, the structure of Mamiellophyceae communities during the year closely followed the variation of the environmental factors with the exception of Leucate A, (Figure 7). The pattern of variation of the viral OTUs shows the same variation as that of the environmental factors (Figure 8). In fact, the similarity between environment and viral community structure was lower than for the Mamiellophyceae community structure. The co-structure between viral OTUs and environment at Leucate A was different between March 2013 and July 2013, while viral communities and environmental variations were clearly linked after July 2013. Differences in the structure were also observed in Leucate B, SOLA, and MOLA but in a less marked fashion (from June to October for Leucate B and August for SOLA and MOLA), and appeared to be linked to higher temperatures.

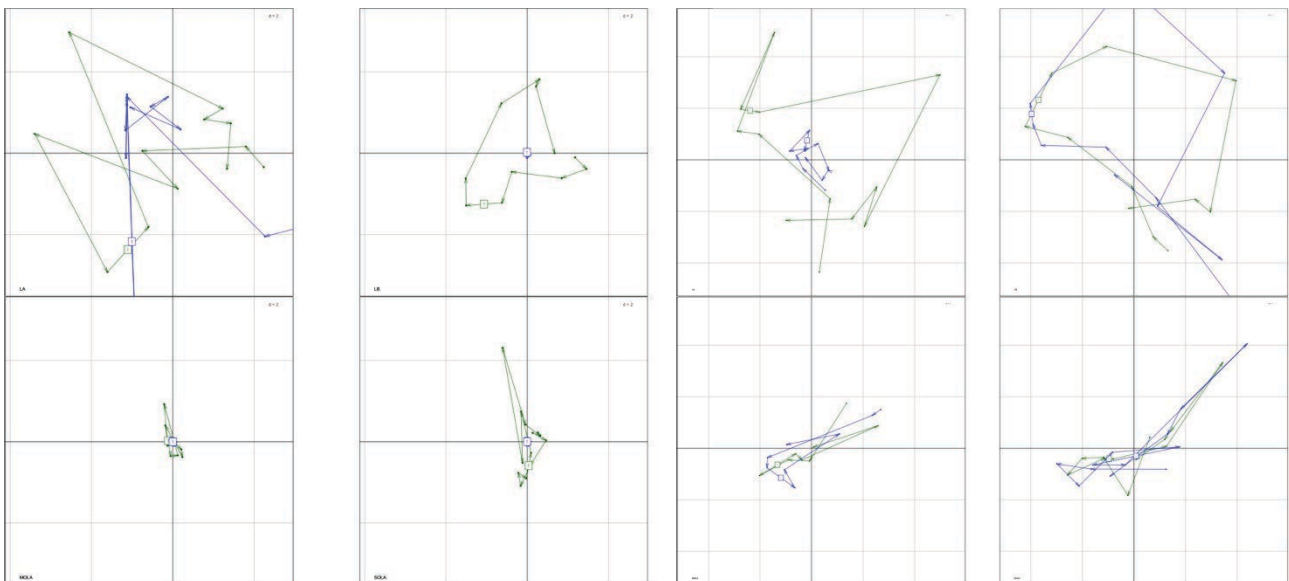


Figure 7: Trajectory factor maps from the STATICO program representing the structure of the variation in the same factorial space for Mamiellophyceae OTUs and environmental variables. Green line = environmental factors temporal structure, blue line= Mamiellophyceae OTUs temporal structure. Right group: data not normalized (Leucate A structure clearly visible), Left group: data log-normalized (the structure for stations other than Leucate A is then clearly visible). For each group, Upper-right = Leucate A, upper-left = Leucate B, Bottom-right = MOLA, Bottom-left = SOLA

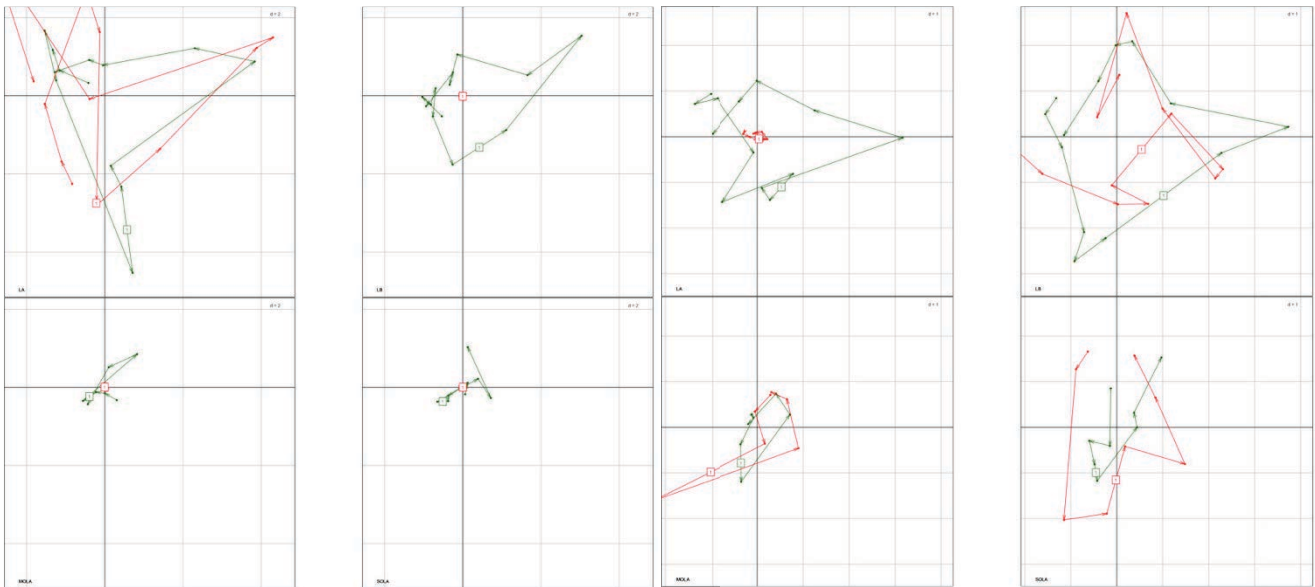


Figure 8: Trajectory factor maps from the STATICO program which represented variation structure in same factorial space of Prasinovirus OTUs and environmental variables. Green line = environmental factors temporal structure, Red line = Prasinovirus OTUs temporal structure. Right: data not normalized, Left : data normalized with $\log(x+1)$. For each group, Upper-right = Leucate A, upper-left = Leucate B, bottom-right = MOLA, bottom-left = SOLA

3.6 Virus distribution and environmental variables

A canonical correspondence analysis (CCA) based on viral OTUs showed that spatio-temporal segregation accounted for 41.1% of the variation (Figure 9). Three groups of OTUs could be distinguished, structured by environmental factors. Lagoon stations were distant from the marine stations, principally based on salinity, cyanobacterial abundance, and the abundance of *Bathycoccus* and *Micromonas* clade B. Within the lagoon, a group was formed by stations from August to October at Leucate A and August at Leucate B, distinct from the other stations. This first group was characterized by higher sea surface temperatures and ammonium concentrations. Moreover, OTUs composed of unknown prasinoviruses/viruses characterized these lagoonal time points.

Ostreococcus viruses and *Bathycoccus* viruses tended to be more abundant in lagoon and marine environments, respectively. In contrast, *Micromonas* viruses formed three different groups, one related to the lagoon stations, one related to the marine stations, and a third one found in both.

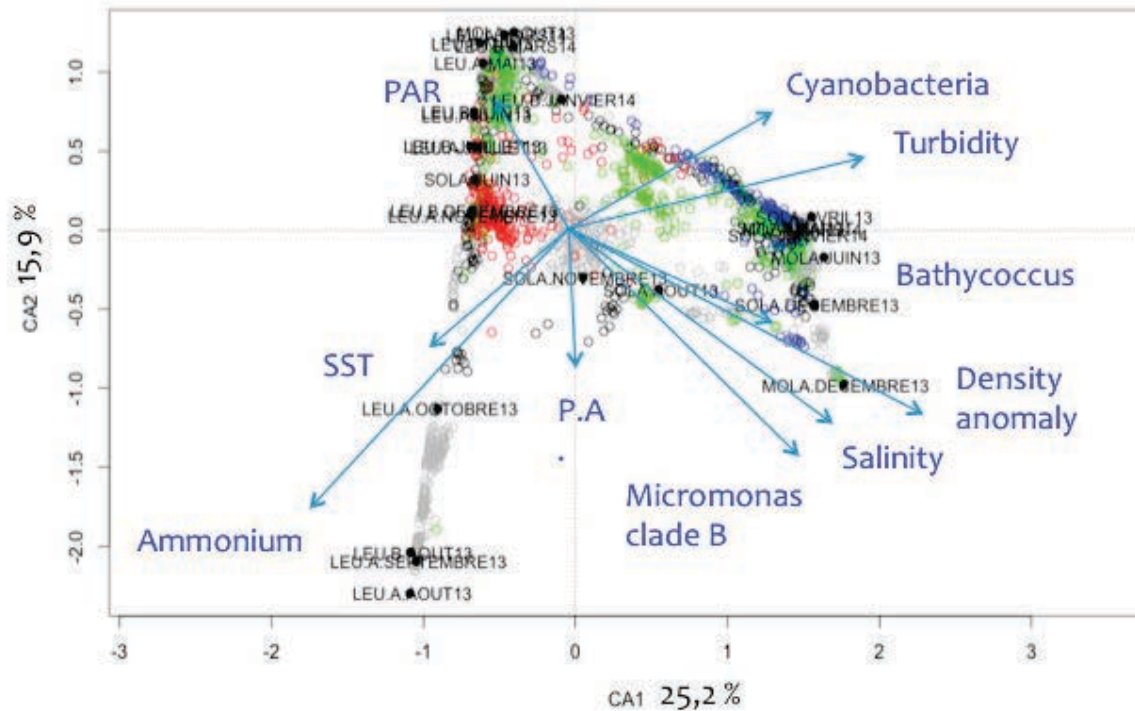


Figure 9: Canonical correspondence analysis on normalized virus OTUs and environmental variables for all sites. Blue arrows represent environmental variables with host abundance. Colored spots represent virus OTUs (Green = *Micromonas* viruses, red = *Ostreococcus* viruses, blue = *Bathycoccus* viruses, grey = unknown OTU, black = unknown Prasinovirus)

The CCA based only on viral OTUs and environmental factors at Leucate A explained 50.2% of the total variation (Figure 10), and showed only a temporal distribution. From January to May, a complex viral community containing all of the prasinoviruses was observed. These time points are distributed along pH and MES gradients and characterized by higher oxygen values than in August, September and October.

June and especially July are characterized by OTUs classified as *Ostreococcus* viruses. During this period, pH and MES concentrations explained much of the distribution of OTUs. In contrast, during August and September, the abundance of known prasinoviruses (OIV, OtV, BpV, MpV) is low (20% of all viral sequences amplified, data not shown). These samples showed a higher relative abundance of unknown viruses with a distribution following a group of correlated variables including temperature, nitrite, ammonium and salinity.

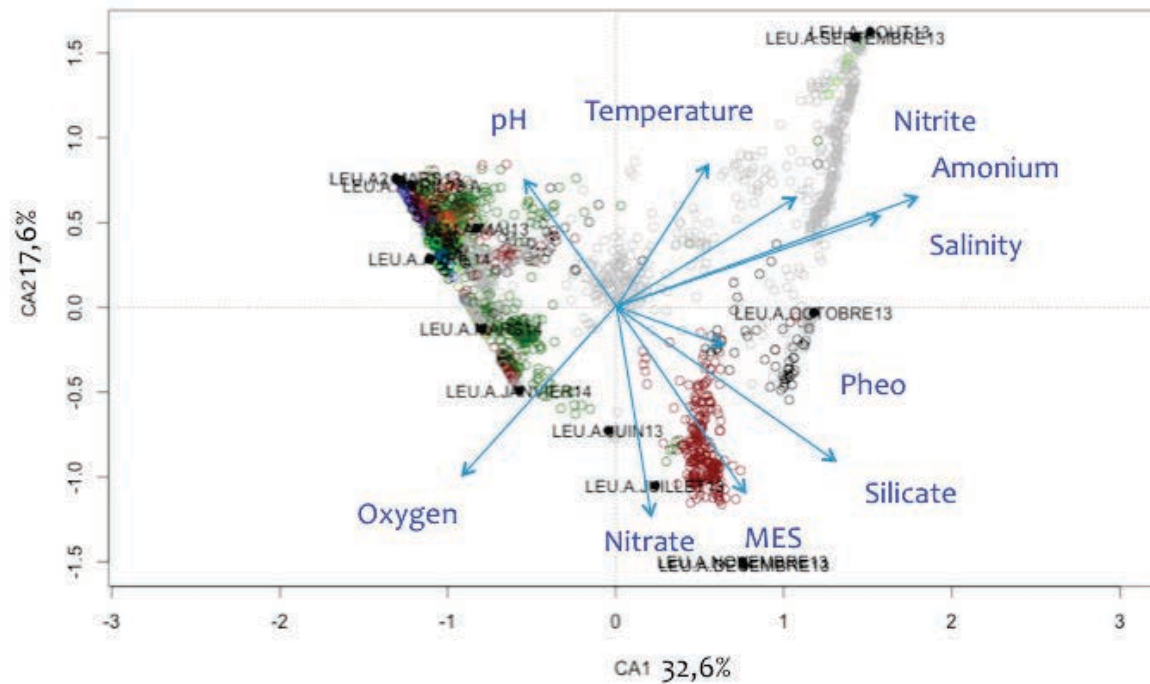


Figure 10: Canonical correspondence analysis on normalized virus OTUs and environmental variables at Leucate A. Blue arrows represent environmental variables. Colored spots represent virus OTUs (green = *Micromonas* viruses, red = *Ostreococcus* viruses, blue = *Bathycoccus* viruses, grey = unknown OTU, black = unknown Prasinovirus)

As for Leucate A, the CCA performed on viral OTUs with environmental factors at SOLA explained 51.6% of variation (Figure 11), with a temporal structure. The end of spring (May and June 2013) was related to the highest *Ostreococcus* virus relative abundances. Mainly *Micromonas* and *Bathycoccus* viruses were observed at the other time points. The distribution observed in August 2013 was different from the other points based on the concentration of pheophytine a. Furthermore, the viral community was not different from the communities found between November 2013 and April 2014 (mainly consisting of MpV, BpV), but the concentration of pheophytine a was higher than for the rest of the year.

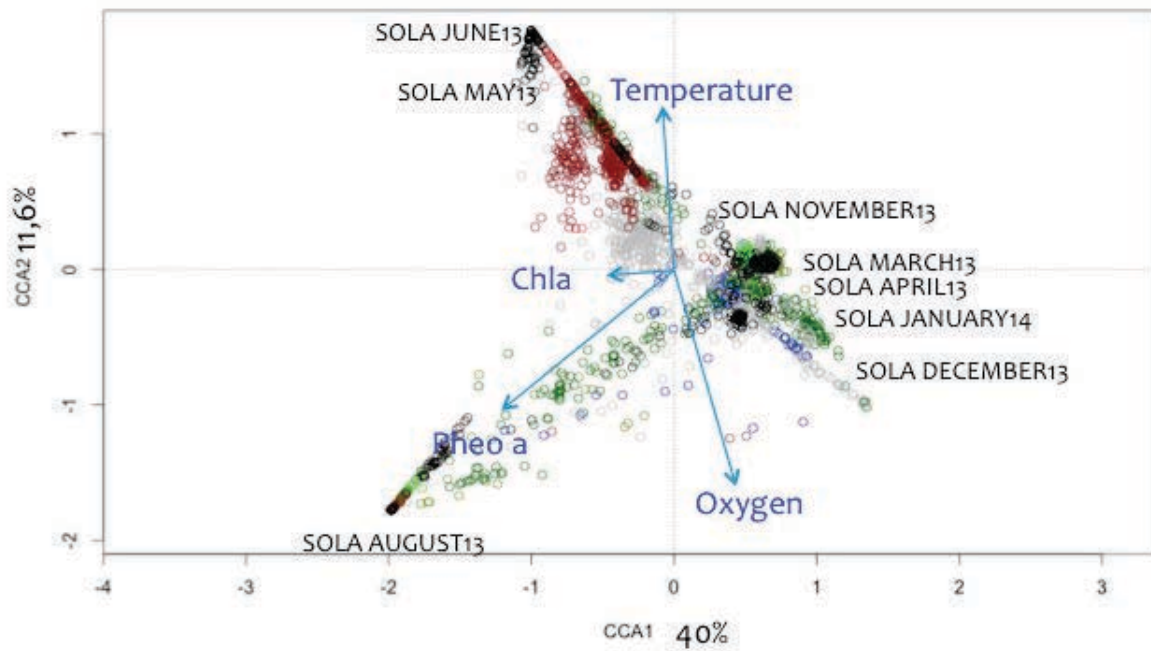


Figure 11: Canonical correspondence analysis on normalized virus OTUs and environmental variables for all sites. Blue arrows represent environmental variables with host abundance. Colored spots represent virus OTUs (Green = *Micromonas* viruses, red = *Ostreococcus* viruses, blue = *Bathycoccus* viruses, grey = unknown OTU, black = unknown Prasinovirus)

The CCA made with the MOLA sample (Figure 12) explains 62% of the variation, and displayed the same temporal structure as in SOLA. *Ostreococcus* viral OTUs were in limited proportion only in August 2013. In addition, pheophytine a, oxygen and temperature explained the data structure as in SOLA, except during December 2013 when a higher number of unknown viruses was found. At MOLA, more OTUs related to unknown Prasinoviruses were found than in the other environments.

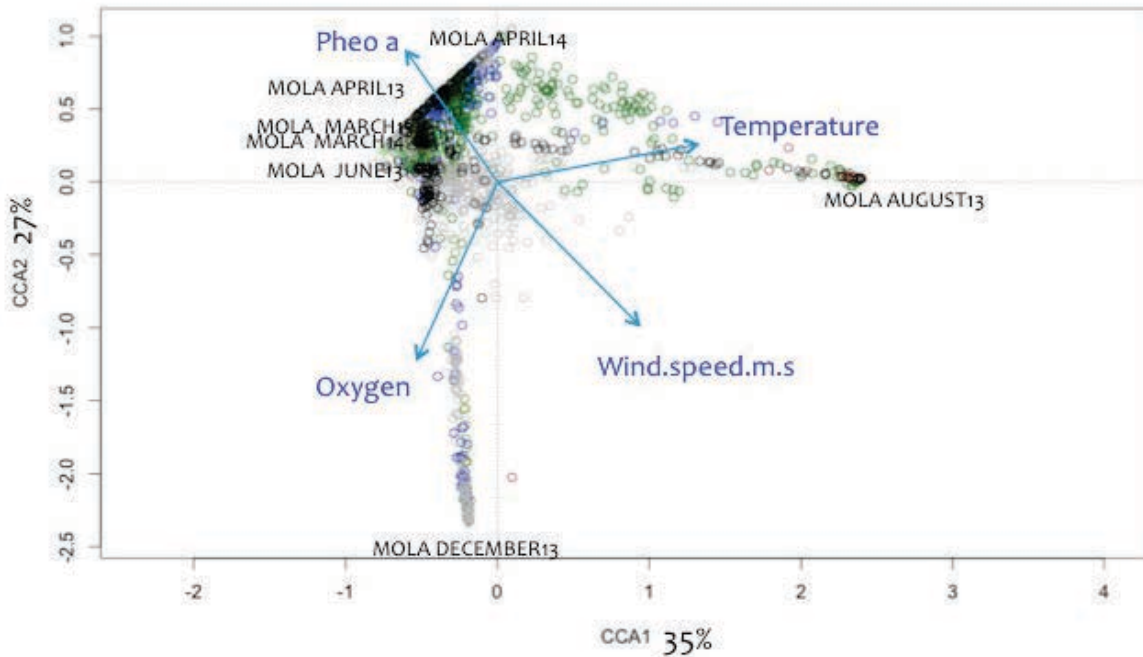


Figure 12: Canonical correspondence analysis on normalized virus OTUs and environmental variables at MOLA. Blue arrows represent environmental variables. Colored spots represent virus OTUs (green = *Micromonas* viruses, red = *Ostreococcus* viruses, blue = *Bathycoccus* viruses, grey = unknown OTU, black = unknown Prasinovirus)

3.7 The link between virus and host communities

The CCA performed on all prasinoviruses and Mamiellophyceae data (Figure 13) suggests complex relationships between these two communities. The distributions of *Micromonas* and of their putative viruses showed a similar structure and a co-occurrence. In comparison, *Bathycoccus* and *Ostreococcus* viruses and their putative hosts tended to exhibit different distributions. For example, around April 2014, many *Bathycoccus* viruses were present but the closest Mamiellophyceae group was not related to *Bathycoccus*. In September 2013, many *Ostreococcus tauri* were present but most viruses were related to *Micromonas* viruses.

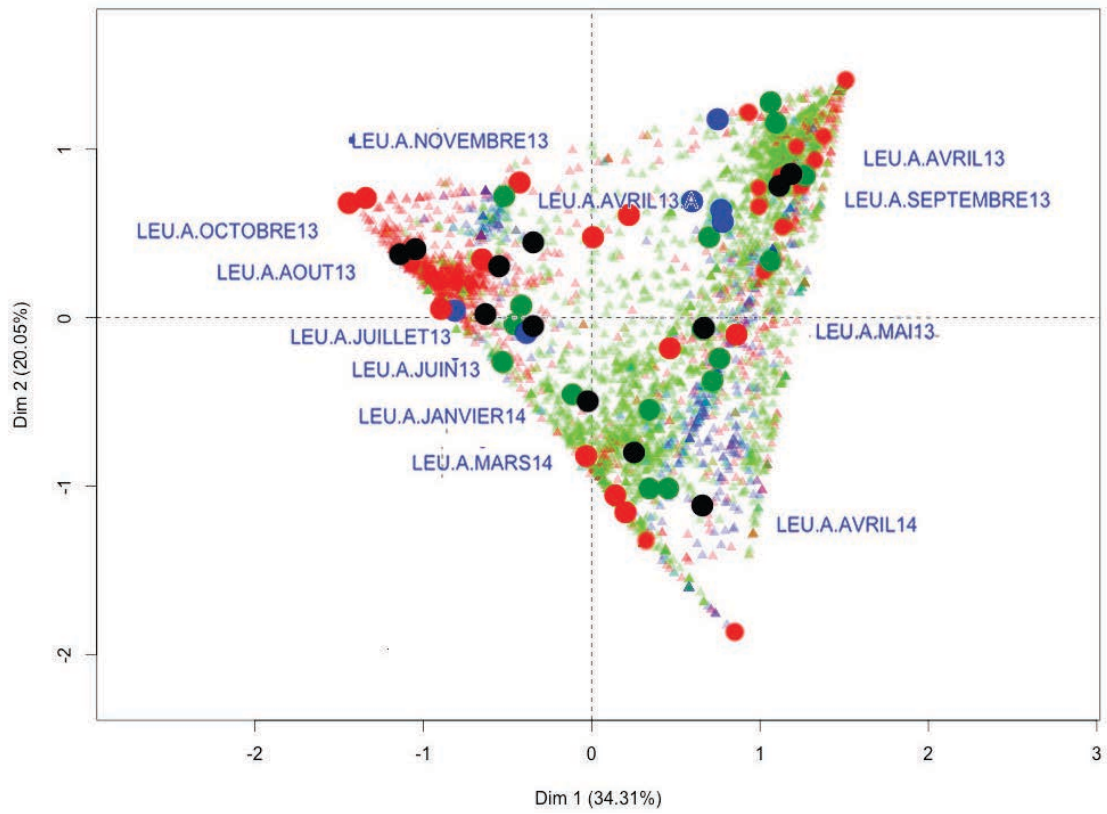


Figure 13: Correspondence analysis on the relative abundances of Prasinovirus and *Mamiellophyceae*. Black spot = Station, Colored spot = *Mamiellophyceae* OTU and Triangles = Prasinovirus OTU. Green means related to *Micromonas*, blue related to *Bathycoccus*, red related to *Ostreococcus*.

DISCUSSION

4.1 Contrasted environmental dynamics

The effect of the environmental factors, and especially meteorological conditions, is clearly different for the four sampling stations as previously noted by Hervé (1978). Geomorphology and water column depth could explain the observed variations of these environmental factors between lagoon and marine systems (Newton and Mudge, 2003). Indeed, the thermal inertia, related to depth, maintains a higher temperature stability in shore and offshore stations throughout the year than in lagoon sites, which displayed a greater seasonality between winter and summer. Furthermore, increasing depth reduces wind impact. In lagoon stations, winds exert a strong impact on resuspended matter (Wilke, 1999) and water currents (Cesmat, 2006). During the summer, sea winds and evaporation generate currents from the sea to the lagoon, but during the other seasons, a strong terrestrial wind, the “Tramontane,” is prevalent and induces water outlet from the lagoon to the sea (Bouin *et al.*, 2012; Cesmat, 2006). This current may impact the connectivity between the sea and the lagoon. During winter, windy weather induces seawater convection at MOLA (Severin *et al.*, 2014). The geographic situation of these sites explains their trophic conditions: Mediterranean lagoons are semi-closed environments, inducing higher nutrient concentrations by confined catchment basin supplies (Jacques *et al.*, 1975; Cépralmar, 2000; Figure 1). In addition, this enclosed state and low depth increase the variations within lagoons throughout the year and between years following the variations of inter-annual meteorological conditions (Cesmat, 2006). SOLA is a mesotrophic station and nutrient influx occurs mainly during winter. In contrast, the offshore station MOLA is far from the shore and thus oligotrophic. This geographical location of MOLA reduces nutrient supplies from the shore and likely promotes specific microbial communities, well adapted to these stable and oligotrophic conditions. Biotic variables such as concentrations of cyanobacteria, picoplankton, *Bathycoccus*, *Micromonas* indicate eutrophic or oligotrophic conditions. Higher values of cyanobacteria and *Bathycoccus* are found in oligotrophic areas than in eutrophic areas and picoplankton and *Micromonas* abundance present higher value in eutrophic than oligotrophic environments. Picoplankton is more abundant in eutrophic environments (Viprey *et al.*, 2008). In contrast, *Bathycoccus* and cyanobacteria are known to be more

abundant in oligotrophic areas than in eutrophic environments (Wu *et al.*, 2014; Viprey *et al.*, 2008). Environmental trophic state seems then to be the main constraint for algal community (Litchman et Klausmeier, 2008). Despite this trophic state, eukaryote communities are related to salinity state as shown on the dendrogram based on V4 and V9 OTU. Lagoons A and B are clustered with offshore stations during spring and summer periods when salinities are similar.

4.2 Spatio-temporal *Prasinovirus* distribution and dispersion between environments

The present data suggest that the spatial dispersion of Prasinovirus communities is not homogeneous. Whereas, more than 40% of the MpV and BpV diversity is found within all sampling stations, OIV and OtV population seem limited to lagoon nutrient-rich systems (respectively 27.1% and 18.8% of their diversity were found in all sites; Figure 4). This suggests processes of general dispersion of viruses among these areas. Previous studies point to such a connection between Mediterranean lagoon and coastal environments, observed among their viral communities (Bellec *et al.*, 2010; Clerissi *et al.*, 2012). However, in spite of this global distribution, OIV and OtV populations seem limited to nutrient-rich environments because, even though a small part of their diversity is shared by all environments (respectively 27.1% and 18.8%), most of them are found in lagoon and coastal stations. So, if the global distribution occurred in the same way for prasinoviruses, population maintenance must be reduced in MOLA and must be influenced by viral dispersion and host community (Zingone *et al.*, 1999).

Prasinovirus abundance is different between lagoon and open sea (Figure 2). This difference is mainly related to the presence of OtV and BpV OTUs (Figure 6). The diversity of OtV is higher in both coastal and lagoonal areas than at the offshore station. In contrast, the BpV diversity is higher at the offshore station than in lagoonal environments. Furthermore, even if most of the BpV and MpV diversities are shared among all the environments, their distribution shows clearly that the distribution of virus groups is related to these areas (Figure 8). MpV exhibited three main groups, that

can be related to host specificity (Baudoux *et al.*, 2015). Limited MpV communities in sea or lagoon areas can have a smaller distribution range and population maintenance is limited during dispersion. In contrast, the MpV communities shared between these environments based on abundance may be able to infect a host range higher than other MpV groups and their dispersion is not limited, as observed by Clerissi *et al.* (2012) on OtV communities.

The spatial distribution of prasinoviruses seems constrained by time, especially for the OtV population. The increase of viral communities during spring (Figure 8) is related to the increase in host communities (Tuomi *et al.*, 2006). During this period, virus communities were composed of all known prasinoviruses, even though OtV is less represented than in the other groups. During the June/July period, *Ostreococcus* viruses dominated Prasinovirus communities in lagoon and coastal areas (Figure 9-10), which could be related to an increase in host abundance (Lebretonchel *et al.*, submitted). During August and September, host abundance strongly decreases, especially in the lagoon, and an abundant population of unknown viruses dominates the viral community. The main explanation for this pattern may be adverse environmental conditions because the abundance of *Ostreococcus* was found to be less variable (between 5,000 and 12,000 cell/ml) as previously noted by Clerissi *et al.* (2014). The latter authors explained their observations by filtration and protocol problems during viral DNA extraction, but this is not the case in the present study. In addition, *Ostreococcus* viruses might disperse from the lagoon to the sea when windy days promote water outlet (Cesmat, 2006). The higher abundance of OtV in lagoon than in the sea areas suggests that this environment can play the role of a natural chemostat to increase viral development.

4.3 Link between *Prasinovirus* and Mamiellophyceae communities

Even though the Mamiellophyceae were found in all environments (Figure 3), *Ostreococcus* is more diverse and abundant in the lagoon, compared to *Bathycoccus*, which exhibited an inverse distribution than *Ostreococcus*. In general, these

distributions are linked to trophic state (Zhu *et al.*, 2005, Viprey *et al.*, 2008). *Ostreococcus tauri* is abundant in Mediterranean lagoons (Rodriguez *et al.*, 2005; Subirana *et al.*, 2013), especially in the Thau lagoon from where it has been isolated (Courties *et al.*, 1994).

Bathycoccus populations were described mainly in oligotrophic areas (Wu *et al.* 2014; Vaultot *et al.*, 2012). (Figures 4-5). Salinity explains the station clustering based on Mamiellophyceae OTUs but this factor is certainly a signal of seasonality, corresponding to an *Ostreococcus* bloom (O’Kelly *et al.*, 2003). The diversity of *Micromonas* is much higher (2797 OTUs) than for *Bathycoccus*, and *Ostreococcus tauri* and *lucimarinus*, (respectively 1271, 1194 and 116 OTUs). *Micromonas* is known to be ubiquitous and particularly present in nutrient-rich environments (Viprey *et al.*, 2008, Wu *et al.*, 2014) as in the English Channel (Not *et al.*, 2004). Since Mediterranean lagoons and the open sea areas are connected, as suggested by the repartition of viral communities, the maintenance of algal communities during dispersion events may be more limited by environmental conditions than viral dispersion. These dispersion characteristics were observed on dendrogram where seasonal patterns were not linked to station clustering (Figure 5). This difference between lagoon groups is probably linked to bloom events, occurring during spring when viruses may be recruited and enhanced by host growth.

Bathycoccus and BpV highlight a particular distribution pattern. If *Bathycoccus* is really adapted to oligotrophic areas (Vaultot *et al.*, 2012), it is not expected to be found in high abundance in lagoons. However, surprisingly, we found this species at Leucate B (the furthest station from the sea) but not at Leucate A. This particular pattern might be due to an autochthonous community or an ancestral community isolated from the sea community, which then adapted to lagoonal environmental conditions, suggesting existence of two different ecotypes as found by Vaultot *et al.* (2012). This is supported by the present observation that 30.6% of the BpV diversity is shared between marine stations and Leucate B (Figure 6). The marine Mamiellophyceae communities seem less variable (this could be because of the higher salinity average at Leucate B than Leucate A, data not shown), and may enhance the viral recruitment during dispersion all over the year and the persistence of populations after dispersion. The percentages of viral and eukaryote OTUs suggest a link between these community.

The temporal analysis of co-occurrence supports the notion that hosts and viruses are usually found together (Figure 13). While this co-occurrence is globally observed, there are some potential missing co-occurrences between viruses and their hosts. In September and April, higher abundances and diversities of *Ostreococcus* (compared to other Mamiellophyceae) are found but their viruses seem absent. April 2014 was characterized by a higher spatial distribution of BpV than other viruses, but *Bathycoccus* was not found. If we considered that only host community drives the viral community, we expect both to be represented at the same time (O’Kelly, 2003), even if small time gaps can be observed (Baudoux *et al.*, 2015). This absence of complete co-occurrence can be related to environmental variations, which can select viral communities, or bring new communities by water current.

4.4 Link between *Prasinovirus* communities and environmental factors

The present study suggests that the Mamiellophyceae community is structured by environmental factors (Figure 3), except for the lagoon station Leucate A, maybe because of factors not considered here, such as water current input occurring during summer (Cesmat *et al.*, 2002). This strong link between Mamiellophyceae and environmental variables is expected because phytoplanktonic communities are mainly driven by environmental conditions (Gao and Song, 2003; Gieskes and Buma, 1997). The development of Mamiellophyceae community is related to the PAR variable, which is linked to the phytoplanktonic bloom (Figure 8). This increase in host abundance could induce the development of viral community (Bratbak *et al.*, 1990; Castberg *et al.*, 2001), and favour viral recruitment and production (Short *et al.*, 2010) in all environments.

Prasinovirus community structure is mainly related to environmental dynamics (Fig 7). However differences in structure were observed between viral community structure and environment during the hot period, suggesting that the community changes are related to temperature. The pattern observed in Leucate A, where viral community structure is not linked with the environment (during May to August), is probably due to seawater input. The decrease of OtV in lagoon during August and September could be

explained by environmental conditions unfavourable to Prasinovirus communities, when temperature is higher and oxygen lower than other month. During this period, the *Ostreococcus* community was monitored by qPCR (Lebretonchel *et al.*, submitted), where they were found highly abundant and is visible in this study on the CA (Figure 12). The decrease in prasinoviruses seems clearly linked to environmental conditions. Temperature and oxygen influence the viral community: an increase in temperature has been shown to promote viral decay (Nagasaki and Yamaguchi, 1998) or enhance exoenzyme activities that could favour capsid degradation (Cornaladesi *et al.*, 2010); anoxic environments are known to harbour high viral concentration but also high viral decay (Cornaladesi *et al.*, 2007) and lower infection rates (Taylor *et al.*, 2003). Such environments may also enhance lysogenic strategies (Muck *et al.*, 2014). These combinations of factors could be related to a higher abundance in unknown viruses and indicate that in such environmental conditions, prasinoviruses exhibit a more important decay than the unknown viruses found here which seems resistant to environmental conditions. In addition, the impact of water current may partly explain the change in viral communities via export or dilution. To illustrate this latter point, an important community of unknown viruses was observed at MOLA in December (Figure 11), and during this period, convection processes are known to occur (Severin *et al.*, 2014). Such processes can induce the export of viruses from the surface to the bottom or advect viruses from the sediment. This community change mediated by current may also explain the variation observed at Leucate A during seawater input.

CONCLUSION

Prasinovirus communities seem mainly driven by host communities. We hypothesize that viral communities can be considered as a seed bank that is dispersed by abiotic factors (wind, water current). After viral dispersion, algal communities select virus recruitment and production. Environmental factors such as trophic conditions, luminosity, and temperature directly impact algal communities, then indirectly influence

prasinovirus communities. Finally, environmental conditions, such as temperature, may strongly influence the maintenance or the decay of viral communities.

Acknowledgements

This work was supported by a “Agence Nationale de la Recherche” grant “DECOVIR” N° ANR- 12-BSV7-0009. Hugo Lebretonchel benefited from a doctoral fellowship from the Université Pierre et Marie Curie.

Chapitre VI – Discussion générale et perspectives

I Discussion générale

Dans ce travail de thèse, l'étude des communautés de Prasinovirus et de Mamiellophyceae en lien avec les conditions environnementales a été étudiée à l'aide de plusieurs approches. Un protocole de qPCR a été développé afin de décrire la dynamique saisonnière de l'abondance de Mamiellophyceae dans les différents sites choisis au sein du Golfe du Lion. Nous avons mis en place un protocole spécifique de tri cellulaire basé sur la taille et la fluorescence des eucaryotes photosynthétiques, permettant la concentration et la sélection des communautés eucaryotes photosynthétiques environnementales. Enfin, nous avons analysé les séquences environnementales (Illumina MiSeq), en adaptant de précédents protocoles dédiés aux gènes ribosomiaux 18S et 16S aux amplicons des gènes viraux (PolB et MCP).

Dans le but d'augmenter la quantité d'OTU assignés dans notre base de données, nous avons complété ce travail par l'isolement et la caractérisation de nouvelles souches de Prasinovirus par le biais de culture de Mamiellophyceae. Les résultats préliminaires de ces travaux ne pouvant faire l'objet d'un chapitre à part entière, nous les présentons succinctement dans la partie ci-dessous.

1) Description de nouveaux Prasinovirus

Nous avons isolé trois nouvelles souches de Prasinovirus. Deux d'entre elles sont spécifiques à *Mamiella sp. nov.* (RCC 391) et la troisième infecte *Crustomastix stigmata* (CCMP 3273). Le genre *Mamiella* se situe à la base des Mamiellales (Marin et Melkonian, 2010). *Crustomastix* quant à lui est le genre le plus proche des Mamiellales au sein des Mamiellophyceae .

Le virus de *Crustomastix*, « CruV1 » a été isolé à partir d'une lagune de la région de Rio (lagune d'Araruama, 22.886°S, 42.168°W), et les virus de *Mamiella* « MaV1 », « MaV2 » ont été isolés dans le Golfe du Mexique, à Saint Petersburg (Floride) à l'ouest (St Peter Beach, 27.710°S, 82.740°W) et à l'est (Port, 27.759°S, 82.633°W). Nous avons caractérisé en partie ces virus grâce à l'utilisation de différentes méthodes: la microscopie électronique (Figure 1), un suivi de cinétique de lyse (M.O.I. = 100, Figure 3)), la détermination de leur taille génomique (Figure 2) ainsi que le séquençage de la polymérase virale.

CruV1 possède une capside icosahédrale de 105 nm de diamètre, commune aux Prasinovirus.

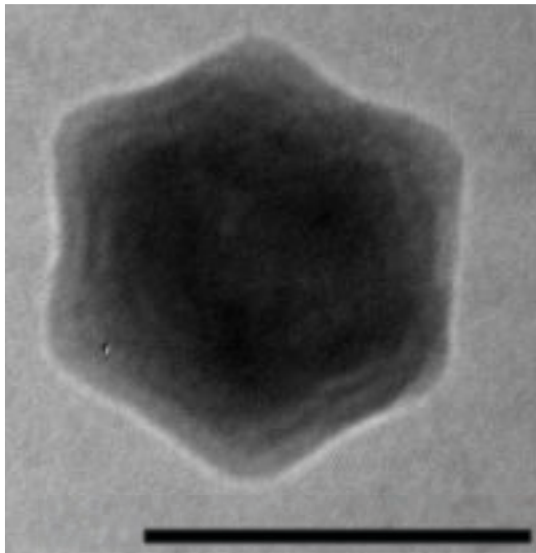


Figure 1 : Cliché de microscopie électronique de CruV1. La barre représente 100 nm.

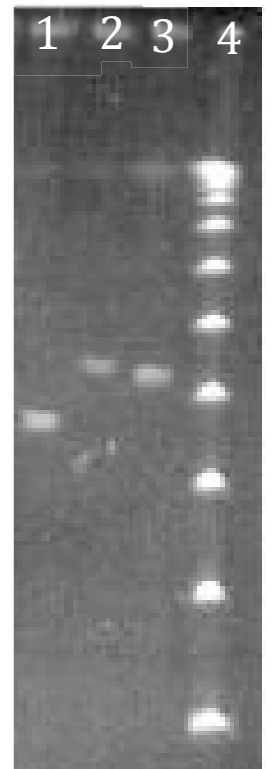


Figure 2 : PFGE sur les virus de CruV1 (1), MaV1 (3), MaV2 (2), marqueur lambda (4).

Cinétique Crustomastix Vs CruV1

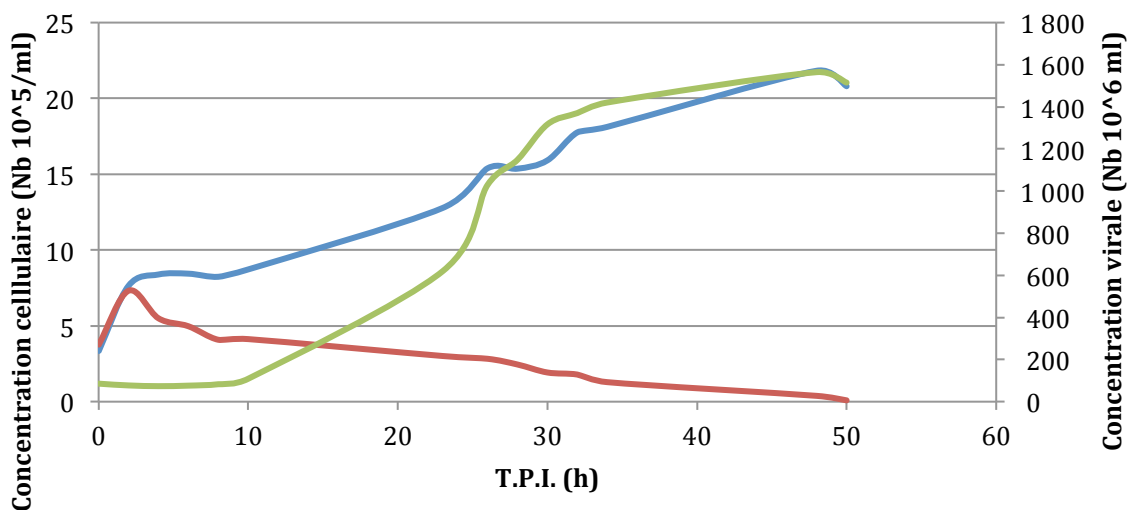


Figure 3 : Cinétique de lyse de CruV1. Culture témoin (bleu), culture infectée (rouge), abondance virale (vert). TPI : Temps post-infection.

Ces trois nouvelles souches de Prasinovirus ont une réplication lytique. CruV1 présente un temps de latence de 8 à 10h, comparable aux souches virales précédemment décrites (Baudoux et Brussaard, 2008 ; Brussaard *et al.*, 1999 ; Brussaard et Martinez, 2008 ; Derelle *et al.*, 2008). Les tailles des génomes de ces souches sont comparables à celles des autres Prasinovirus déjà identifiés (Figure 4), MaV1 et 2, avec des tailles respectives de 201 kb et 209 kb. En revanche, CruV1 possède un génome de taille plus réduite, 178 kb, inférieure à celle des plus petits génomes connus chez les Prasinovirus, OtV5 (186 kb, Derelle *et al.*, 2008) et MpV1 (184 kb, Moreau *et al.*, 2010).

Le séquençage partiel de la polymérase a été utilisé pour constituer la phylogénie de ces virus. Elle a révélé une position phylogénétique particulière pour ces nouveaux virus. La polymérase virale est un « bon » marqueur car ce gène contient des domaines conservés mais aussi des régions variables (Chen et Suttle, 1995 ; Tidona *et al.*, 1998) et permet une bonne représentation des relations phylogénétiques au sein des NCLDV (Yutin *et al.*, 2013). Les amorces AVS1 et AVS2 (Chen et Suttle, 1995) n'ont pas permis l'amplification de la polymérase (elles ont pourtant été établies pour amplifier la polymérase de Chlorovirus, a priori plus divergents), et ces séquences ont été obtenues par l'utilisation des amorces AAS1-AS4 (Clerissi *et al.*, 2014) utilisées durant notre étude des communautés naturelles. Ce changement d'amorces peut bien sûr être dû à une zone d'hybridation trop divergente de la séquence des amorces AVS1 et AVS1, mais aussi à la présence d'une intéine au sein de la séquence, limitant ainsi l'amplification du gène cible (Clerissi *et al.*, 2014). Le *blast* dans NCBI de ces séquences nous a indiqué qu'elles n'ont pas encore été rapportées. Seuls MaV1 et 2 possèdent des séquences proches respectivement à 86 % (*Micromonas pusilla* virus 12T, Mer du Nord) et 90 % (Unknown phycodnavirus, Hawaii). La phylogénie obtenue montre que les trois nouvelles souches isolées semblent être à la base des Prasinovirus (Figure 5). Cette position phylogénétique était présagée de par la position de leurs hôtes respectifs (*Mamiella gilva* et *Crustomastix minuta*) à la base des Mamiellales (Figure 5). Un signal de cospéciation significatif a en effet été estimé pour les Prasinovirus infectant les genres *Ostreococcus*, *Bathycoccus* et *Micromonas* (Bellec *et al.*, 2014). Cette phylogénie basée sur une séquence de 350 pb peut potentiellement être modifiée lors de l'acquisition prochaine des génomes entiers de ces trois nouvelles souches virales.

L'intérêt de l'isolement de ces virus était d'obtenir plus de séquences afin d'affiner nos analyses, en ajoutant ces séquences dans la base de référence utilisée pour l'identification des OTUs viraux. Cependant, aucune de ces séquences n'a été retrouvée dans les séquences environnementales obtenues lors de notre suivi temporel.

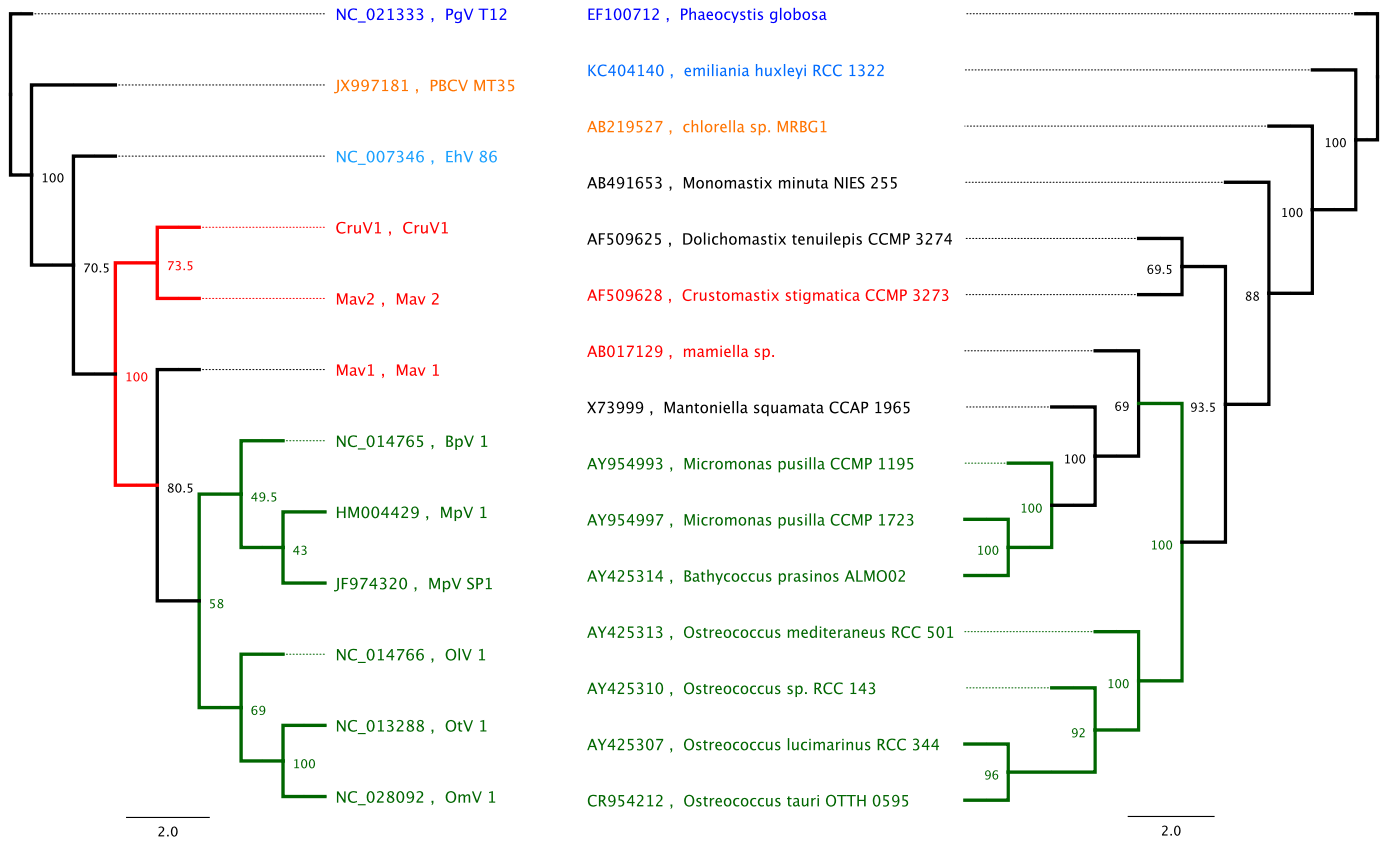


Figure 4 : Phylogénie des Prasinovirus et des *Mamiellophyceae*. Phylogénie obtenue à partir de la polymérase virale amplifiée avec AASI-AAS4 (gauche). Phylogénie des *Mamiellophyceae* estimée à partir du gène ribosomal 18S (droite).

2) Dynamique des Prasinovirus et Mamiellophyceae dans des environnements contrastés

a. Interaction hôte-virus et hypothèse de distribution

L'ensemble des Prasinovirus précédemment décrits dans l'environnement ont un mode de réplication lytique (Derelle *et al.*, 2008 ; Weynberg *et al.*, 2009, 2010 ; Mayer et Taylor, 1979 ; Moreau *et al.*, 2010). Une étude menée sur les virus d'*Ostreococcus tauri* (Clerissi *et al.*, 2012) suggère une forte spécificité de ces virus pour leur hôte. Cette même étude suggère qu'il en est de même pour les virus de *Bathycoccus*, qui semblent confinés à cette espèce d'hôte. L'étude menée en Manche sur les virus de *Micromonas* (Chapitre 3) suggère également une spécificité importante de ces virus, puisque 35 des 45 virus sélectionnés pour cette étude sont clade-spécifiques. Toutefois, cette spécificité n'est pas homogène entre les groupes viraux. En effet, seuls les virus infectant le clade C de *Micromonas* sont strictement inféodés à ce clade alors qu'environ 35% des virus infectant les clades A et B peuvent infecter ces deux clades.

Le suivi temporel en parallèle des différents clades de *Micromonas* et de leurs virus a révélé un lien entre les abondances virales et algales (Chapitre 3). Les abondances virales suivent les dynamiques algales avec un décalage de deux semaines. L'abondance virale montre une forte décroissance sur une période courte (4-5 semaines) après la fin de l'efflorescence algale. Ainsi, les conditions environnementales (biotiques et abiotiques) pourraient avoir un effet direct sur la destruction des populations virales. Par exemple, l'augmentation de la température serait associée à une augmentation de la destruction des virus (Danovaro et Middelboe, 2010), tout comme les fortes luminosités (UVB ; Noble et Fuhrman, 1997).

Ces dynamiques semblent récurrentes et ont également été observées pour les virus infectant *Ostreococcus tauri* (Bellec *et al.*, 2010b). Dans cette étude, Bellec *et al.* (2010b), mettent en évidence une structure forte dans la distribution spatiale des virus d'*Ostreococcus tauri*. Ils sont plus abondants dans les lagunes que dans les environnements hauturiers, ce qui coïncide avec la distribution spatiale de leur hôte (Bratbak *et al.*, 1990 ; Casterberg *et al.*, 2001 ; Danovaro *et al.*, 2011).

Ces différentes informations permettent de formuler, si la dynamique des hôtes est connue, plusieurs hypothèses sur la distribution spatio-temporelle des Prasinovirus. Et ainsi nous permettent de distinguer deux points :

- soit l'abondance de l'hôte implique l'abondance du virus et les communautés prasinovirus ne sont contrôlées que par les dynamiques des hôtes
- soit les facteurs environnementaux sont susceptibles d'affecter les communautés de Prasinovirus indépendamment de la dynamique de leurs hôtes

Nous avons confronté ces hypothèses aux données obtenues par notre échantillonnage mensuel d'un an, au sein d'environnements contrastés. Cela nous a permis d'étudier l'impact des conditions environnementales sur les communautés des virus de *Micromonas*, *Ostreococcus* et *Bathycoccus*.

b. Diversité et distribution saisonnières des communautés de Mamiellophyceae

i) Diversité des Mamiellophyceae

Au sein des différents environnements du Golfe du Lion, les genres *Ostreococcus*, *Micromonas* et *Bathycoccus* présentent les plus fortes diversités (Chapitres 4 et 5). À l'exception du clade B d'*Ostreococcus*, tous les genres de pico-Mamiellophyceae sont observés dans les stations lagunaires.

Lors d'un groupement à 99% pour la région V4, malgré un nombre élevé d'OTUs, le genre *Ostreococcus* est représenté dans 2 à 3 principaux OTUs suivant les stations. En revanche, *Bathycoccus* présente une abondance relative plus homogène concernant la répartition des séquences assignées entre les différents OTUs. Cette diversité estimée pour *Bathycoccus* dans l'ensemble des sites avec des abondances relatives différentes est similaire au résultat trouvé par Viprey *et al.* (2008) indiquant que ce genre contribuerait de façon homogène à la diversité des picoeucaryotes photosynthétiques en Mer Méditerranée.

La distribution spatio-temporelle des OTUs dans les différents environnements met en évidence un lien clair avec les conditions trophiques (Chapitre 5). En considérant la totalité des séquences eucaryotes, la zone marine s'avère plus diversifiée (2960 OTUs) que la zone lagunaire (1835 OTUs). Inversement, les Mamiellophyceae présentent une plus forte diversité moyenne sur l'année dans les lagunes (168 OTUs) qu'en mer ouverte (98 OTUs). Cette diversité lagunaire n'est cependant pas homogène au sein de la lagune. Le nombre d'OTU au site A est environ 2 fois plus important qu'au site B (respectivement 221 et 114 OTUs). Si *Micromonas* est présent dans l'ensemble des sites, *Ostreococcus* est limité aux sites lagunaires et *Bathycoccus* présente un profil particulier avec une présence plus marquée dans les sites marins, mais aussi au site B de la lagune de Leucate.

Les membres nano-eucaryotiques des Mamiellophyceae (*Crustomastix*, *Dolichomastix*, *Mantoniella* et *Monomastix*) présentent une faible diversité. La diversité des Mamiellophyceae que nous observons dans ces différents environnements coïncide avec les précédentes observations rapportées dans la littérature. En effet, l'ubiquité du genre *Micromonas*, qui semble lié à une radiation adaptative, a abouti à une diversité élevée au sein de ce groupe (Slapeta *et al.*, 2006), malgré une faible diversité au niveau de l'ADN ribosomique (Worden *et al.*, 2009). La prédominance des Mamiellales dans les environnements eutrophes (Viprey *et al.*, 2008 ; Wu *et al.*, 2013 ; Zhu *et al.*, 2005) par rapport aux milieux oligotrophes implique une plus forte diversité.

ii) Abondance des Mamiellophyceae

Les Mamiellophyceae présentent une distribution particulière avec une prédominance en milieu eutrophe (Chapitres 4-5). Les genres picoeucaryotiques, et plus particulièrement le genre *Micromonas*, constituent entre 20 et 80 % de la communauté dans les environnements étudiés (Chapitre 3 et 4). Les différentes espèces d'*Ostreococcus* contribuent fortement à l'abondance des Mamiellophyceae dans les lagunes avec des abondances 20 à 40 fois supérieures à celles des stations marines. Le genre *Bathycoccus*, caractérisé comme un genre plus abondant en zone oligotrophe (Viprey *et al.*, 2008 ; Wu *et al.*, 2014), a été retrouvé plus abondant au point B et qu'au

point A (Chapitre 5). Il est probable qu'en plus d'être isolé du reste de la lagune, le point B possède une population algale autochtone favorisée par des valeurs de salinité plus importantes. Les membres nanoeucaryotiques quant à eux ne représentent qu'une faible proportion de la communauté des Mamiellophyceae (4 à 7 %).

Suivant le gradient trophique et la saison, nous avons observé un changement au niveau des différentes communautés. La population d'*Ostreococcus* décroît avec la venue de conditions environnementales moins favorables. Cette diminution d'*Ostreococcus* est accompagnée d'une augmentation de la population de *Bathycoccus* (Chapitre 4 et 5). Ce changement des communautés est cohérent avec les différentes études portant sur la distribution de ces genres en lagunes et en Mer Méditerranée (Courties *et al.*, 1994 ; Viprey *et al.*, 2008 ; Zhu *et al.*, 2008 ; Wu *et al.*, 2014). *Micromonas* et *Bathycoccus* se développent principalement en hiver et au début du printemps, contrastant ainsi avec *Ostreococcus*.

Ces dynamiques présentent malgré tout de fortes variations inter-annuelles dans les lagunes, liées principalement aux conditions météorologiques variables (Hervé 1978). Même si *Ostreococcus tauri* reste présent tout au long de l'année, un accroissement particulièrement important a été observé entre juin et octobre 2013 à la station de Leucate A (proche du grau) mais extrêmement réduit à la deuxième station de Leucate (Chapitre 4). Ce changement particulier pourrait être dû à l'entrée d'eau de mer dans la lagune (Ladagnous et Le Bec, 1998). Cet apport permettrait ainsi l'arrivée de communautés marines au sein de la lagune eutrophe, ce qui confirmerait les résultats du séquençage environnemental direct précédemment décrit (Chapitre 5). Cette efflorescence n'a pas été observée au point B, suggérant que ce site semble relativement isolé du reste de la lagune d'un point de vue courantologique..

c. Diversité et variations saisonnières des communautés de Prasinovirus

i) Hypothèse sur la distribution des Prasinovirus

La prise en compte des dynamiques des Mamiellophyceae avec les interactions observées entre les Prasinovirus et leurs hôtes nous permettent de formuler une hypothèse en trois points quand aux évolutions spatio-temporelles de ces virus, si l'environnement ne joue pas de rôle prépondérant sur les populations virales, et suivant les différentes caractéristiques de la population d'hôtes.

(i) La première évolution est que l'abondance d'un hôte coïncide avec l'abondance de son virus spécifique. Nous pouvons alors nous attendre à une distribution hétérogène des communautés virales entre la zone lagunaire et zone de mer ouverte. La présence de *Micromonas* dans l'ensemble des environnements étudiés devrait se traduire par la présence homogène de leur virus dans l'ensemble des environnements. En revanche, *Ostreococcus* (abondant et limité aux zones lagunaires et côtières) et *Bathycoccus* (majoritairement présent en milieu oligotrophe) devraient entraîner la présence de leurs virus. Selon cette hypothèse, les tendances spatiales seraient associées aux tendances temporelles. *Ostreococcus tauri* montre une importante efflorescence pendant l'été (juin à octobre), donc leurs virus devraient présenter une forte abondance ainsi qu'une forte diversité durant cette période.

(ii) La deuxième évolution suggérée par les interactions entre les Mamiellophyceae et leurs virus concerne la diversification des communautés virales lors du développement des communautés algales. Compte-tenu de l'efflorescence de *Micromonas* au hiver-printemps, *Bathycoccus* au printemps et d'*Ostreococcus* en été, les communautés virales devraient théoriquement présenter une augmentation de diversité et d'abondance durant ces périodes.

(iii) La troisième évolution implique que la destruction des virions, suite à l'arrêt du développement des hôtes. Ainsi une diminution de l'abondance virale devrait être observée, du printemps à l'hiver pour les virus de *Micromonas* (MpV) et *Bathycoccus* (BpV) et de l'automne au printemps pour les virus d'*Ostreococcus* (OtV).

Ces différentes dynamiques sont schématisées sur la Figure 5.

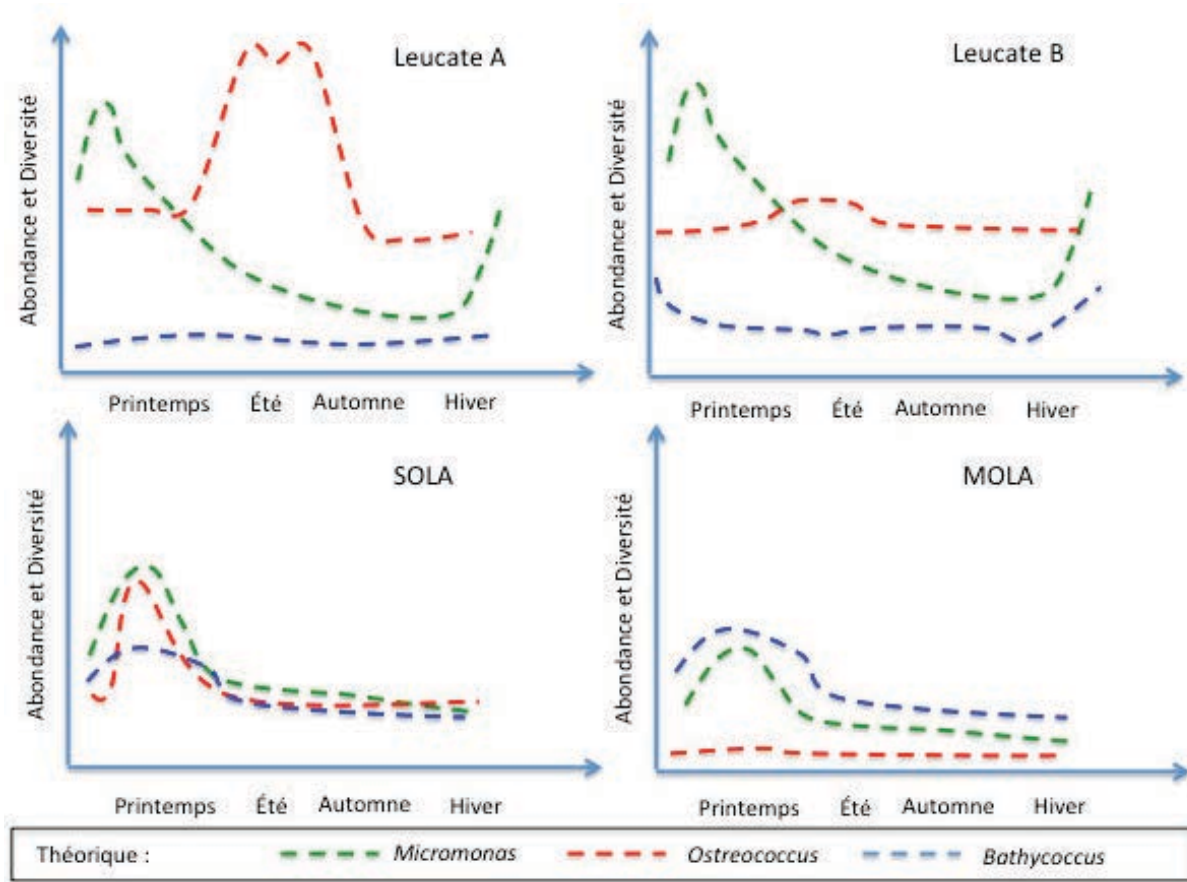


Figure 5 : dynamiques théoriques des groupes de Prasinovirus dans les environnements étudiés.

ii) Diversité

La majorité des études portant sur les Prasinovirus, et principalement sur les communautés environnementales, se sont basées jusqu'à présent sur des protocoles culture-dépendant pour l'étude de certains groupes viraux (Bellec *et al.*, 2009, Chapitre 3). Le protocole que nous avons mis en place ici nous a permis d'obtenir un nombre de séquences important (environ 18 millions de séquences par gène), et une description approfondie de l'ensemble des communautés de Prasinovirus.

Au sein des sites étudiés, les virus de *Micromonas* et *Ostreococcus tauri* sont les plus diversifiés (Chapitre 5), mais leur diversité n'est pas répartie de la même manière. Une majorité des OTUs assignés à MpV est partagée par tous les sites (43,4%) et représente en moyenne annuelle 25 à 34% des populations de Prasinovirus dans l'ensemble des

sites. Une augmentation de cette diversité est observée en début de printemps. Ces virus présentent la répartition la plus large au sein des environnements. Leurs hôtes sont adaptés à de nombreux environnements puisqu'ils sont identifiés dans les environnements eutrophes (lagune et Manche, Not *et al.*, 2004), en zone mésotrophique (SOLA) et oligotrophique (MOLA) (Chapitre 3 - 5). Ces picoeucaryotes présentent une importante diversité (Marin et Melkonian, 2010), et leurs virus présentent eux aussi la plus grande diversité au sein des Prasinovirus (Chapitre 5). Leur abondance ainsi que les mécanismes d'infection sont en relation étroite avec les dynamiques des communautés d'hôtes infectés (Chapitre 3). Les virus de *Micromonas* semblent pour la plupart spécifiques de clades d'hôtes.

En revanche, la diversité des virus assignés à OtV est principalement partagée entre les zones côtière et lagunaire. Dans ces environnements, elle représente une part homogène de la diversité (14 à 17%).

La diversité des BpV présente une particularité, en effet même si une majorité des OTUs est partagée par tous les sites échantillonnés (40%), la moitié des OTUs restants (30%) n'est partagée que par les stations marines et le point B des lagunes. Cette particularité suggère, comme pour les hôtes, que le site B possède une population particulière, limitée à cette zone (Chapitre 5), représentant en moyenne 3,5% du nombre d'OTU observés. En revanche, avec l'éloignement de la côte allant du site A de la lagune à SOLA puis à MOLA, une augmentation de la diversité de ce groupe viral est observée (indiquée par la proportion d'OTUs retrouvés : 1,8%, 6,3% et 10,6 %).

Lors de cette étude, 39,6% des OTUs formés n'ont pas pu être assignés aux principaux groupes de Prasinovirus connus (BpV, OtV, OIV, OmV, MpV) ou récemment découverts (Cruv1, MaV1 et 2, durant cette thèse). Ces OTUs représentent en moyenne une part constante de la population virale dans l'ensemble des sites (36 à 43%), même si ces séquences sont globalement plus représentées dans les points lagunaires et principalement en août et septembre où ils peuvent représenter plus de 80% de la diversité virale.

iii) Abondance relative

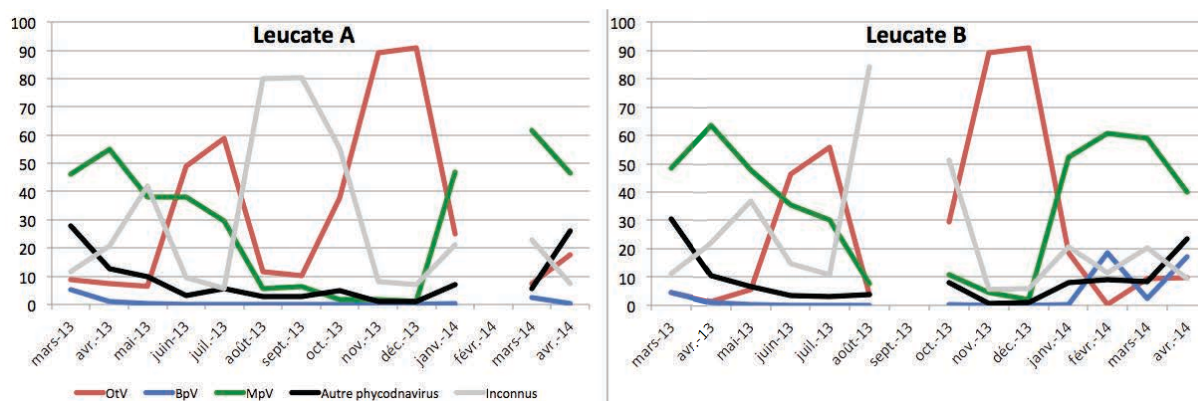


Figure 6: abondances relatives (%) des différents groupes de virus aux sites A et B de la lagune, basées sur la polymérase virale (résultats non présentés dans les différents articles de thèse, mais qui seront intégrés dans l'article correspondant au Chapitre 5)

Les zones eutrophes comme les sites lagunaires favorisent les efflorescences d'*Ostreococcus* (Chapitre 4-5) induisant ainsi la production de virus type OtV (Chapitre 5). Cette abondance d'OtV avait été observée précédemment, montrant que les virus infectieux d'*Ostreococcus* étaient en moyenne 100 fois plus abondants dans la lagune de Leucate qu'à la station côtière SOLA (Bellec *et al.*, 2010b).

Les abondances relatives des virus d'*Ostreococcus tauri* (Figure 6) dans les lagunes montrent une augmentation depuis le printemps jusqu'en décembre. Toutefois, une diminution importante du nombre relatif de séquences à 10% en août et septembre a été observée. Deux hypothèses peuvent expliquer cette diminution importante. La première serait le développement important de populations virales pouvant être amplifiées par les amorces utilisées. Ainsi, la proportion relative de la population de Prasinovirus connue serait réduite à des valeurs minimales du fait de la présence d'une autre population virale. La seconde hypothèse serait une diminution réelle de la population de de Prasinovirus par un facteur biotique ou abiotique. Trois arguments favorisent la seconde hypothèse :

(i) le nombre de séquences obtenues globalement plus bas durant cette période (Annexe 1 et 2 ; moyennes de séquences appariées et après nettoyage pour la polymérase virale : $116 \cdot 10^3$, moyenne sur cette période : site A = $67 \cdot 10^3$; site B : $69 \cdot 10^3$; moyenne de

séquences appariées et après nettoyage pour la MCP : $85 \cdot 10^3$, moyenne sur cette période : site A = $20 \cdot 10^3$; site B : $22 \cdot 10^3$;

(ii) la stabilité des autres groupes viraux. Les virus d'*Ostreococcus* montrent une diminution d'un facteur 4 (site A) et 17 (site B) entre les mois de juillet et août. Les autres groupes viraux incluant les virus *Ostreococcus lucimarinus*, *Bathycoccus* et les virus assignés au groupe des Phycodnaviridae en dehors de Prasinovirus, ne présentent pas de baisse aussi importante (facteur 1,6 pour les phycodnaviridae), montrant même une augmentation d'un facteur 2 à 7, respectivement pour BpV et OIV. Les virus de *Micromonas* montrent une diminution d'un facteur 4, mais cette diminution est graduelle jusqu'à décembre ;

(iii) ce phénomène de décroissance a déjà été rapporté dans cette lagune. En effet, Bellec et al (2010b), grâce à l'utilisation de cultures, ont montré sur des échantillons de 2007 une diminution d'un facteur 40 du nombre de virus infectant *Ostreococcus* entre juin et juillet ($20\,000$ pfu.ml à 520 pfu.mL). Cette diminution semble récurrente car visible l'année suivante. Enfin, elle a à nouveau été observée en 2011 (Clerissi *et al.*, 2014).

Ainsi cette baisse importante du nombre d'OtV ne serait pas liée à un phénomène de biais mais à un phénomène saisonnier et récurrent.

En Mer Méditerranée et plus précisément au site hauturier MOLA, les virus de *Bathycoccus* sont plus abondants, ce qui est probablement la conséquence de la présence de leurs hôtes (Chapitre 5). Ces hôtes sont principalement connus pour être présents en milieu oligotrophe (Wu *et al.*, 2014 ; Viprey et al, 2008) et sont donc supposés être distribués suivant un gradient trophique allant des abondances les plus faibles en zone eutrophe aux plus importantes en zone oligotrophe. Cela semble être le cas pour les stations de Leucate A, SOLA et MOLA. En revanche, une communauté particulière de BpV et de leurs hôtes co-existe au point B de la lagune de Leucate. Les abondances relatives de ces virus peuvent atteindre 10% au printemps. Cette observation implique que la communauté d'hôtes est la « clé » du développement des communautés virales, en induisant leur recrutement et leur maintien, puisque le caractère eutrophe et la dispersion par les courants ne peuvent pas expliquer à eux seuls la distribution des BpV dans ces milieux (Chapitre 5).

Malgré la répartition ubiquiste des MpV, l'abondance des OTUs formés pour ces virus montre des groupes inféodés au milieu lagunaire, d'autres à la station hauturière et enfin d'autres présents dans chaque environnement (Chapitre 5). Malgré cet aspect, l'abondance relative de ces virus dans les lagunes révèlent qu'ils sont majoritaires au printemps (50 à 60% des séquences obtenues). Les abondances relatives de ce groupe nous a révélé une augmentation de leur abondance durant l'hiver jusqu'au printemps.

Les communautés virales sont en lien avec les communautés d'hôtes (Bratbak *et al.*, 1990 ; Castberg *et al.*, 2001), qui induisent la sélection et la production des particules virales. Dans un premier temps, les conditions trophiques et les variations saisonnières des différents environnements sélectionnent les communautés d'hôtes.

d. Comparaison des hypothèses sur la distribution des Prasinovirus avec les observations spatio-temporelles

D'une manière globale, nos hypothèses sur les dynamiques virales, élaborées simplement à partir des dynamiques d'hôtes ainsi que les interactions hôtes-virus, sont supportées. Il est toutefois nécessaire de préciser que ces dynamiques théoriques ne sont basées que sur des hypothèses densité-dépendante en relation avec les hôtes. Des différences auraient pu être observées avec la prise en compte des *burst size*. En effet, les virus de *Micromonas* ont des *burst size* plus importantes mais aussi plus variables (entre 50 et 700) (Chapitre 3) en comparaison des virus d'*Ostreococcus* (25 ; Derelle *et al.*, 2008). Les virus de *Micromonas* et *Bathycoccus* ont des dynamiques explicables par la présence de leur hôte. Ainsi, pour ces groupes viraux les hypothèses suivant lesquelles (i) la présence de l'hôte implique celle du virus et (ii) l'activité de l'hôte influence majoritairement l'abondance virale, semble justes. Nous observons un couplage étroit dans les dynamiques hôte-virus qui se traduit par des cinétiques de décroissance différentes entre la Manche et le milieu lagunaire. Ainsi, en milieu lagunaire, les populations de *Micromonas* restent toujours présentes à des quantités importantes, induisant potentiellement une production virale et une disparition très progressive (5

mois) de MpV dans ce milieu. Inversement, la disparition est très rapide (1 mois) en Manche.

En revanche, pour les virus infectant *Ostreococcus tauri* des différences majeures sont visibles entre la distribution théorique attendue et les observations mesurées. Si globalement ces dynamiques de distribution concordent avec l'abondance de l'hôte dans les stations SOLA et MOLA, elles ne sont pas explicables par la seule présence de leur hôte pour les sites lagunaires. En effet, l'importante abondance d'*Ostreococcus tauri* en période estivale au site A (Chapitre 4) devrait théoriquement favoriser les communautés OtV durant cette période. Or durant les mois d'août et de septembre, ce groupe viral est très peu abondant (cette thèse, Bellec *et al.*, 2010b, Clerissi *et al.*, 2014). De plus, l'absence de dynamiques marquées pour *Ostreococcus tauri* au site B, ainsi que leur abondance restreinte comparée au site A, ne devraient pas favoriser de dynamique similaire entre les sites A et B.

Ces dynamiques particulières n'étant pas liées au nombre d'hôtes pour les virus OtV, il est possible que des facteurs environnementaux aient un impact significatif sur ce groupe. Le caractère récurrent de ce phénomène semble révéler un processus écologique potentiellement important dans cet écosystème.

La comparaison des résultats obtenus par l'acquisition de cette base de données environnementales décrivant les communautés de Prasinovirus nous permet d'identifier des périodes charnières, liées à l'impact des facteurs environnementaux.

e. Vers l'identification d'un contrôle environnemental

Nous avons observé dans ce travail que l'environnement semble impacter indirectement l'abondance et la diversité virale par le contrôle des communautés d'hôtes. Les efflorescences phytoplanctoniques induisent un recrutement et une production des virions (Chapitres 3 et 5). La diversité plus importante des Mamiellophyceae au site lagunaire proche du grau suggérerait une influence importante de l'apport d'eau marine à ce point. En effet, la diversification des communautés autochtones lagunaires pourrait

être due à l'apport de communautés marines lors de l'entrée d'eau dans la lagune. Ces communautés d'hôtes sont contrôlées par la concentration en nutriment (Fuller *et al.*, 2006 ; Mackey *et al.*, 2009 ; Kirkham *et al.*, 2011), imposant une distribution à chaque groupe phytoplanctonique en terme de diversité ou d'abondance. Un effet direct de ces facteurs environnementaux sur les abondances virales a toutefois été suggéré concernant les concentrations en phosphate. La présence de gène codant pour des transporteurs de phosphate dans le génome de prasinovirus a été révélée (Monier *et al.*, 2012). Des tests de lyse en condition ont montré un effet limitant des faibles concentrations en phosphate sur la lyse virale (Bratbak *et al.*, 1993). Cet effet pourrait être dû au besoin du virus d'une quantité importante de phosphate pour se répliquer. En effet, les virus possèdent un ratio acide nucléique/protéine élevé (Wilson *et al.*, 1996 ; Monier *et al.*, 2012,) impliquant un coût en phosphate important. Les conditions trophiques pourraient ainsi déterminer des régions plus productives en virus comme les lagunes eutrophes. Ainsi ces régions pourraient être des « chemostats naturels » pour certains virions durant les périodes les plus actives de l'année, lors des efflorescences phytoplanctoniques. C'est à partir de ces zones que les particules virales seraient dispersées par les courants dans les environnements alentours. Par la suite, les particules virales se comporteraient comme une « banque de graines », qui sont sélectionnées et produites, si et seulement si des hôtes sensibles sont présents. Dans le cas de l'absence de l'hôte ou présent en quantité limitée, les virus seraient dégradés, mais resteraient tout de même actifs en quantités minimale attendant le retour d'activité de leurs hôtes.

Dans notre étude le phosphate, ainsi que les nutriments en général, n'atteignent jamais de si faibles concentrations en zone lagunaire. Lors de la diminution drastique des communautés d'OtV, les concentrations de phosphate présentaient même les plus fortes valeurs enregistrées (0,62 et 0,34 $\mu\text{mole.L}$). Nous écartons donc l'hypothèse du phosphate comme facteur déclenchant de la diminution d'abondance relative d'OtV en été. Un déséquilibre du rapport N/P pourrait être à l'origine de cette diminution. Le rapport de Redfield (Redfield *et al.*, 1963) indique une limitation en phosphore lorsque le rapport N/P est supérieur à 10 et montre une limitation en azote si il est inférieur à 5. Dans l'environnement lagunaire, ce rapport est continuellement supérieur à 9, indiquant une possible limitation en phosphate. Il est toutefois nécessaire de préciser qu'il est difficile d'attribuer un effet à ce déséquilibre, puisqu'il ne possède pas de réel sens dans

un contexte d'eutrophisation, où les sels nutritifs sont continuellement en excès par rapport aux besoins de la croissance des algues.

Nous avons en revanche pu inférer deux impacts directs des facteurs environnementaux sur ces communautés de Prasinovirus.

Le premier est lié au vent, car durant les mois d'été où le vent d'autan (de la mer vers la lagune) est actif, une entrée d'eau est favorisée (rapport IFREMER, 1998). Durant cette période, la structure des communautés à Leucate A n'est pas liée à la dynamique saisonnière des facteurs environnementaux mesurés contrairement au reste de l'année. Ce fait est probablement dû à l'apport de virus marins à l'intérieur de la lagune. Le vent semble donc avoir un rôle dans la création de courants favorisant la dispersion des particules virales marines vers la lagune, pouvant par la suite être sélectionnées en fonction de la population d'hôtes du site d'arrivée. Si les virus marins (inconnus) sont capables d'infecter *Ostreococcus*, cela expliquerait l'augmentation d'abondance relative de clades inconnus en août et septembre. Mais cette hypothèse implique une homogénéisation complète de la lagune, ce qui ne semble pas le cas

Le deuxième aspect environnemental semblant impacter directement les populations virales est lié à la température et/ou à la salinité..

Suite à l'augmentation de l'abondance virale dans la lagune au printemps, nous avons observé une diminution importante des populations des virus d'*Ostreococcus* (août et septembre ; Chapitre 5). La baisse des OtV n'est pas liée à la disparition de leurs hôtes puisque nous les avons retrouvés en quantité abondante à la même période (Chapitre 3). L'influence de la température et/ou de la salinité pourrait expliquer ce processus, puisqu'elles atteignent des valeurs maximales durant cette période.

Ces deux descripteurs ont déjà été rapportés comme impactant les communautés virales. L'augmentation de la température est associée à une augmentation exponentielle de la dégradation virale, soit directement par déstabilisation de la capsid virale, soit par la stimulation des activités exoenzymatiques présentes dans l'environnement (Danoravo, 2010 ; Corinaldesi *et al.*, 2010). L'étude des communautés virales en prenant en compte la salinité à dans un premier temps révélée une plus grande abondance virale dans les milieux à faible salinité (Danoravo *et al.*, 2011). Cette salinité plus importante durant ces mois pourrait influencer le pH et les conditions

d'absorptions des virions par modification des forces electro-statiques (Harvey et Ryan, 2004). Ces propriétés d'absorption se renforcent généralement avec une augmentation du pH (Charle, 1984 ; Grant *et al.*, 1993), Il vrai qu'une légère augmentation du pH à été observée lors de notre suivi mais ces augmentions sont moins importantes que lors des periodes hivernales, ou il atteint ces valeurs maximales.

Ainsi les conditions environnementales pourraient contraindre l'abondance virale à travers l'augmentation ou la diminution des taux de dégradation permettant d'établir un équilibre entre les communautés virales et les communautés d'hôtes.

Ces impacts directs de l'environnement sont hypothétiques, est la relation avec les Prasinovirus n'a pu être montré que sur la base des descripteurs environnemntaux mesurés. L'utilisation de la méthode STATICO (Chapitre 5) nous a permis une visualisation en parallèle de nos jeux de données regroupant les descripteurs environnementaux ainsi que les informations sur les Mamiellophyceae et les Prasinovirus. Ainsi nous avons pu observer la co-structure entre ces jeux de données (Chapitre 5, Figures 7 et 8). Ces analyses montrent des différences entre ces co-structures, suggérant que les descripteurs environnementaux mesurés ne sont pas suffisants pour expliquer la variation des communautés virales et algales. La variabilité du point A pour les deux communautés indique qu'un évènement important modifie la structure de la communauté, suggérant ainsi l'impact de l'échange d'eau entre la mer et la lagune. D'autre part, la structure de la communauté virale durant les mois d'août et de septembre est différente de la structure des descripteurs environnementaux. Cela suggère que malgré l'impact probable de la température sur la destruction de la communauté d'OtV, ce n'est pas suffisant pour expliquer entièrement ce phénomène. Anisi il serait intéressant de mesurer plus de facteurs environnementaux comme la production primaire, la présence de métaux comme le fer, mais aussi la présence de nanoflagellés qui ont déjà été rapportés comme étant responsable de la destruction virale par phagocytose (Suttle et chen, 1992 ; Gonzales et Suttle, 1993 ; Manage *et al.*, 2002 ; Bettarel *et al.*, 2005). En revanche, cette activité de prédation ne s'appliquerait pas aux hôtes puisqu'ils restent à des concentrations élevées durant cette période (Chapitre 4). Un ciblage spécifique des virus d'*Ostreococcus* pourrait permettre une simplification de l'analyse et ainsi d'identifier clairement l'impact des facteurs

environnementaux sur ce groupe, notamment par des études en laboratoire permettant de faire varier ces paramètres.

3) Limite méthodologique pour l'étude des communautés hôtes-virus

- a. Utilisation de la qPCR et du séquençage haut débit pour l'étude des communautés eucaryotiques

Lors de ce travail de thèse, nous avons étudié deux régions de l'ADNr 18S (V4 et V9, Stoeck *et al.*, 2010). Ces régions, bien que particulièrement informatives, ne permettent pas d'atteindre la même résolution taxonomique. En effet, la région V9 est plus petite et contient moins de variabilité que la région V4. Malgré cette différence, l'utilisation de ces deux régions pour l'étude des communautés eucaryotiques nous a permis de vérifier la similarité des analyses par rapport à ces deux régions et ainsi de contrôler les biais éventuels. Suite au séquençage, une grande quantité de séquences a été obtenue (21 à 24 millions).

Bathycoccus, *Micromonas* et *Ostreococcus* ne possédant que quelques copies des gènes ribosomiques (2-4 copies, Derelle *et al.*, 2006 ; Zhu *et al.*, 2005), l'utilisation d'amorces universelles pour le 18S pourrait favoriser l'amplification d'autres groupes taxonomiques qui possèdent plus de copies (Zhu *et al.*, 2005). Malgré un tri cellulaire diminuant la fraction n'ayant pas d'intérêt (Shi *et al.*, 2009), une grande part des séquences récupérées est assignée à des groupes différents des *Mamiellophyceae*, principalement des Dinoflagellés ou des *Trebouxiophyceae*. Ce phénomène n'est pas lié à leur abondance, puisque que la qPCR effectuée sur ces échantillons a révélé que les *Mamiellophyceae* sont dominantes dans ces environnements. Ce résultat supporte que le nombre de copies de 18S retrouvées chez ces organismes, 100 à 1000 fois plus important que chez les Mamiellales (Zhu *et al.* 2005), est à l'origine de leur dominance dans les séquences.

L'inconvénient principal à l'utilisation du séquençage haut débit est qu'il est délicat d'utiliser les données obtenues d'un point de vue quantitatif. En effet, la technique

lorsqu'elle est réalisée sur plusieurs échantillons simultanément, nécessite de placer les produits amplifiés par PCR suivant des proportions stoechiométriques équivalentes. Pour chaque banque de séquences, la normalisation *a posteriori* du nombre de séquences pour chaque OTU par le nombre de séquence total obtenu par banque, est essentiel afin de comparer les résultats obtenus entre les différentes banques. Ainsi le séquençage haut débit ne permet d'obtenir que des données semi-quantitatives.

Le nombre d'OTU obtenus lors du séquençage n'est pas obligatoirement représentatif de la diversité réelle, en effet suivant le seuil de groupement choisi le nombre d'OTUs retrouvés peut fortement varier (e.g. pour la région V4, le genre *Ostreococcus* : 442 OTUs à 99%, contre 5 OTUs à 98%). Malgré cette différence, les courbes de rarefaction réalisées pour ces seuils de regroupement suggèrent que le plateau est atteint et donc que l'effort d'échantillonnage a été suffisant pour capter l'ensemble des communautés eucaryotiques photosynthétiques. L'augmentation artificielle du nombre d'OTU est probablement due aux erreurs produites lors de l'amplification de la région cible ou lors du séquençage (Schloss *et al.*, 2009). En revanche, la formation OTU avec un seuil similaire de l'ensemble de la base de séquence obtenue, possédant la même profondeur de séquençage, autorise la comparaison de la diversité entre les différents échantillons.

Afin de palier ce défaut, nous avons mis en place un protocole de qPCR permettant de suivre les dynamiques du nombre de copies de la région V4. La mise au point de cette technique sur 12 représentants des Mamiellophyceae (incluant les différents clades) nous a permis de décrire de manière quantitative les variations d'abondances de ces communautés. Ces deux approches complémentaires ont abouti à une description à haute résolution de la dynamique des communautés de Mamiellophyceae dans différents environnements du Golfe du Lion.

b. L'étude des communautés de Prasinovirus par séquençage haut débit

Durant ce travail de thèse, notre approche a consisté à étudier la distribution et la diversité des virions à l'état libre dans l'environnement grâce à l'amplification et au séquençage d'ADN extrait à partir d'échantillons floculés au FeCl₃. Cependant, cette technique induit quelques biais qu'il est nécessaire d'identifier.

Le protocole basé sur le chlorure de fer (John *et al.*, 2010) permet de récupérer l'ensemble des particules libres possédant une charge ionique positive, principalement les entités virales. Cependant, toutes les particules virales sont collectées indépendamment de leur état de dégradation ou de leur pouvoir infectieux, et il nous est difficile d'estimer la part de virus actifs dans nos échantillons. Derelle *et al.*, (2008) ont montré que pour OtV5 seul 35 % d'entre eux pouvaient s'adsorber sur les hôtes, mais ces capacités d'adsorption ne tiennent pas compte des mécanismes de résistance mis en place par les hôtes face à ces virus (Thomas *et al.*, 2011). De ce fait, la présence d'une séquence virale dans les amplicons n'implique pas obligatoirement que le virus soit actif.

De plus, la préservation ou la dégradation des virions face aux conditions biotiques et abiotiques influence la qualité et la quantité d'ADN viral extrait. L'augmentation de l'activité exoenzymatique serait largement responsable de la dégradation de la capsid virale lors de l'accroissement de la température (Corinaldesi *et al.*, 2010). Cette dégradation pourrait permettre la libération de l'ADN viral dans l'environnement, expliquant ainsi que 17 à 30 % de l'ADN libre dans les environnements aquatiques serait d'origine virale (Nielsen *et al.*, 2007). Cependant cet ADN libre est rapidement hydrolysé par des nucléases extra et intra-cellulaires (Paul *et al.*, 1987). Dans des échantillons estuariens, le turn-over de ces molécules a été rapporté comme court (6,5 h, Paul *et al.* (1987), et England *et al.* (2005) indiquent qu'une quantité d'ADN viral purifiée et hautement concentrée, ajoutée dans un mésocosme en milieu naturel, n'était plus détectable après 1 semaine.

Nous pouvons donc supposer que malgré cette incertitude, si de l'ADN libre est amplifié par nos amorces, il doit provenir de virus dégradés récemment. L'utilisation lors de la mise en place des techniques de séquençage à haut débit d'amorces amplifiant spécifiquement les communautés Prasinovirus, et le fait que ces gènes étudiés ne sont

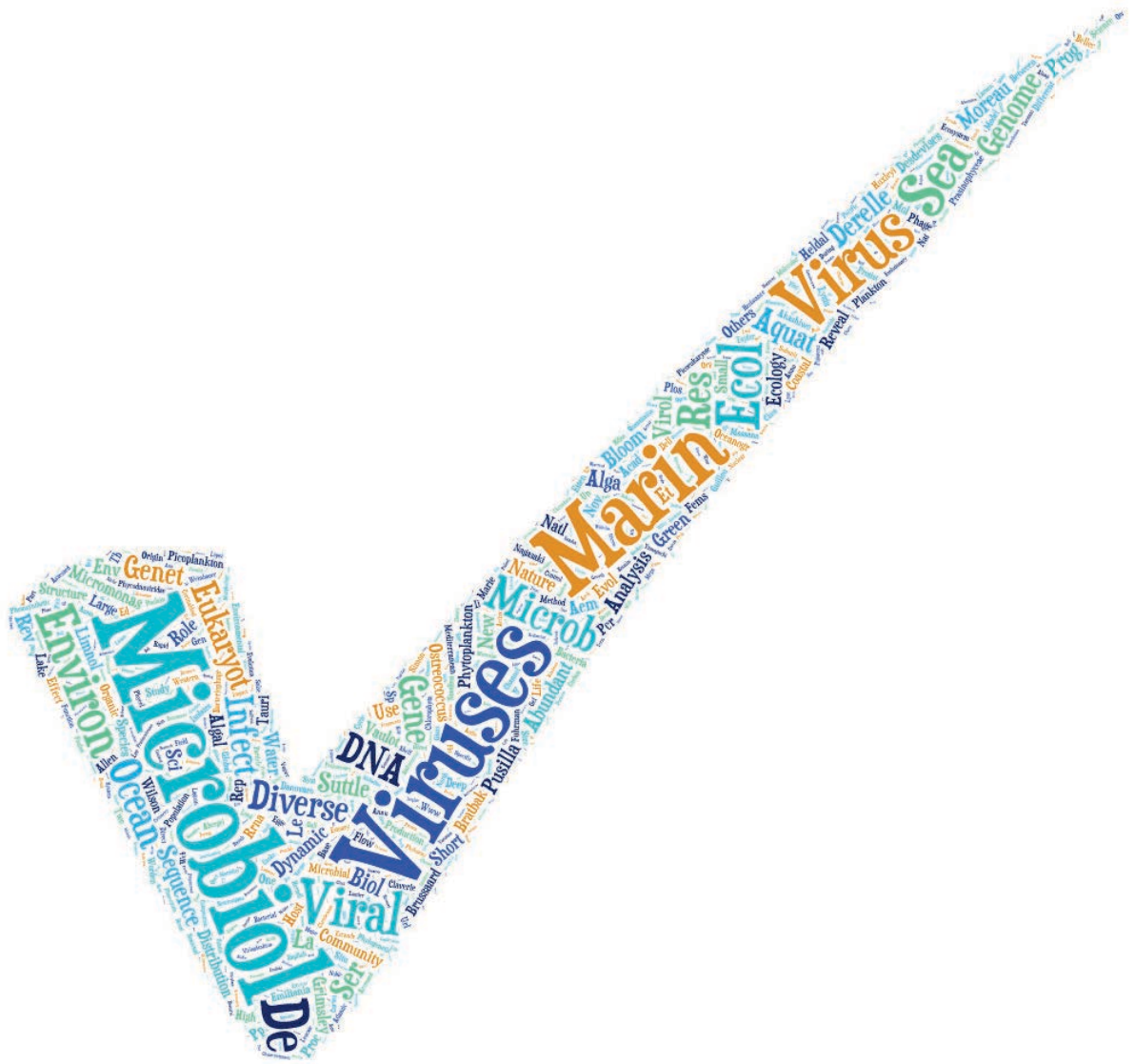
présents qu'à une seule copie (Derelle *et al.*, 2008), nous permet de réduire l'aspect semi-quantitatif de la méthode, mais sans toutefois la supprimer. Enfin, l'étude de la décroissance des populations virales de Prasinovirus en mésocosmes implanté en milieu naturel pourrait nous permettre de mieux comprendre le comportement des virions libres face aux variations environnementales.

II Conclusion et perspectives

Les principaux résultats de ce travail de thèse suggèrent que l'évolution ainsi que la stratégie de réplication des prasinovirus est liés avec l'évolution et les caractéristiques de leurs hôtes. Les relations entre les patrons phylogénétiques des virus et des hôtes pourraient résulter d'une co-spéciation entre ces entités que nous pourrions étudier plus en détail grâce aux nouvelles souches virales isolées. Par l'utilisation de techniques de génomique environnementale au sein de 3 sites contrastés, nous avons mis en évidence que les virus d'*Ostreococcus* et de *Micromonas* sont les Prasinovirus les plus diversifiés. Les environnements eutrophes, à forte productivité, influencent de façon indirecte l'abondance des virus par la sélection des communautés d'hôtes. Les conditions environnementales permettent la dispersion des particules virales mais contraignent dans un même temps leur abondance, en influençant leur condition de préservation. Dans le cadre d'études environnementales, la mise en place d'une détection *in silico* assignant des séquences virales à des séquences hôtes pourrait permettre d'identifier de nouveaux groupes viraux et ainsi permettre de s'affranchir, dans une certaine limite, des techniques d'isolement parfois difficiles à mettre en place.

Les génomes complets des nouveaux prasinovirus infectant les microalgues *Crustomastix* et *Mamiella* permettront d'obtenir non seulement la séquence complète de l'ADN polymérase B pour affiner la phylogénie réalisée durant cette thèse et même d'étudier de façon plus approfondie leur évolution à travers des approches phylogénomiques, mais également d'approfondir l'étude comparative des Prasinovirus.

L'acquisition de la base de données regroupant les abondances relatives des communautés d'hôtes et de virus ainsi que 31 descripteurs environnementaux permet actuellement le développement d'une méthode innovante afin d'identifier les couples hôte-virus à partir de séquences obtenues par séquençage massif dans l'environnement. **Des détails sur le développement et le principe de cette méthode sont disponibles Annexe 5.**



Références

Références

A

- Abedon, S.T., Herschler, T.D., Stopar, D., 2001. Bacteriophage latent-period evolution as a response to resource availability. *Appl Env. Microbiol.* 67, 4233–4241.
- Ackermann, H.-W., 2001. Frequency of morphological phage descriptions in the year 2000. *Arch. Virol.* 146, 843–857.
- Adolph, K.W., Haselkorn, R., 1972. Photosynthesis and the development of blue-green algal virus N-1. *Virology* 47, 370–374. doi:10.1016/0042-6822(72)90272-3
- Amaral-Zettler, L.A., McCliment, E.A., Ducklow, H.W., Huse, S.M., 2009. A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA genes. *PLoS One* 4, e6372.
- Anonyme, 1984. Etang côtiers du languedoc-Roussillon - Vol. 1 : Situation actuelle - C.E.E. - programme méditerranéen intégré.
- 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). *Revue des Travaux de l'Institut des Pêches Maritimes* 33, 335–443.
- Auguet, J.C., Montanié, H., Lebaron, P., 2006. Structure of virioplankton in the Charente Estuary (France): transmission electron microscopy versus pulsed field gel electrophoresis. *Microb. Ecol.* 51, 197–208. doi:10.1007/s00248-005-0043-0
- Auguet, J.C., Montanié, H., Delmas, D., Hartmann, H.J., Huet, V., 2005. Dynamic of virioplankton abundance and its environmental control in the Charente estuary (France). *Microb. Ecol.* 50, 337–349.
- Azam, F., 1983. The ecological role of water-column microbes in the sea. *Mar. Ecol. Prog. Ser.* 10, 257–263. doi:10.3354/meps010257

B

- Baldauf, S.L., 2008. An overview of the phylogeny and diversity of eukaryotes. *J. Syst. Evol.* 46, 263–273.
- Balzano, S., Marie, D., Gourvil, P., Vaultot, D., 2012. Composition of the summer photosynthetic pico and nanoplankton communities in the Beaufort Sea assessed by T-RFLP and sequences of the 18S rRNA gene from flow cytometry sorted samples. *ISME J.*
- Baudoux, A.-C., Brussaard, C.P., 2008. Influence of irradiance on virus–algal host interactions¹. *J. Phycol.* 44, 902–908.
- Baudoux, A.C., Brussaard, C.P.D., 2005. Characterization of different viruses infecting the marine harmful algal bloom species *Phaeocystis globosa*. *Virology* 341, 80–90.
- Baudoux, A.-C., Lebretonchel, H., Dehmer, H., Latimier, M., Edern, R., Rigaut-Jalabert, F., Ge, P., Guillou, L., Foulon, E., Bozec, Y., others, 2015. Interplay between the genetic clades of *Micromonas* and their viruses in the Western English Channel. *Environ. Microbiol. Rep.* 7, 765–773.
- Bec, B., Collos, Y., Souchu, P., Vaquer, A., Lautier, J., Fiandrino, A., Benau, L., Orsoni, V., Laugier, T., 2011. Distribution of picophytoplankton and nanophytoplankton along an anthropogenic eutrophication gradient in French Mediterranean coastal lagoons. *Aquat. Microb. Ecol.* 63, 29–45. doi:10.3354/ame01480
- Bellec, L., Desdevises, Y., 2015. Quand virus et hôtes évoluent ensemble: la fidélité est-elle la règle? *Virologie* 19, 140–148.
- Bellec, L., Grimsley, N., Desdevises, Y., 2009. Phylogenetic analysis of new Prasinoviruses (Phycodnaviridae) that infect the green unicellular algae *Ostreococcus*, *Bathycoccus* and *Micromonas*. *Env. Microbiol Rep.* 1, 114–123.
- 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., Derelle, E., Moreau, H., Desdevises, Y., 2010b. Abundance, spatial distribution and genetic diversity of *Ostreococcus tauri* viruses in two different environments. *Environ. Microbiol. Rep.* 2, 313–321. doi:10.1111/j.1758-2229.2010.00138.x
- 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.
- Berdjeb, L., Jacquet, S., 2009. La viriosphère: quelle place dans le fonctionnement et l'évolution des écosystèmes aquatiques. *Virologie* 13, 189–199.
- Bergh, O., Børsheim, K.Y., Bratbak, G., Heldal, M., 1989. High abundance of viruses found in aquatic environments. *Nature* 340, 467–468. doi:10.1038/340467a0
- Bettarel, Y., Sime-Ngando, T., Bouvy, M., Arfi, R., Amblard, C., 2005. Low consumption of virus-sized particles by heterotrophic nanoflagellates in two lakes of the French Massif Central. *Aquatic Microbial Ecology* 39, 205–209.
- Béthoux, J.P., Morin, P., Ruiz-Pino, D.P., 2002. Temporal trends in nutrient ratios: chemical evidence of Mediterranean ecosystem changes driven by human activity. *Deep Sea Research Part II: Topical Studies in Oceanography* 49, 2007–2016.
- Biegala, I.C., Not, F., Vaultot, D., Simon, N., 2003. Quantitative Assessment of Picoeukaryotes in the Natural Environment by Using Taxon-Specific Oligonucleotide Probes in Association with Tyramide Signal Amplification-Fluorescence In Situ Hybridization and Flow Cytometry. *Appl. Environ. Microbiol.* 69, 5519–5529. doi:10.1128/AEM.69.9.5519-5529.2003
- Billler, S. J., F. Schubotz, S. E. Roggensack, A. W. Thompson, R., E. Summons, and S. W. Chisholm, 2014. Bacterial Vesicles in Marine Ecosystems. *Science* 343, no. 6167: 183-186.

- 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. *Marine Ecology-Progress Series* 97, 1.
- Boucher, N., Vaultot, D., Partensky, F., 1991. Flow cytometric determination of phytoplankton DNA in cultures and oceanic populations. *Mar. Ecol. Prog. Ser.* 71, 75–84.
- Bouin, M.-N., Caniaux, G., Traullé, O., Legain, D., Le Moigne, P., 2012a. Long-term heat exchanges over a Mediterranean lagoon. *J. Geophys. Res. Atmospheres* 117, D23104. doi:10.1029/2012JD017857
- Boujelben, I., Yarza, P., Almansa, C., Villamor, J., Maalej, S., Antón, J., Santos, F., 2012. Virioplankton community structure in Tunisian solar salterns. *Applied and environmental microbiology* 78, 7429–7437.
- Boutiere, H., 1980. Introduction à la connaissance des milieux lagunaires. *Océanis* 5, 823–832.
- Bratbak, G., Heldal, M., Thingstad, T.F., Tuomi, P., 1996. Dynamics of virus abundance in coastal seawater. *FEMS Microbiology Ecology* 19, 263–269.
- Bratbak, G., Egge, J.K., Heldal, M., 1993a. Viral mortality of the marine alga *Emiliana huxleyi* (Haptophyceae) and termination of algal blooms. *Mar. Ecol. Prog. Ser.* 93.
- Bratbak, G., Heldal, M., Norland, S., Thingstad, T.F., 1990. Viruses as partners in spring bloom microbial trophodynamics. *Appl. Environ. Microbiol.* 56, 1400–1405.
- Breitbart, M., Salamon, P., Andresen, B., Mahaffy, J.M., Segall, A.M., Mead, D., Azam, F., Rohwer, F., 2002. Genomic analysis of uncultured marine viral communities. *Proc. Natl. Acad. Sci.* 99, 14250–14255.
- Brittain, S.M., Wang, J., Babcock-Jackson, L., Carmichael, W.W., Rinehart, K.L., Culver, D.A., 2000. Isolation and Characterization of Microcystins, Cyclic Heptapeptide Hepatotoxins from a Lake Erie Strain of *Microcystis aeruginosa*. *J. Gt. Lakes Res.* 26, 241–249. doi:10.1016/S0380-1330(00)70690-3

- Brown, C.M., Campbell, D.A., Lawrence, J.E., 2007. Resource dynamics during infection of *Micromonas pusilla* by virus MpV-Sp1. *Environmental microbiology* 9, 2720–2727.
- Brussaard, C.P., 2005. Viral Control of Phytoplankton Populations-a Review 1. *J. Eukaryot. Microbiol.* 51, 125–138.
- Brussaard, C.P.D., 2004. Optimization of procedures for counting viruses by flow cytometry. *Appl. Env. Microbiol* 70, 1506–1513.
- Brussaard, C.P.D., Kempers, R.S., Kop, A.J., Riegman, R., Heldal, M., 1996. Virus-like particles in a summer bloom of *Emiliana huxleyi* in the North Sea. SCSIO OpenIR 10.
- Brussaard, C.P.D., Riegman, R., Noordeloos, A.A.M., Cadée, G.C., Witte, H., Kop, A.J., Nieuwland, G., Van Duyl, F.C., Bak, R.P.M., 1995. Effects of grazing, sedimentation and phytoplankton cell lysis on the structure of a coastal pelagic food web. *Mar. Ecol. Prog. Ser. Oldendorf* 123, 259–271.
- Brussaard, C.P.D., Thyrraug, R., Marie, D., Bratbak, G., 1999. Flow Cytometric Analyses of Viral Infection in Two Marine Phytoplankton Species, *Micromonas pusilla* (Prasinophyceae) and *Phaeocystis pouchetii* (Prymnesiophyceae). *J. Phycol.* 35, 941–948. doi:10.1046/j.1529-8817.1999.3550941.x
- Brussaard, C.P., Martinez, J.M., 2008. Algal bloom viruses. *Plant Viruses* 2, 1–13.
- Bucher, R.W., 1952. Contribution to our knowledge of the smaller marine algae.

C

- Campbell, M.S., Wright, A.C., 2003. Real-Time PCR Analysis of *Vibrio vulnificus* from Oysters. *Appl. Environ. Microbiol.* 69, 7137–7144. doi:10.1128/AEM.69.12.7137-7144.2003
- Cardol, P., Bailleul, B., Rappaport, F., Derelle, E., Béal, D., Breyton, C., Bailey, S., Wollman, F.A., Grossman, A., Moreau, H., others, 2008. An original adaptation of photosynthesis in the marine green alga *Ostreococcus*. *Proc. Natl. Acad. Sci.* 105, 7881–7886.

- Castberg, T., Larsen, A., Sandaa, R.A., Brussaard, C.P.D., Egge, J.K., Heldal, M., Thyrhaug, R., Van Hannen, E.J., Bratbak, G., 2001. Microbial population dynamics and diversity during a bloom of the marine coccolithophorid *Emiliana huxleyi* (Haptophyta). *Mar. Ecol. Prog. Ser.* 221, 39–46.
- Cavalier-Smith, T., 1985. Selfish DNA and the origin of introns. *Nature* 315, 283–284.
- Cépralmar, 2000. Bilan écologique et diagnostic de l'étang de Salse-Leucate.
- Cesmat, L., 2006. Etude des processus hydrodynamiques et écophysologiques de la dynamique d'une algue invasive: *Valonia aegagropila* (C. Agardh) dans la lagune de Salses-Leucate. Université Montpellier II.
- Charles, P., 1984. Applied and theoretical aspects of virus adsorption to surfaces. *Adv. Appl. Microbiol.* 30, 133.
- Chen, F., Suttle, C.A., 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.
- Chen, F., Suttle, C.A., 1995. Amplification of DNA polymerase gene fragments from viruses infecting microalgae. *Appl. Environ. Microbiol.* 61, 1274–1278.
- Chen, F., Suttle, C.A., Short, S.M., 1996. Genetic diversity in marine algal virus communities as revealed by sequence analysis of DNA polymerase genes. *Appl. Environ. Microbiol.* 62, 2869–2874.
- Chessel, D., Dufour, A.B., Thioulouse, J., 2004. The ade4 package-I-One-table methods. *R News* 4, 5–10.
- Chrétiennot-Dinet, M.-J., Courties, C., Vaquer, A., Neveux, J., Claustre, H., Lautier, J., Machado, M.C., 1995. A new marine picoeucaryote: *Ostreococcus tauri* gen. et sp. nov. (Chlorophyta, Prasinophyceae). *Phycologia* 34, 285–292. doi:10.2216/i0031-8884-34-4-285.1
- Clarke, K.R., 1993. Non-parametric multivariate analyses of changes in community structure. *Aust. J. Ecol.* 18, 117–143. doi:10.1111/j.1442-9993.1993.tb00438.x

- Clasen, J.L., Brigden, S.M., Payet, J.P., Suttle, C.A., 2008. Evidence that viral abundance across oceans and lakes is driven by different biological factors. *Freshw. Biol.* 53, 1090–1100.
- Clasen, J.L., Suttle, C.A., 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.
- Claverie, J.-M., Abergel, C., 2014. Pithovirus: un virus géant venu du fond des âges. *Biofutur* 352, 50–51.
- Claverie, J.-M., Abergel, C., 2010. Mimivirus: the emerging paradox of quasi-autonomous viruses. *Trends Genet.* 26, 431–437. doi:10.1016/j.tig.2010.07.003
- Claverie, J.-M., Abergel, C., 2009. Mimivirus and its virophage. *Annu. Rev. Genet.* 43, 49–66.
- 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. *Applied and environmental microbiology* 80, 3150–3160.
- Clerissi, C., Grimsley, N., Subirana, L., Maria, E., Oriol, L., Ogata, H., Moreau, H., Desdevises, Y., 2014. Prasinovirus distribution in the Northwest Mediterranean Sea is affected by the environment and particularly by phosphate availability. *Virology* 466, 146–157.
- Clerissi, C., Grimsley, N., Desdevises, Y., 2013. Genetic exchanges of inteins between prasinoviruses (Phycodnaviridae). *Evolution* 67, 18–33.
- Clerissi, C., Desdevises, Y., Grimsley, N., 2012. Prasinoviruses of the Marine Green Alga *Ostreococcus tauri* Are Mainly Species Specific. *J. Virol.* 86, 4611–4619. doi:10.1128/JVI.07221-11
- Clerissi, C., Desdevises, Y., Romac, S., Audic, S., Vargas, C., Acinas, S.G., Casotti, R., Poulain, J., Wincker, P., Hingamp, P., others, 2015. Deep sequencing of amplified Prasinovirus and host green algal genes from an Indian Ocean transect reveals interacting trophic dependencies and new genotypes. *Environ. Microbiol. Rep.*

- Clokier, M.R., Millard, A.D., Letarov, A.V., Heaphy, S., 2011. Phages in nature. *Bacteriophage* 1, 31–45.
- Cochlan, W.P., Wikner, J., Steward, G.F., Smith, D.C., Azam, F., 1993. Spatial distribution of viruses, bacteria and chlorophyll a in neritic, oceanic and estuarine environments. *Mar. Ecol.-Prog. Ser.* 92, 77–77.
- Corinaldesi, C., Dell'Anno, A., Danovaro, R., 2007a. Viral infection plays a key role in extracellular DNA dynamics in marine anoxic systems. *Limnol. Oceanogr.* 52, 508–516.
- Corinaldesi, C., Dell'Anno, A., Danovaro, R., 2007b. Early diagenesis and trophic role of extracellular DNA in different benthic ecosystems. *Limnol. Oceanogr.* 52, 1710–1717.
- Corinaldesi, C., Dell'Anno, A., Magagnini, M., Danovaro, R., 2010. Viral decay and viral production rates in continental-shelf and deep-sea sediments of the Mediterranean Sea. *FEMS Microbiol. Ecol.* 72, 208–218.
- Cottrell, M.T., Suttle, C.A., 1995a. Dynamics of a lytic virus infecting the photosynthetic marine picoflagellate *Micromonas pusilla*. *Limnol Ocean.* 40, 730–739.
- Cottrell, M.T., Suttle, C.A., 1995b. Genetic diversity of algal viruses which lyse the photosynthetic picoflagellate *Micromonas pusilla* (Prasinophyceae). *Appl. Environ. Microbiol.* 61, 3088–3091.
- Cottrell, M.T., Suttle, C.A., 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.
- Countway, P.D., Caron, D.A., 2006. Abundance and Distribution of *Ostreococcus* sp. in the San Pedro Channel, California, as Revealed by Quantitative PCR. *Appl. Environ. Microbiol.* 72, 2496–2506. doi:10.1128/AEM.72.4.2496-2506.2006
- Courties, C., Vaquer, A., Troussellier, M., Lautier, J., Chrétiennot-Dinet, M.J., Neveux, J., Machado, C., Claustre, H., 1994. Smallest eukaryotic organism. *Nature* 370, 255–255. doi:10.1038/370255a0

- Culley, A.I., Welschmeyer, N.A., 2002. The abundance, distribution, and correlation of viruses, phytoplankton, and prokaryotes along a Pacific Ocean transect. *Limnol. Oceanogr.* 1508–1513.
- Cuvelier, M.L., Allen, A.E., Monier, A., McCrow, J.P., Messié, M., Tringe, S.G., Woyke, T., Welsh, R.M., Ishoey, T., Lee, J.-H., Binder, B.J., DuPont, C.L., Latasa, M., Guigand, C., Buck, K.R., Hilton, J., Thiagarajan, M., Caler, E., Read, B., Lasken, R.S., Chavez, F.P., Worden, A.Z., 2010. Targeted metagenomics and ecology of globally important uncultured eukaryotic phytoplankton. *Proc. Natl. Acad. Sci.* 107, 14679–14684. doi:10.1073/pnas.1001665107
- D
- Danovaro, R., Corinaldesi, C., Dell'anno, A., Fuhrman, J.A., Middelburg, J.J., Noble, R.T., Suttle, C.A., 2011. Marine viruses and global climate change. *FEMS Microbiol. Rev.* 35, 993–1034. doi:10.1111/j.1574-6976.2010.00258.x
- Danovaro, R., Middelboe, M., 2010. aquatic viral ecology.
- Danovaro, R., Serresi, M., 2000. Viral Density and Virus-to-Bacterium Ratio in Deep-Sea Sediments of the Eastern Mediterranean. *Appl. Environ. Microbiol.* 66, 1857–1861. doi:10.1128/AEM.66.5.1857-1861.2000
- Dasilva, C.R., Li, W.K.W., Lovejoy, C., 2013. Phylogenetic diversity of eukaryotic marine microbial plankton on the Scotian Shelf Northwestern Atlantic Ocean. *J. Plankton Res. fbt123*. doi:10.1093/plankt/fbt123
- Dawkins, R., 2006. *The Selfish Gene: 30th Anniversary Edition*. OUP Oxford.
- Derelle, E., Ferraz, C., Escande, M.-L., Eychenié, S., Cooke, R., Piganeau, G., Desdevises, Y., Bellec, L., Moreau, H., Grimsley, N., 2008. Life-cycle and genome of OtV5, a large DNA virus of the pelagic marine unicellular green alga *Ostreococcus tauri*. *PLoS One* 3, e2250.
- Derelle, E., Ferraz, C., Rombauts, S., Rouzé, P., Worden, A.Z., Robbens, S., Partensky, F., Degroeve, S., Echeynié, S., Cooke, R., others, 2006. Genome analysis of the smallest free-living eukaryote *Ostreococcus tauri* unveils many unique features. *Proc. Natl. Acad. Sci.* 103, 11647–11652.

- Derelle, E., Monier, A., Cooke, R., Worden, A.Z., Grimsley, N.H., Moreau, H., 2015. Diversity of Viruses Infecting the Green Microalga *Ostreococcus lucimarinus*. *J. Virol.* 89, 5812–5821.
- Díez, B., Pedrós-Alió, C., Massana, R., 2001. Study of Genetic Diversity of Eukaryotic Picoplankton in Different Oceanic Regions by Small-Subunit rRNA Gene Cloning and Sequencing. *Appl. Environ. Microbiol.* 67, 2932–2941. doi:10.1128/AEM.67.7.2932-2941.2001
- Doyle, J., 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull* 19, 11–15.
- Dublanchet, A., 2009. Des virus pour combattre les infections: La phagothérapie : renouveau d'un traitement au secours des antibiotiques. Favre.
- Dunigan, D.D., Fitzgerald, L.A., Van Etten, J.L., 2006. Phycodnaviruses: a peek at genetic diversity. *Virus Res.* 117, 119–132.
- E
- Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi:10.1093/nar/gkh340
- Eikrem, W., Throndsen, J., 1990. Phycologia Online - The ultrastructure of *Bathycoccus* gen. nov. and *B. prasinos* sp. nov., a non-motile picoplanktonic alga (Chlorophyta, Prasinophyceae) from the Mediterranean and Atlantic [WWW Document]. URL <http://www.phycologia.org/doi/abs/10.2216/i0031-8884-29-3-344.1> (accessed 7.20.15).
- England, L.S., Pollok, J., Vincent, M., Kreutzweiser, D., Fick, W., Trevors, J.T., Holmes, S.B., 2005. Persistence of extracellular baculoviral DNA in aquatic microcosms: extraction, purification, and amplification by the polymerase chain reaction (PCR). *Mol. Cell. Probes* 19, 75–80.
- Evans, C., Archer, S.D., Jacquet, S., Wilson, W.H., 2003. Direct estimates of the contribution of viral lysis and microzooplankton grazing to the decline of a *Micromonas spp.* population. *Aquat Microb. Ecol* 30, 207–219.

F

- Finsterbusch, T., Mankertz, A., 2009. Porcine circoviruses--small but powerful. *Virus Res.* 143, 177–183. doi:10.1016/j.virusres.2009.02.009
- Flint, S.J., Enquist, L.W., Racaniello, V.R., Skalka, A.M., Barnum, D.R., de Evaluación, E., 2000. *Principles of Virology: Molecular Biology, Patogenesis and.* ASM Press. Washington DC USA.
- Flores, C.O., Meyer, J.R., Valverde, S., Farr, L., Weitz, J.S., 2011. , in: *Statistical Structure of Host–phage Interactions.* *Proc Natl Acad Sci USA* 108. pp. 288–297.
- Fofonoff, N.P., Millard, R.C., 1983. Algorithms for computation of fundamental properties of seawater.
- Forterre, P., 2010. Giant viruses: conflicts in revisiting the virus concept. *Intervirology* 53, 362–378. doi:10.1159/000312921
- Forterre, P., 2006. The origin of viruses and their possible roles in major evolutionary transitions. *Virus Res.* 117, 5–16. doi:10.1016/j.virusres.2006.01.010
- Forterre, P., Soler, N., Krupovic, M., Marguet, E., & Ackermann, H. W., 2013. Fake virus particles generated by fluorescence microscopy. *Trends in Microbiology*, 1(21), 1-5.
- Foulon, E., Not, F., Jalabert, F., Cariou, T., Massana, R., Simon, N., 2008. Ecological niche partitioning in the picoplanktonic green alga *Micromonas pusilla*: evidence from environmental surveys using phylogenetic probes. *Env. Microbiol* 10, 2433–2443.
- Frada, M., Probert, I., Allen, M.J., Wilson, W.H., De Vargas, C., 2008. The “Cheshire Cat” escape strategy of the coccolithophore *Emiliana huxleyi* in response to viral infection. *Proc. Natl. Acad. Sci.* 105, 15944–15949.
- Fuhrman, J., 1992. Bacterioplankton Roles in Cycling of Organic Matter: The Microbial Food Web, in: Falkowski, P.G., Woodhead, A.D., Vivirito, K. (Eds.), *Primary Productivity and Biogeochemical Cycles in the Sea*, Environmental Science Research. Springer US, pp. 361–383.

G

- Galluzzi, L., Penna, A., Bertozzini, E., Vila, M., Garcés, E., Magnani, M., 2004. Development of a Real-Time PCR Assay for Rapid Detection and Quantification of *Alexandrium minutum* (a Dinoflagellate). *Appl. Environ. Microbiol.* 70, 1199–1206. doi:10.1128/AEM.70.2.1199-1206.2004
- Gao, X., Song, J., 2005. Phytoplankton distributions and their relationship with the environment in the Changjiang Estuary, China. *Mar. Pollut. Bull.* 50, 327–335. doi:10.1016/j.marpolbul.2004.11.004
- Gerba, C.P., 1984. Microorganisms as groundwater tracers, in: *Groundwater Pollution Microbiology*. John Wiley & Sons, pp. 225–33.
- Gastrich, M.D., Leigh-Bell, J.A., Gobler, C.J., Anderson, O.R., Wilhelm, S.W., Bryan, M., 2004. Viruses as potential regulators of regional brown tide blooms caused by the alga, *Aureococcus anophagefferens*. *Estuaries* 27, 112–119.
- Gieskes, W.W.C., Buma, A.G.J., 1997. UV damage to plant life in a photobiologically dynamic environment: the case of marine phytoplankton, in: Rozema, J., Gieskes, W.W.C., Geijn, S.C.V.D., Nolan, C., Boois, H.D. (Eds.), *UV-B and Biosphere, Advances in Vegetation Science*. Springer Netherlands, pp. 16–25.
- Glansdorff, N., Xu, Y., Labedan, B., 2008. The last universal common ancestor: emergence, constitution and genetic legacy of an elusive forerunner. *Biol. Direct* 3, 29. doi:10.1186/1745-6150-3-29
- Gobler, C.J., Hutchins, D.A., Fisher, N.S., Coper, E.M., Sañudo-Wilhelmy, S.A., 1997. Release and bioavailability of C, N, P Se, and Fe following viral lysis of a marine chrysophyte. *Limnol. Oceanogr.* 42, 1492–1504.
- Gogarten, J.P., Senejani, A.G., Zhaxybayeva, O., Olendzenski, L., Hilario, E., 2002. Inteins: structure, function, and evolution. *Annual Reviews in Microbiology* 56, 263–287.
- Gonzalez, P., Labarère, J., 1998. Sequence and Secondary Structure of the Mitochondrial Small-Subunit rRNA V4, V6, and V9 Domains Reveal Highly Species-Specific Variations within the Genus *Agrocybe* [WWW Document]. URL <http://aem.asm.org/content/64/11/4149.short> (accessed 7.20.15).

- Gonzalez, J.M., Suttle, C.A., 1993. Grazing by marine nanoflagellates on viruses and virus-sized particles: ingestion and digestion. *Marine Ecology Progress Series MESED*, 94.
- Grant, S.B., List, E.J., Lidstrom, M.E., 1993. Kinetic analysis of virus adsorption and inactivation in batch experiments. *Water Resour. Res.* 29, 2067–2085.
- Grob, C., Ulloa, O., Claustre, H., Huot, Y., Alarcón, G., Marie, D., 2007. Contribution of picoplankton to the total particulate organic carbon concentration in the eastern South Pacific. *Biogeosciences* 4, 837–852.
- Guillou, L., Eikrem, W., Chrétiennot-Dinet, M.-J., Le Gall, F., Massana, R., Romari, K., Pedrós-Alió, C., Vaulot, D., 2004. Diversity of Picoplanktonic Prasinophytes Assessed by Direct Nuclear SSU rDNA Sequencing of Environmental Samples and Novel Isolates Retrieved from Oceanic and Coastal Marine Ecosystems. *Protist* 155, 193–214. doi:10.1078/143446104774199592
- H
- Hara, S., Koike, I., Terauchi, K., Kamiya, H., Tanoue, E., 1996. abundance of viruses in deep ocean waters. *Mar. Ecol. Prog. Ser.* 269–277.
- Harvey, R.W., Ryan, J.N., 2004. Use of PRD1 bacteriophage in groundwater viral transport, inactivation, and attachment studies. *FEMS Microbiol. Ecol.* 49, 3–16.
- Henrich, E., 1988. Contribution à l'étude de la capacité biotique de la partie sud de l'étang de Salses-Leucatee, thèse de l'université de Barcelone
- Hensen, V., 1887. Über die Bestimmung des Planktons oder des im Meere treibenden Materials an Pflanzen und Tieren. Schmidt & Klaunig.
- Hérelle, F., princeps : Sur un microbe invisible antagoniste des bacilles dysentériques *Comptes-Rendus de l'Académie des Sciences, Paris*, 1917. 165 :p. 373-5
- Herve, P., 1978. Ichthyofaune comparée de deux étangs littoraux du Roussillon: Canet-St Nazaire et Salses-Leucate. Univ. PM Curie, Paris.

Hingamp, P., Grimsley, N., Acinas, S.G., Clerissi, C., Subirana, L., Poulain, J., Ferrera, I., Sarmiento, H., Villar, E., Lima-Mendez, G., 2013. Exploring nucleo-cytoplasmic large DNA viruses in Tara Oceans microbial metagenomes. *The ISME journal* 7, 1678–1695.

Hofer, J.S., Sommaruga, 2001. Seasonal dynamics of viruses in an alpine lake: importance of filamentous forms. *Aquatic Microbial Ecology* 26, 1–11.

Hwang, C.Y., Cho, B.C., 2002. Virus-infected bacteria in oligotrophic open waters of the East Sea, Korea. *Aquat. Microb. Ecol.* 30, 1–9.

I

Ishizaka, J., Harada, K., Ishikawa, K., Kiyosawa, H., Furusawa, H., Watanabe, Y., Ishida, H., Suzuki, K., Handa, N., Takahashi, M., 1997. Size and taxonomic plankton community structure and carbon flow at the equator, 175°E during 1990–1994. *Deep Sea Res. Part II Top. Stud. Oceanogr., A JGFOS Process Study in the Equatorial Pacific* 44, 1927–1949. doi:10.1016/S0967-0645(97)00019-2

Iyer, L.M., Aravind, L., Koonin, E.V., 2001. Common origin of four diverse families of large eukaryotic DNA viruses. *J. Virol.* 75, 11720–11734.

Iyer, L.M., Balaji, S., Koonin, E.V., Aravind, L., 2006. Evolutionary genomics of nucleo-cytoplasmic large DNA viruses. *Virus Res.* 117, 156–184.

J

Jacques, G., Çahet, G., Fiala, M., Neveux, J., Panouse, M., 1975. Caractéristiques du milieu pélagique des étangs de Bages-Sigean et de Salses-Leucate pendant l'été 1974. *Vie Milieu* 25, 1–18.

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. *Aquatic Microbial Ecology* 27, 111–124.

- Jacquet, S., Miki, T., Noble, R., Peduzzi, P., Wilhelm, S., 2010. Viruses in aquatic ecosystems: important advancements of the last 20 years and prospects for the future in the field of microbial oceanography and limnology. *Advances in Oceanography and Limnology* 1, 97–141.
- Jamindar, S., Polson, S.W., Srinivasiah, S., Waidner, L., Wommack, K.E., 2012. Evaluation of two approaches for assessing the genetic similarity of virioplankton populations as defined by genome size. *Applied and environmental microbiology* 78, 8773–8783.
- Jeffrey, W.H., Kase, J.P., Wilhelm, S.W., 2000. UV radiation effects on heterotrophic bacterioplankton and viruses in marine ecosystems. *The effects of UV radiation in the marine environment* 10, 206.
- Jiang, S., Fu, W., Chu, W., Fuhrman, J.A., 2003. The vertical distribution and diversity of marine bacteriophage at a station off Southern California. *Microb. Ecol.* 45, 399–410.
- Jin, S., Zhang, B., Weisz, O.A., Montelaro, R.C., 2005. Receptor-mediated entry by equine infectious anemia virus utilizes a pH-dependent endocytic pathway. *J. Virol.* 79, 14489–14497.
- John, S.G., Mendez, C.B., Deng, L., Poulos, B., Kauffman, A.K.M., Kern, S., Brum, J., Polz, M.F., Boyle, E.A., Sullivan, M.B., 2011. A simple and efficient method for concentration of ocean viruses by chemical flocculation. *Environ. Microbiol. Rep.* 3, 195–202.
- Johnson, P.W., Sieburth, J.M., 1982. In-Situ Morphology and Occurrence of Eucaryotic Phototrophs of Bacterial Size in the Picoplankton of Estuarine and Oceanic Waters1. *J. Phycol.* 18, 318–327. doi:10.1111/j.1529-8817.1982.tb03190.x
- Johnson, W.E., 2010. Endless Forms Most Viral. *PLoS Genet* 6, e1001210. doi:10.1371/journal.pgen.1001210

K

- Krom, M.D., Kress, N., Brenner, S., Gordon, L.I., 1991. Phosphorus limitation of primary productivity in the eastern Mediterranean Sea. *Limnology and Oceanography* 36, 424–432.
- Krueger, A.P., Fong, J., 1937. The relationship between bacterial growth and phage production. *J. Gen. Physiol.* 21, 137.
- Kutateladze, M., Adamia, R., 2008. Phage therapy experience at the Eliava Institute. *Médecine Mal. Infect., Numéro spécial CEMI 2007 et 2008* 38, 426–430. doi:10.1016/j.medmal.2008.06.023

L

- Ladagnous, H., Le Bec, C., 1997. Lagune de Salses-Leucate. I.-Synthèse bibliographique, contamination bactériologique et approche biogéochimique.
- Laghdass, M., Catala, P., Caparros, J., Oriol, L., Lebaron, P., Obernosterer, I., 2012. High contribution of SAR11 to microbial activity in the North West Mediterranean Sea. *Microbial ecology* 63, 324–333.
- Larsen, A., Castberg, T., Sandaa, R.A., Brussaard, C.P.D., Egge, J., Heldal, M., Paulino, A., Thyraug, R., Van Hannen, E.J., Bratbak, G., 2001. Population dynamics and diversity of phytoplankton, bacteria and viruses in a seawater enclosure. *Mar. Ecol. Prog. Ser.* 221, 47–57.
- Larsen, J.B., Larsen, A., Bratbak, G., Sandaa, R.-A., 2008. Phylogenetic analysis of members of the Phycodnaviridae virus family, using amplified fragments of the major capsid protein gene. *Appl. Environ. Microbiol.* 74, 3048–3057.
- La Scola, B., Desnues, C., Pagnier, I., Robert, C., Barrassi, L., Fournous, G., Merchat, M., Suzan-Monti, M., Forterre, P., Koonin, E., Raoult, D., 2008. The virophage as a unique parasite of the giant mimivirus. *Nature* 455, 100–104. doi:10.1038/nature07218

- Litchman, E., Klausmeier, C.A., 2008. Trait-based community ecology of phytoplankton. *Annu. Rev. Ecol. Evol. Syst.* 615–639.
- López-Bueno, A., Tamames, J., Velázquez, D., Moya, A., Quesada, A., Alcamí, A., 2009. High diversity of the viral community from an Antarctic lake. *Science* 326, 858–861.
- Luisi, P.L., 1998. About various definitions of life. *Orig Life Evol Biosph* 28:613–622
- Lwoff, J., 1957, The concept of virus. *Gen. Microbiol.* 1957 Oct; 17(2) : 239-53.
- M
- MacArthur, R.H., 1967. *The theory of island biogeography* (Vol. 1). Princeton University Press.
- Malaterre, C., 2010. Lifeness signatures and the roots of the tree of life. *Biol. Philos.* 25, 643–658. doi:10.1007/s10539-010-9220-8
- Malin, G., Turner, S.M., Liss, P.S., 1992. Sulfur: the plankton/climate connection. *J. Phycol.* 28, 590–597.
- Malin, G., Wilson, W.H., Bratbak, G., Liss, P.S., Mann, N.H., 1998. Elevated production of dimethylsulfide resulting from viral infection of cultures of *Phaeocystis pouchetii*. *Limnol. Oceanogr.* 43, 1389–1393. doi:10.4319/lo.1998.43.6.1389
- Manage, P.M., Kawabata, Z., Nakano, S., Nishibe, Y., 2002. Effect of heterotrophic nanoflagellates on the loss of virus-like particles in pond water. *Ecological Research* 17, 473–479.
- 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.
- Manton, I., Parke, M., 1960. Further observations on small on small green flagellates with special reference to possible relatives of *chromulina pussilla*. *J Mar Biol Assoc* 35, 275–298.

- 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. doi:10.1016/j.protis.2009.10.002
- Marrec, P., Cariou, T., collin, E., durand, A., Latimier, M., Macé, E., 2013. Seasonal and latitudinal variability of the CO₂ system in the western English Channel based on Voluntary Observing Ship (VOS) measurements. *Mar Chem* 155, 29–41.
- Martínez, J.M., Schroeder, D.C., Larsen, A., Bratbak, G., Wilson, W.H., 2007. Molecular dynamics of *Emiliana huxleyi* and cooccurring viruses during two separate mesocosm studies. *Appl. Environ. Microbiol.* 73, 554–562.
- Martinez Martinez, J., Boere, A., Gilg, I., van Lent, J.W.M., Witte, H.J., L, van B.J.D., Brussaard, C.P.D., 2015. New lipid envelope-containing dsDNA virus isolates infecting *Micromonas pusilla* reveal a separate phylogenetic group. *Aquat Microb Ecol* 74, 17–28.
- Massana, R., 2011. Eukaryotic picoplankton in surface oceans. *Annu. Rev. Microbiol.* 65, 91–110.
- Mayer, J.A., Taylor, F.J.R., 1979. A virus which lyses the marine nanoflagellate *Micromonas pusilla*. *Nature* 281, 299–301.
- Middelboe, M., Jorgensen, N., Kroer, N., 1996. Effects of viruses on nutrient turnover and growth efficiency of noninfected marine bacterioplankton. *Appl. Environ. Microbiol.* 62, 1991–1997.
- Middelboe, M., Riemann, L., Steward, G.F., Hansen, V., Nybroe, O., 2003. Virus-induced transfer of organic carbon between marine bacteria in a model community. *Aquat. Microb. Ecol.* 33, 1–10.
- Mizumoto, H., Tomaru, Y., Takao, Y., Shirai, Y., Nagasaki, K., 2007. Intraspecies host specificity of a single-stranded RNA virus infecting a marine photosynthetic protist is determined at the early steps of infection. *Journal of virology* 81, 1372–1378.

- Mojica, K.D.A., Brussaard, C.P.D., 2014. Factors affecting virus dynamics and microbial host-virus interactions in marine environments. *FEMS Microbiol. Ecol.* n/a-n/a. doi:10.1111/1574-6941.12343
- Moon-van der Staay, S.Y., De Wachter, R., Vault, D., 2001. Oceanic 18S rDNA sequences from picoplankton reveal unsuspected eukaryotic diversity. *Nature* 409, 607–610. doi:10.1038/35054541
- 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.
- Moreau, H., Verhelst, B., Couloux, A., Derelle, E., Rombauts, S., Grimsley, N., Van Bel, M., Poulain, J., Katinka, M., Hohmann-Marriott, M.F., others, 2012. Gene functionalities and genome structure in *Bathycoccus* prasinus reflect cellular specializations at the base of the green lineage. *Genome Biol* 13, R74.
- Moreira, D., López-García, P., 2009. Ten reasons to exclude viruses from the tree of life. *Nat. Rev. Microbiol.* 7, 306–311. doi:10.1038/nrmicro2108
- Muck, S., Griessler, T., Köstner, N., Klimiuk, A., Winter, C., Herndl, G.J., 2014. Fracture zones in the Mid Atlantic Ridge lead to alterations in prokaryotic and viral parameters in deep-water masses. *Front. Microbiol.* 5. doi:10.3389/fmicb.2014.00264

N

- Nagasaki, K., Ando, M., Itakura, S., Imai, I., Ishida, Y., 1994. Viral mortality in the final stage of *Heterosigma akashiwo* (Raphidophyceae) red tide. *J. Plankton Res.* 16, 1595–1599.
- Nagasaki, K., Tarutani, K., Yamaguchi, M., 1999. Growth characteristics of *Heterosigma akashiwo* virus and its possible use as a microbiological agent for red tide control. *Appl. Environ. Microbiol.* 65, 898–902.

- Nagasaki, K., Yamaguchi, M., 1998. Effect of temperature on the algicidal activity and the stability of HaV (Heterosigma akashiwo virus). *Aquat. Microb. Ecol.* 15, 211–216. doi:10.3354/ame015211
- Nagasaki, K., Ando, M., Itakura, S., Imai, I., Ishida, Y., 1994. Viral mortality in the final stage of *Heterosigma akashiwo* (Raphidophyceae) red tide. *Journal of Plankton Research* 16, 1595–1599.
- Newton, A., Mudge, S.M., 2003. Temperature and salinity regimes in a shallow, mesotidal lagoon, the Ria Formosa, Portugal. *Estuar. Coast. Shelf Sci.* 57, 73–85. doi:10.1016/S0272-7714(02)00332-3
- Nielsen, K.M., Johnsen, P. al J., Bensasson, D., Daffonchio, D., 2007. Release and persistence of extracellular DNA in the environment. *Environ. Biosafety Res.* 6, 37–53.
- Noble, R.T., Fuhrman, J.A., 2000a. Rapid virus production and removal as measured with fluorescently labeled viruses as tracers. *Appl. Environ. Microbiol.* 66, 3790–3797.
- Noble, R.T., Fuhrman, J.A., 1997. Virus decay and its causes in coastal waters. *Applied and Environmental Microbiology* 63, 77–83.
- Not, F., Latasa, M., Scharek, R., Viprey, M., Karleskind, P., Balagué, V., Ontoria-Oviedo, I., Cumino, A., Goetze, E., Vaultot, D., others, 2008. Protistan assemblages across the Indian Ocean, with a specific emphasis on the picoeukaryotes. *Deep Sea Research Part I: Oceanographic Research Papers* 55, 1456–1473.
- Not, F., Latasa, M., Marie, D., Cariou, T., Vaultot, D., Simon, N., 2004. A single species, *Micromonas pusilla* (Prasinophyceae), dominates the eukaryotic picoplankton in the western English Channel. *Appl. Environ. Microbiol.* 70, 4064–4072.

O

- O'Kelly, C.J., Sieracki, M.E., Thier, E.C., Hobson, I.C., 2003. A Transient Bloom of *Ostreococcus* (chlorophyta, Prasinophyceae) in West Neck Bay, Long Island, New York. *J. Phycol.* 39, 850–854. doi:10.1046/j.1529-8817.2003.02201.x

P

- Palenik, B., Grimwood, J., Aerts, A., Rouzé, P., Salamov, A., Putnam, N., Dupont, C., Jorgensen, R., Derelle, E., Rombauts, S., others, 2007. The tiny eukaryote *Ostreococcus* provides genomic insights into the paradox of plankton speciation. *Proc. Natl. Acad. Sci.* 104, 7705–7710.
- Parada, V., Herndl, G.J., Weinbauer, M.G., 2006. Viral burst size of heterotrophic prokaryotes in aquatic systems. *J. Mar. Biol. Assoc. U. K.* 86, 613–621.
- Parada, V., Sintes, E., van Aken, H.M., Weinbauer, M.G., Herndl, G.J., 2007. Viral Abundance, Decay, and Diversity in the Meso- and Bathypelagic Waters of the North Atlantic. *Appl. Environ. Microbiol.* 73, 4429–4438. doi:10.1128/AEM.00029-07
- Parvathi, A., Zhong, X., Jacquet, S., 2012. Dynamics of various viral groups infecting autotrophic plankton in Lake Geneva. *Advances in Oceanography and Limnology* 3, 171–191.
- Paul, J.H., Jeffrey, W.H., DeFlaun, M.F., 1987. Dynamics of extracellular DNA in the marine environment. *Appl. Environ. Microbiol.* 53, 170–179.
- Pearson, H., 2008. “Virophage” suggests viruses are alive. *Nat. News* 454, 677–677. doi:10.1038/454677a
- Pernthler, J., Pernthaler, Annelie, Rudolf, Amann, 2003. Automated Enumeration of Groups of Marine Picoplankton after Fluorescence In Situ Hybridization.
- Philippe, N., Legendre, M., Dautre, G., Couté, Y., Poirot, O., Lescot, M., Arslan, D., Seltzer, V., Bertaux, L., Bruley, C., Garin, J., Claverie, J.-M., Abergel, C., 2013. Pandoraviruses: Amoeba Viruses with Genomes Up to 2.5 Mb Reaching That of Parasitic Eukaryotes. *Science* 341, 281–286. doi:10.1126/science.1239181
- Pienaar, R.N., 1976. Virus-like particles in three species of phytoplankton from San Juan. *Isl. Wash. Phycol.* 15, 185–190.
- Piganeau, G., Grimsley, N., Moreau, H., 2011. Genome diversity in the smallest marine photosynthetic eukaryotes. *Res. Microbiol.* 162, 570–577.

- Poorvin, L., Rinta-Kanto, J.M., Hutchins, D.A., Wilhelm, S.W., 2004. Viral release of iron and its bioavailability to marine plankton. *Limnol. Oceanogr.* 49, 1734–1741.
- Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14, 817–818.
- Poulin, R., 1998. Large-scale patterns of host use by parasites of freshwater fishes. *Ecol Lett.* 1, 118–128.
- Proctor, L.M., 1991. Roles of viral infection in organic particle flux. *Mar Ecol Prog Ser* 69, 133–142. doi:10.3354/meps069133
- Proctor, L.M., Fuhrman, J.A., 1990. Viral mortality of marine bacteria and cyanobacteria. *Nature* 343, 60–62. doi:10.1038/343060a0
- Prokopowich, C.D., Gregory, T.R., Crease, T.J., 2003. The correlation between rDNA copy number and genome size in eukaryotes - *Genome*. *Genome* 46, 48–50. doi:10.1139/G02-103
- R
- Raoult, D., Audic, S., Robert, C., Abergel, C., Renesto, P., Ogata, H., La Scola, B., Suzan, M., Claverie, J.-M., 2004. The 1.2-megabase genome sequence of Mimivirus. *Science* 306, 1344–1350.
- Reiter, D.M., Frierson, J.M., Halvorson, E.E., Kobayashi, T., Dermody, T.S., Stehle, T., 2011. Crystal structure of reovirus attachment protein σ 1 in complex with sialylated oligosaccharides. *PLoS Pathog.* 7, 1002166.
- Riemann, L., Middelboe, M., 2002. Stability of bacterial and viral community compositions in Danish coastal waters as depicted by DNA fingerprinting techniques. *Aquat. Microb. Ecol.* 27, 219–232.
- Rodríguez, F., Derelle, E., Guillou, L., Le Gall, F., Vaulot, D., Moreau, H., 2005. Ecotype diversity in the marine picoeukaryote *Ostreococcus* (Chlorophyta, Prasinophyceae). *Environ. Microbiol.* 7, 853–859.

- Rohwer, F., Edwards, R., 2002. The Phage Proteomic Tree: a genome-based taxonomy for phage. *Journal of bacteriology* 184, 4529–4535.
- Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* 61, 539–542.
- Rosario, K., Nilsson, C., Lim, Y.W., Ruan, Y., Breitbart, M., 2009. Metagenomic analysis of viruses in reclaimed water. *Environmental Microbiology* 11, 2806–2820.
- Rowe, J.M., DeBruyn, J.M., Poorvin, L., LeCleir, G.R., Johnson, Z.I., Zinser, E.R., Wilhelm, S.W., 2012. Viral and bacterial abundance and production in the Western Pacific Ocean and the relation to other oceanic realms. *FEMS Microbiol. Ecol.* 79, 359–370.
- Ruardij, P., Veldhuis, M.J., Brussaard, C.P., 2005. Modeling the bloom dynamics of the polymorphic phytoplankter *Phaeocystis globosa*: impact of grazers and viruses. *Harmful Algae* 4, 941–963.
- Rueckert, R.R., 1996. Picornaviridae: the viruses and their replication. *Fields Virol.* 1, 609–654.
- Ruiz-Mirazo K, Pereto J, Moreno A, 2004. A universal definition of life: autonomy and open-ended evolution. *Orig Life Evol Biosph* 34(3):323–346
- S
- Sahlsten, E., 1998. Seasonal abundance in Skagerrak-Kattegat coastal waters and host specificity of viruses infecting the marine photosynthetic flagellate *Micromonas pusilla*. *Aquat Microb. Ecol* 16, 103–108.
- Salter, I., Galand, P.E., Fagervold, S.K., Lebaron, P., Obernosterer, I., Oliver, M.J., Suzuki, M.T., Tricoire, C., 2015. Seasonal dynamics of active SAR11 ecotypes in the oligotrophic Northwest Mediterranean Sea. *The ISME journal* 9, 347–360.
- Sano, E., Carlson, S., Wegley, L., Rohwer, F., 2004. Movement of viruses between biomes. *Appl. Environ. Microbiol.* 70, 5842–5846.

- Säwström, C., Anesio, M.A., Granéli, W., Laybourn-Parry, J., 2007. Seasonal viral loop dynamics in two large ultraoligotrophic Antarctic freshwater lakes. *Microbial ecology* 53, 1–11.
- Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., others, 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 *Emiliana huxleyi* bloom. *Appl. Environ. Microbiol.* 69, 2484–2490.
- Severin, T., Conan, P., de Madron, X.D., Houpert, L., Oliver, M.J., Oriol, L., Caparros, J., Ghiglione, J.F., Pujo-Pay, M., 2014. Impact of open-ocean convection on nutrients, phytoplankton biomass and activity. *Deep Sea Res. Part Oceanogr. Res. Pap.* 94, 62–71.
- Shi, X.L., Marie, D., Jardillier, L., Scanlan, D.J., Vaultot, D., 2009. Groups without cultured representatives dominate eukaryotic picophytoplankton in the oligotrophic South East Pacific Ocean. *PLoS One* 4, e7657–e7657.
- Short, C.M., Rusanova, O., Short, S.M., 2010. Quantification of virus genes provides evidence for seed-bank populations of phycodnaviruses in Lake Ontario, Canada. *ISME J.* 5, 810–821.
- Short, S.M., 2012. The ecology of viruses that infect eukaryotic algae: Algal virus ecology. *Environ. Microbiol.* 14, 2253–2271. doi:10.1111/j.1462-2920.2012.02706.x
- Short, S.M., Short, C.M., 2009. Quantitative PCR reveals transient and persistent algal viruses in Lake Ontario, Canada. *Environ. Microbiol.* 11, 2639–2648. doi:10.1111/j.1462-2920.2009.01988.x
- Short, S.M., Short, C.M., 2008. Diversity of algal viruses in various North American freshwater environments *Aquat. Microb. Ecol* 51, 13–21.
- Short, S.M., Suttle, C.A., 2003. Temporal dynamics of natural communities of marine algal viruses and eukaryotes. *Aquat. Microb. Ecol.* 32, 107–119.

- Sieburth, J.M., Johnson, P.W., Hargraves, P.E., 1988. Ultrastructure and ecology of *Aureococcus anophageferens* gen. et sp. nov.(Chrysophyceae): the dominant picoplankton during a bloom in Narragansett Bay, Rhode Island, summer 19851. *Journal of Phycology* 24, 416–425.
- Sime-Ngando, T., 2014. Environmental bacteriophages: viruses of microbes in aquatic ecosystems. *Front Microbiol* 5, 355.
- Simon, N., Cras, A.-L., Foulon, E., Lemée, R., 2009. Diversity and evolution of marine phytoplankton. *C. R. Biol., La théorie de Darwin revisitée par la biologie d'aujourd'hui / Darwin's theory revisited by today's biology* 332, 159–170. doi:10.1016/j.crv.2008.09.009
- Šlapeta, J., López-García, P., Moreira, D., 2006. Global Dispersal and Ancient Cryptic Species in the Smallest Marine Eukaryotes. *Mol. Biol. Evol.* 23, 23–29. doi:10.1093/molbev/msj001
- Sobsey, M.D., Wallis, C., Melnick, J.L., 1975. Studies on the survival and fate of enteroviruses in an experimental model of a municipal solid waste landfill and leachate. *Appl. Microbiol.* 30, 565–574.
- Speel, E.J.M., Hopman, A.H.N., Komminoth, P., 1999. Amplification Methods to Increase the Sensitivity of In Situ Hybridization: Play CARD(S). *J. Histochem. Cytochem.* 47, 281–288. doi:10.1177/002215549904700302
- Spencer, R., 1955. A Marine Bacteriophage. *Nature* 175, 690–691. doi:10.1038/175690a0
- Steward, G.F., 2001. Fingerprinting viral assemblages by pulsed field gel electrophoresis (PFGE). *Methods Microbiol.* 30, 85–103.
- Stoddard, L.I., Martiny, J.B., Marston, M.F., 2007. Selection and characterization of cyanophage resistance in marine *Synechococcus* strains. *Appl Env. Microbiol* 73, 5516–5522.

- Stoeck, T., Bass, D., Nebel, M., Christen, R., Jones, M.D.M., Breiner, H.-W., Richards, T.A., 2010. Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol. Ecol.* 19, 21–31. doi:10.1111/j.1365-294X.2009.04480.x
- Subirana, L., Péquin, B., Michely, S., Escande, M.-L., Meilland, J., Derelle, E., Marin, B., Piganeau, G., Desdevises, Y., Moreau, H., Grimsley, N.H., 2013. Morphology, genome plasticity, and phylogeny in the genus *ostreococcus* reveal a cryptic species, *O. mediterraneus* sp. nov. (Mamiellales, Mamiellophyceae). *Protist* 164, 643–659. doi:10.1016/j.protis.2013.06.002
- Suttle, C.A., 2007. Marine viruses - major players in the global ecosystem. *Nat. Rev Microbiol* 5, 801–812.
- Suttle, C.A., 2005. Viruses in the sea. *Nature* 437, 356–361.
- Suttle, C.A., 2002. Cyanophages and their role in the ecology of cyanobacteria, in: *The Ecology of Cyanobacteria*. Springer, pp. 563–589.
- Suttle, C.A., 2000. Ecological, evolutionary, and geochemical consequences of viral infection of cyanobacteria and eukaryotic algae. *Viral ecology* 1, 247–296.
- Suttle, C.A., Chan, A.M., 1994. Dynamics and distribution of cyanophages and their effect on marine *Synechococcus spp.* *Applied and environmental microbiology* 60, 3167–3174.
- Suttle, C.A., Chen, F., 1992. Mechanisms and rates of decay of marine viruses in seawater. *Applied and Environmental Microbiology* 58, 3721–3729.
- Suzuki, M.T., Taylor, L.T., DeLong, E.F., 2000. Quantitative Analysis of Small-Subunit rRNA Genes in Mixed Microbial Populations via 5' -Nuclease Assays. *Appl. Environ. Microbiol.* 66, 4605–4614. doi:10.1128/AEM.66.11.4605-4614.2000
- Swofford, D.L., 2003. PAUP: Phylogenetic analysis using parsimony. Version 4. Sinauer Associates, Sunderland, Massachusetts.

T

- Takemura, M., 2001. Poxviruses and the Origin of the Eukaryotic Nucleus. *J. Mol. Evol.* 52, 419–425. doi:10.1007/s002390010171
- Tarutani, K., Nagasaki, K., Yamaguchi, M., 2006. Virus adsorption process determines virus susceptibility in *Heterosigma akashiwo* (Raphidophyceae). *Aquat. Microb. Ecol* 42, 209–213.
- Tarutani, K., Nagasaki, K., Yamaguchi, M., 2000. Viral Impacts on Total Abundance and Clonal Composition of the Harmful Bloom-Forming Phytoplankton *Heterosigma akashiwo*. *Appl. Environ. Microbiol.* 66, 4916–4920.
- Taylor, G.T., Hein, C., Iabichella, M., 2003. Temporal variations in viral distributions in the anoxic Cariaco Basin. *Aquat. Microb. Ecol.* 30, 103–116.
- Taylor, G.T., Iabichella, M., Ho, T.-Y., Scranton, M.I., Thunell, R.C., Muller-Karger, F., Varela, R., 2001. Chemoautotrophy in the redox transition zone of the Cariaco Basin: a significant midwater source of organic carbon production. *Limnol. Oceanogr.* 46, 148–163.
- Thingstad, T.F., 2000. Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. *Limnol. Oceanogr.* 1320–1328.
- Thioulouse, J., Chessel, D., 1987. Les analyses multitableaux en écologie factorielle. I: De la typologie d'état à la typologie de fonctionnement par l'analyse triadique. *Acta Oecologica Oecologia Gen.* 8, 463–480.
- Thioulouse, J., Simier, M., Chessel, D., 2004. Simultaneous analysis of a sequence of paired ecological tables. *Ecology* 85, 272–283.
- Thomas, R., Grimsley, N., Escande, M.-L., Subirana, L., Derelle, E., Moreau, H., 2011a. Acquisition and maintenance of resistance to viruses in eukaryotic phytoplankton populations. *Env. Microbiol* 13, 1412–1420.
- Thronsen, J., 1976. Occurrence and productivity of small marine flagellates. *Nor. J. Bot.*

- Thronsdén, J., Kristiansen, S., 1991. *Micromonas pusilla* (Prasinophyceae) as part of pico- and nanoplankton communities of the Barents Sea - THRONSDÉN - 2007 - Polar Research - Wiley Online Library [WWW Document].
- Thurber, R.V., Haynes, M., Breitbart, M., Wegley, L., Rohwer, F., 2009. Laboratory procedures to generate viral metagenomes. *Nat. Protoc.* 4, 470–483.
- Tidona, C.A., Schnitzler, P., Kehm, R., Darai, G., 1998. Is the major capsid protein of iridoviruses a suitable target for the study of viral evolution? *Virus Genes* 16, 59–66.
- Tijdens, M., Hoogveld, H.L., Kamst-van Agterveld, M.P., Simis, S.G., Baudoux, A.-C., Laanbroek, H.J., Gons, H.J., 2008. Population dynamics and diversity of viruses, bacteria and phytoplankton in a shallow eutrophic lake. *Microbial ecology* 56, 29–42.
- Tomaru, Y., Tarutani, K., Yamaguchi, M., Nagasaki, K., 2004. Quantitative and qualitative impacts of viral infection on a *Heterosigma akashiwo* (Raphidophyceae) bloom in Hiroshima Bay, Japan. *Aquat. Microb. Ecol.* 34, 227–238.
- Tuomi, P., Fagerbakke, K.M., Bratbak, G., Heldal, M., 1995. Nutritional enrichment of a microbial community: The effects on activity, elemental composition, community structure and virus production. *FEMS Microbiol. Ecol.* 16, 123–134. doi:10.1111/j.1574-6941.1995.tb00276.x

V

- Van Boekel, W.H.M., 1992. Lysis-induced decline of *Phaeocystis* spring bloom and coupling with the microbial foodweb. *Mar Ecol Prog Ser* 81, 269–276. doi:10.3354/meps081269
- 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 Etten, J.L., Meints, R.H., 1999. Giant viruses infecting algae. *Annu. Rev. Microbiol.* 53, 447–494.

- Van Valen, L., 1973. A new evolutionary law. *Evol. Theory* 1, 1–30.
- Vargas, C. de, Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., Lara, E., Berney, C., Bescot, N.L., Probert, I., Carmichael, M., Poulain, J., Romac, S., Colin, S., Aury, J.-M., Bittner, L., Chaffron, S., Dunthorn, M., Engelen, S., Flegontova, O., Guidi, L., Horák, A., Jaillon, O., Lima-Mendez, G., Lukeš, J., Malviya, S., Morard, R., Mulo, M., Scalco, E., Siano, R., Vincent, F., Zingone, A., Dimier, C., Picheral, M., Searson, S., Kandels-Lewis, S., Coordinators, T.O., Acinas, S.G., Bork, P., Bowler, C., Gorsky, G., Grimsley, N., Hingamp, P., Iudicone, D., Not, F., Ogata, H., Pesant, S., Raes, J., Sieracki, M.E., Speich, S., Stemann, L., Sunagawa, S., Weissenbach, J., Wincker, P., Karsenti, E., Boss, E., Follows, M., Karp-Boss, L., Krzic, U., Reynaud, E.G., Sardet, C., Sullivan, M.B., Velayoudon, D., 2015. Eukaryotic plankton diversity in the sunlit ocean. *Science* 348, 1261605. doi:10.1126/science.1261605
- Vaulot, D., Lepere, C., Toulza, E., la Iglesia, R. De, Poulain, J., Gaboyer, F., Moreau, H., Vandepoele, K., Ulloa, O., Gavory, F., others, 2012. Metagenomes of the picoalga *Bathycoccus* from the Chile coastal upwelling. *PLoS One* 7, e39648–e39648.
- Viprey, M., Guillou, L., Ferréol, 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. doi:10.1111/j.1462-2920.2008.01602.x
- W
- Wang, K., Chen, F., 2004. Genetic diversity and population dynamics of cyanophage communities in the Chesapeake Bay. *Aquatic microbial ecology* 34, 105–116.
- Waterbury, J.B., Valois, F.W., 1993. Resistance to co-occurring phages enables marine *Synechococcus* communities to coexist with cyanophages abundant in seawater. *Appl. Environ. Microbiol.* 59, 3393–3399.
- Weinbauer, M.G., 2004. Ecology of prokaryotic viruses. *FEMS Microbiol. Rev.* 28, 127–181.

- 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.
- Weinbauer, M.G., Fuks, D., Puskaric, S., Peduzzi, P., 1995. Diel, seasonal, and depth-related variability of viruses and dissolved DNA in the Northern Adriatic Sea. *Microb. Ecol.* 30, 25–41.
- Weithoff, G., 2003. The concepts of “plant functional types” and “functional diversity” in lake phytoplankton – a new understanding of phytoplankton ecology? *Freshw. Biol.* 48, 1669–1675. doi:10.1046/j.1365-2427.2003.01116.x
- Weitz, J. S., Stock, C. A., Wilhelm, S. W., Bourouiba, L., Coleman, M. L., Buchan, A., *et al.* (2015). A multitrophic model to quantify the effects of marine viruses on microbial food webs and ecosystem processes, 1–13. <http://doi.org/10.1038/ismej.2014.220>
- Weitz, J.S., Poisot, T., Meyer, J.R., Flores, C.O., Valverde, S., Sullivan, M.B., 2013. Phage–bacteria infection networks. *Trends Microbiol* 21, 82–91.
- Wells, L.E., Deming, J.W., 2006. Effects of temperature, salinity and clay particles on inactivation and decay of cold-active marine Bacteriophage 9A. *Aquat. Microb. Ecol.* 45, 31–39.
- Weynberg, K.D., Allen, M.J., Ashelford, K., Scanlan, D.J., Wilson, W.H., 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.
- Wiens, J.J., 2011. The niche, biogeography and species interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences* 366, 2336–2350.
- Wilhelm, S.W., Matteson, A.R., 2008. Freshwater and marine virioplankton: a brief overview of commonalities and differences. *Freshw. Biol.* 53, 1076–1089.

- Wilhelm, S.W., Jeffrey, W.H., Suttle, C.A., Mitchell, D.L., 2002. Estimation of Biologically Damaging UV Levels in Marine Surface Waters with DNA and Viral Dosimeters. *Photochemistry and Photobiology* 76, 268–273. doi:10.1562/0031-8655(2002)076<0268:EOBDUL>2.0.CO;2
- Wilhelm, S.W., Suttle, C.A., 1999. Viruses and Nutrient Cycles in the Sea Viruses play critical roles in the structure and function of aquatic food webs. *BioScience* 49, 781–788. doi:10.2307/1313569
- Wilhelm, S.W., Weinbauer, M.G., Suttle, C.A., Jeffrey, W.H., 1998. The role of sunlight in the removal and repair of viruses in the sea. *Limnology and Oceanography* 43, 586–592.
- Wilke, M., 1999. Spatio-temporal dynamics of physico-chemical and chemical factors in the water of a heavily transformed Mediterranean coastal lagoon, the Etang de Salses-Leucate (France). *Vie et milieu* 49, 177–191.
- 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* 1–42.
- Winnepenninckx, B., Backeljau, T., De Wachter, R., 1994. Extraction of high molecular weight DNA from molluscs. *Trends Genet.* 407.
- Winter, C., Moeseneder, M.M., Herndl, G.J., Weinbauer, M.G., 2008. Relationship of geographic distance, depth, temperature, and viruses with prokaryotic communities in the eastern tropical Atlantic Ocean. *Microb. Ecol.* 56, 383–389.
- Wommack, K.E., Colwell, R.R., 2000. Virioplankton: viruses in aquatic ecosystems. *Microbiol. Mol. Biol. Rev.* 64, 69–114.
- Wommack, K.E., Ravel, J., Hill, R.T., Colwell, R.R., 1999. Hybridization analysis of Chesapeake Bay virioplankton. *Appl. Environ. Microbiol.* 65, 241–250.
- Worden, A.Z., 2006. Picoeukaryote diversity in coastal waters of the Pacific Ocean. *Aquat Microb. Ecol* 43, 165–175.

- Worden, A.Z., Lee, J.-H., Mock, T., Rouzé, P., Simmons, M.P., Aerts, A.L., Allen, A.E., Cuvelier, M.L., Derelle, E., Everett, M.V., 2009. Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes *Micromonas*. *Science* 324, 268–272.
- Worden, A.Z., Not, F., 2008. Ecology and Diversity of Picoeukaryotes, in: Kirchman, D.L. (Ed.), *Microbial Ecology of the Oceans*. John Wiley & Sons, Inc., pp. 159–205.
- Wu, W., Huang, B., Zhong, C., 2013. Photosynthetic picoeukaryote assemblages in the South China Sea from the Pearl River estuary to the SEATS station. *Aquat. Microb. Ecol.* 71, 271–284.
- Y
- Yutin, N., Wolf, Y.I., Raoult, D., Koonin, E.V., 2009. Eukaryotic large nucleo-cytoplasmic DNA viruses: clusters of orthologous genes and reconstruction of viral genome evolution. *Virol J* 6, 223.
- Z
- Zhu, F., Massana, R., Not, F., Marie, D., Vaulot, D., 2005. Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. *FEMS Microbiol. Ecol.* 52, 79–92. doi:10.1016/j.femsec.2004.10.006
- Zimmermann, J., Jahn, R., Gemeinholzer, B., 2011. Barcoding diatoms: evaluation of the V4 subregion on the 18S rRNA gene, including new primers and protocols. *Org. Divers. Evol.* 11, 173–192. doi:10.1007/s13127-011-0050-6
- Zingone, A., Natale, F., Biffali, E., Borra, M., Forlani, G., Sarno, D., 2006. Diversity in morphology, infectivity, molecular characteristics and induced host resistance between two viruses infecting *Micromonas pusilla*. *Aquat Microb. Ecol* 45, 1–14.
- Zingone, A., Sarno, D., Forlani, G., 1999. Seasonal dynamics in the abundance of *Micromonas pusilla* (Prasinophyceae) and its viruses in the Gulf of Naples (Mediterranean Sea). *J. Plankton Res.* 21, 2143–2159.ja

Annexes

Annexes

Annexe 1 : Détail sur le traitement et groupement des sequences de polymerase virale

Annexe 2 : Détail sur le traitement et groupement des sequences de *Major Capsid protein*

Annexe 3 : Détail sur le traitement et groupement des sequences de la region V4 du 18S

Annexe 4 : Détail sur le traitement et groupement des sequences de la region V9 du 18S

Annexe 5 : Développement d'une methode d'identification in silico des couples hôte-virus à partir de données de séquençage massif

Annexe 6 : *Curriculum vitae*

Annexe 7 : Participation à des congrès : Résumé de 2 congrès

Annexe 8 : Poster de médiation scientifique : atelier plancton

Annexe 9 : Poster de médiation scientifique : atelier virus

Annexe 1 : détail des traitement et groupement des séquences de polymérase virale

Station	Nombre de séquences total	Nombre de séquences après assemblage	Nombre de séquences jointes après nettoyage	Nombre d'OTU formés présents
LEU_A-MARS13	247 418	115 415	65 729	3 844
LEU_A-AVRIL13-A	478 796	225 555	160 436	4 573
LEU_A-AVRIL13-B	574 262	274 397	156 125	5 078
LEU_A-MAI13	153 942	71 643	45 660	3 278
LEU_A-JUIN13	403 168	192 657	139 674	3 831
LEU_A-JUILLET13	191 176	90 163	59 587	2 365
LEU_A-AOUT13	229 234	106 002	74 376	3 242
LEU_A-SEPTEMBRE13	280 138	113 295	60 373	3 240
LEU_A-OCTOBRE13	782 420	364 247	239 105	4 485
LEU_A-NOVEMBRE13	265 482	125 360	78 781	2 548
LEU_A-DECEMBRE13	296 030	140 326	92 320	2 435
LEU_A-JANVIER14	483 520	229 102	157 425	4 514
LEU_A-MARS14	409 464	194 482	137 245	4 308
LEU_A-AVRIL14	391 974	187 435	135 326	4 004
LEU_B-MARS13-A	198 992	92 463	51 811	3 386
LEU_B-MARS13-B	208 228	97 380	52 399	3 464
LEU_B-AVRIL13	342 042	161 765	115 774	3 873
LEU_B-MAI13	31 118	14 412	9 656	1 544
LEU_B-JUIN13	301 324	143 913	100 068	3 619
LEU_B-JUILLET13	275 256	128 225	83 128	3 050
LEU_B-AOUT13	235 138	106 838	69 135	3 177
LEU_B-OCTOBRE13	287 102	135 628	92 939	3 549
LEU_B-NOVEMBRE13	637 776	303 343	197 495	3 412
LEU_B-DECEMBRE13	722 838	343 622	221 771	3 577
LEU_B-JANVIER14	818 630	385 140	264 381	5 214
LEU_B-FEVRIER14	494 360	233 287	160 515	3 490
LEU_B-MARS14	541 026	254 659	170 324	5 440
LEU_B-AVRIL14	602 080	284 538	184 784	4 864
MOLA-MARS13	309 284	145 709	113 136	3 077
MOLA-AVRIL13-A	484 980	230 050	158 685	3 334
MOLA-AVRIL13-B	537 024	255 178	172 122	3 684
MOLA-JUIN13	32 178	14 881	6 022	1 211
MOLA-AOUT13	274 628	130 886	96 079	3 481
MOLA-DECEMBRE13	340 096	126 167	78 514	2 596
MOLA-MARS14	445 152	211 707	150 299	3 317
MOLA-AVRIL14	306 634	144 724	105 632	2 881
SOLA-MARS13	469 586	222 873	166 668	3 954
SOLA-AVRIL13-A	314 618	150 158	99 476	3 567
SOLA-AVRIL13-B	337 884	160 585	114 591	3 611
SOLA-MAI13-A	138 506	65 254	45 177	2 921
SOLA-MAI13-B	536 218	253 858	149 718	4 822
SOLA-JUIN13	673 060	316 265	204 649	5 259
SOLA-AOUT13	268 502	124 767	68 049	3 453
SOLA-NOVEMBRE13	238 904	98 249	51 555	3 016
SOLA-DECEMBRE13	229 302	95 670	60 533	2 601
SOLA-JANVIER14	369 318	176 791	120 221	3 065
TOTAL	17 188 808	8 039 064	5 337 468	91 120
Moyenne	373 670	174 762	116 032	3 549

Annexe 2 : détail des traitement et groupement des séquences de la MCP

Stations	Nombre de séquences total	Nombre de séquences après assemblage	Nombre de séquences jointes après nettoyage	Nombre d'OTU formés présents
LEU_A-MARS13-A	333 312	147 313	67 399	3 718
LEU_A-MARS13-B	368 986	168 686	76 165	3 942
LEU_A-AVRIL13-A	943 566	427 157	94 574	3 378
LEU_A-AVRIL13-B	411 370	187 194	195 485	4 716
LEU_A-MAI13	81 676	34 026	15 009	1 800
LEU_A-JUIN13	143 594	62 632	33 678	2 368
LEU_A-JUILLET13	100 864	45 149	25 140	1 839
LEU_A-AOUT13	126 206	48 336	18 339	2 343
LEU_A-SEPTEMBRE13	272 982	72 365	23 322	2 526
LEU_A-OCTOBRE13	278 352	116 076	56 681	2 366
LEU_A-NOVEMBRE13	531 366	240 613	110 731	3 921
LEU_A-DECEMBRE13	222 980	98 783	45 816	2 751
LEU_A-JANVIER14	591 390	260 990	122 142	4 868
LEU_A-MARS14	237 716	107 174	45 011	2 701
LEU_A-AVRIL14	412 258	179 850	85 905	3 297
LEU_B-MARS13-A	195 854	88 066	41 724	2 769
LEU_B-MARS13-B	447 520	201 855	93 031	4 034
LEU_B-AVRIL13	385 186	174 883	88 549	3 184
LEU_B-MAI13	51 642	20 541	9 438	1 377
LEU_B-JUIN13	189 904	84 768	44 684	2 610
LEU_B-JUILLET13	434 024	182 738	91 487	3 353
LEU_B-AOUT13	220 664	52 481	21 901	2 004
LEU_B-OCTOBRE13	345 744	153 337	75 865	3 949
LEU_B-NOVEMBRE13	570 662	258 047	122 827	3 969
LEU_B-DECEMBRE13	667 124	298 132	142 078	3 966
LEU_B-JANVIER14	574 938	253 763	121 501	4 363
LEU_B-FEVRIER14	575 572	262 542	149 119	3 876
LEU_B-MARS14	440 320	184 677	79 486	4 083
LEU_B-AVRIL14	669 052	299 756	168 704	4 819
MOLA-MARS13	342 496	154 370	82 412	3 283
MOLA-AVRIL13-A	570 892	259 831	156 985	4 086
MOLA-AVRIL13-B	631 770	285 521	146 704	3 963
MOLA-JUIN13	548 230	251 094	151 406	4 184
MOLA-DECEMBRE13	495 594	226 153	110 092	4 610
MOLA-MARS14	485 234	218 893	117 019	3 103
MOLA-AVRIL14	469 362	213 743	129 861	3 012
SOLA-MARS13	327 968	143 516	77 261	3 145
SOLA-AVRIL13-A	643 478	286 204	143 268	4 203
SOLA-AVRIL13-B	557 770	252 249	130 262	4 085
SOLA-MAI13-A	182 622	82 732	48 413	2 841
SOLA-MAI13-B	402 484	127 497	55 454	3 222
SOLA-JUIN13	435 270	194 541	100 023	3 988
SOLA-AOUT13	100 062	43 446	24 939	2 531
SOLA-NOVEMBRE13	206 812	60 545	20 513	2 201
SOLA-DECEMBRE13	293 230	79 258	24 154	1 700
SOLA-JANVIER14	644 204	290 564	164 371	3 952
TOTAL	18 162 302	7 882 087	3 948 928	12 795
Moyenne	394 833	171 350	85 846	3 326

Annexe 3 : détail des traitement et groupement des séquences de la région V4 du 18S

Station	Nombre de séquences total	Nombre de séquences après assemblage	Nombre de séquences jointes après nettoyage	Nombre d'OTU formés présents
LEU_A-AVRIL13	409 104	193 102	155 502	720
LEU_A-MAI13	478 140	224 764	174 305	850
LEU_A-JUIN13	903 752	423 146	338 008	2 395
LEU_A-JUILLET13	492 660	229 880	181 560	1 580
LEU_A-AOUT13	281 230	132 603	102 929	991
LEU_A-SEPTEMBRE13	494 116	233 317	183 235	1 570
LEU_A-OCTOBRE13-A	553 356	259 668	205 586	2 302
LEU_A-OCTOBRE13-B	241 540	112 935	85 993	997
LEU_A-NOVEMBRE13	335 022	157 428	115 927	1 498
LEU_A-JANVIER14	666 184	312 918	230 968	4 963
LEU_A-FEVRIER14	432 840	203 105	148 138	3 015
LEU_A-MARS14	350 382	164 243	113 774	3 122
LEU_A-AVRIL14-A	280 374	131 624	92 205	917
LEU_A-AVRIL14-B	403 664	188 411	131 093	2 780
LEU_B-MARS13	399 784	187 507	147 600	729
LEU_B-AVRIL13-A	430 712	202 619	155 473	1 187
LEU_B-MAI13	310 742	146 499	114 251	1 055
LEU_B-JUIN13	477 364	225 389	180 446	1 454
LEU_B-JUILLET13	436 540	204 006	156 372	1 255
LEU_B-AOUT13	294 268	138 428	99 890	921
LEU_B-SEPTEMBRE13	362 896	171 303	134 295	693
LEU_B-DECEMBRE13	462 076	218 543	173 555	1 956
LEU_B-JANVIER14-A	898 872	420 941	316 774	5 426
LEU_B-MARS14	282 582	132 364	91 369	2 148
LEU_B-AVRIL14	490 546	230 465	172 291	3 613
MOLA-AVRIL13	411 896	193 828	154 231	2 263
MOLA-JUIN13	506 188	238 873	190 769	1 339
MOLA-JUILLET13	758 792	355 852	280 016	2 348
MOLA-AOUT13	419 012	196 073	156 217	1 967
MOLA-SEPTEMBRE13	434 614	204 965	161 355	2 765
MOLA-OCTOBRE13	668 854	315 629	249 340	2 425
MOLA-DECEMBRE13-A	308 796	144 057	112 099	1 781
MOLA-DECEMBRE13-B	548 960	259 197	201 364	2 459
MOLA-AVRIL14	935 370	439 402	335 430	6 302
SOLA-AVRIL13	381 720	178 858	138 669	1 925
SOLA-MAI13-A	250 706	117 693	93 452	600
SOLA-MAI13-B	359 950	169 784	137 038	1 981
SOLA-JUIN13	509 146	239 571	196 744	1 932
SOLA-JUILLET13	591 542	277 964	216 291	2 341
SOLA-AOUT13	618 374	288 140	227 606	2 439
SOLA-SEPTEMBRE13	555 210	261 102	204 597	2 700
SOLA-OCTOBRE13	776 004	363 225	281 480	4 386
SOLA-NOVEMBRE13	394 224	185 316	148 245	1 691
SOLA-DECEMBRE13	708 378	334 832	263 003	3 096
SOLA-JANVIER14	593 596	277 725	209 015	4 945
SOLA-FEVRIER14	361 496	170 047	134 382	2 646
SOLA-MARS14	1 090 778	514 479	397 709	8 254
SOLA-AVRIL14	963 436	448 925	335 152	6 767
TOTAL	24 315 788	11 420 745	8 825 743	86 630
moyenne	506 579	237 932	183 870	2 448

Annexe 4 : détail des traitement et groupement des séquences de la région V9 du 18S

Station	Nombre de séquences total	Nombre de séquences après assemblage	Nombre de séquences jointes après nettoyage	Nombre d'OTU formés présents
LEU_A-AVRIL13	496 736	197 481	131 056	1 585
LEU_A-MAI13	549 884	220 894	150 617	1 360
LEU_A-JUIN13	624 672	241 366	198 822	3 088
LEU_A-JUILLET13	520 312	207 171	171 261	2 226
LEU_A-AOÛT13	301 766	118 590	80 781	1 735
LEU_A-SEPTEMBRE13	189 392	285 160	160 785	3 065
LEU_A-OCTOBRE13	741 828	74 669	44 736	535
LEU_A-DECEMBRE13	220 580	86 860	61 179	595
LEU_A-JANVIER14	366 902	139 203	99 264	1 935
LEU_A-FEVRIER14	426 910	160 475	97 301	1 840
LEU_A-MARS14	439 324	165 001	109 315	2 525
LEU_A-AVRIL14	317 744	121 006	87 862	2 020
LEU_B-MARS13	391 856	156 144	93 344	1 285
LEU_B-AVRIL13-A	390 096	152 793	117 924	2 147
LEU_B-AVRIL13-B	373 832	147 301	85 645	837
LEU_B-MAI13	344 394	137 833	75 682	1 573
LEU_B-JUIN13	482 202	187 080	130 494	1 863
LEU_B-JUILLET13	361 902	139 789	42 549	1 183
LEU_B-AOÛT13	362 898	144 020	112 108	1 568
LEU_B-SEPTEMBRE13	340 568	134 471	93 169	2 049
LEU_B-OCTOBRE13	408 992	161 218	83 569	2 257
LEU_B-NOVEMBRE13	372 108	147 785	107 261	2 414
LEU_B-DECEMBRE13	343 792	136 293	115 115	1 766
LEU_B-JANVIER14-A	567 694	213 658	119 036	1 867
LEU_B-JANVIER14-B	636 168	254 044	217 290	2 330
LEU_B-MARS14	596 660	232 812	171 620	2 983
LEU_B-AVRIL14	463 622	179 588	143 517	2 461
MOLA-AVRIL13	476 548	182 230	141 768	4 024
MOLA-JUIN13	597 310	239 161	194 798	2 740
MOLA-JUILLET13	617 384	236 503	173 769	2 876
MOLA-AOÛT13	428 236	165 315	121 374	3 214
MOLA-SEPTEMBRE13	584 428	222 097	146 513	4 944
MOLA-OCTOBRE13	407 792	151 939	108 659	3 261
MOLA-DECEMBRE13	255 768	101 271	68 796	901
MOLA-AVRIL14	469 984	177 969	116 961	2 733
SOLA-AVRIL13	454 386	175 707	130 713	3 478
SOLA-MAI13	716 336	286 646	232 013	2 503
SOLA-JUIN13	467 622	180 463	124 017	2 606
SOLA-JUILLET13	371 282	143 799	92 285	2 591
SOLA-AOÛT13	423 590	163 423	108 721	3 385
SOLA-SEPTEMBRE13	471 214	181 423	132 848	4 164
SOLA-OCTOBRE13	313 414	119 977	82 437	3 789
SOLA-NOVEMBRE13	636 960	243 932	187 655	4 642
SOLA-DECEMBRE13	419 562	159 397	125 366	3 032
SOLA-JANVIER14	482 338	184 358	149 223	3 347
SOLA-FEVRIER14	327 264	125 921	90 175	1 953
SOLA-MARS14	391 234	150 489	116 931	4 098
SOLA-AVRIL14	767 256	298 659	234 065	5 213
TOTAL	21 712 742	8 433 384	5 980 389	53 527
moyenne	452 349	175 696	124 591	2 512

Annexe 5 : Développement d'une méthode d'identification *in silico* des couples hôte-virus à partir de données de séquençage massif

Le fonctionnement des écosystèmes est étroitement lié à la structure des communautés qui les composent (Larsen *et al.*, 2005). La structure de ces communautés n'est pas aléatoire et est influencée par différents types d'associations biologiques (e.g. Weitz *et al.*, 2013 ; Weitz *et al.*, 2015) : parasitisme, prédation, compétition, mutualisme. La caractérisation et la compréhension des patrons d'associations sont importantes au niveau écologique puisqu'ils reflètent les processus de niche qui contrôlent la coexistence et la diversité biologique au sein des environnements.

Grâce aux techniques de séquençage massif, de grandes quantités de données sont aujourd'hui disponibles. En revanche, aucune méthode actuelle ne permet d'identifier au sein de ces séquences les couples hôte-parasite.

La structure co-phylogénétique mise en évidence entre les Prasinovirus et les *Mamiellophyceae* (Bellec *et al.*, 2009, Clerissi *et al.* 2012) fait de ce système hôte-virus un candidat idéal pour la mise en place d'une méthode d'assignation à partir de données environnementales.

Le Laboratoire de Biométrie et Biologie Evolutive de Lyon développe avec notre collaboration une méthode d'analyse basée sur le principe de co-inertie des données d'abondances relatives, obtenues lors de notre suivi mensuel des communautés.

Dans un premier temps, afin de valider la méthode d'assignation, il a fallu mettre en place une simulation des données d'abondance suivant différents types d'interactions (virus strictement spécifiques, virus à spectre large ou réduit, et virus à spectre aléatoire ; Figure 1)

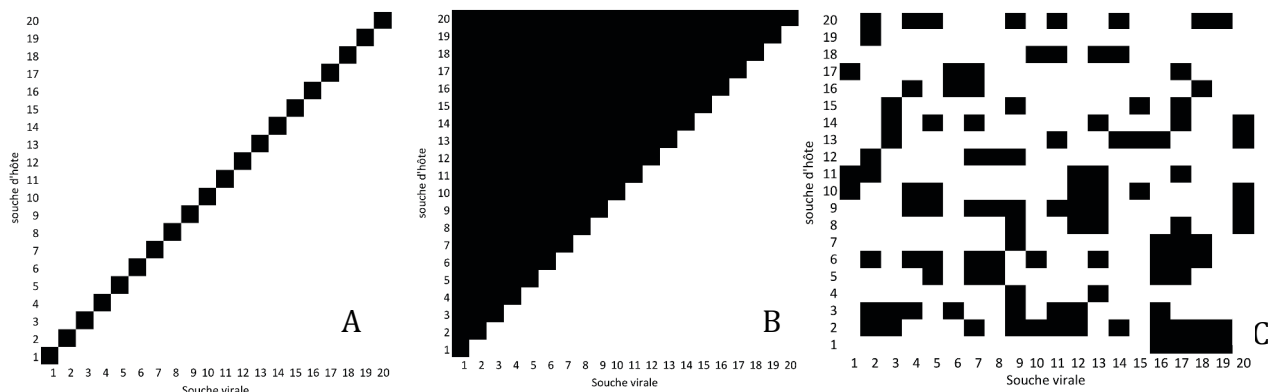


Figure 1 : Matrice d'interactions entre souches de virus et d'hôtes. A : Souches spécifiques, B : Souches généralistes ou spécialistes, C : Associations aléatoires.

Un jeu de données correspondant aux hôtes est simulé suivant une réponse unimodale face à un gradient environnemental (gradient trophique). Suite à ces simulations, l'abondance virale est inférée à partir d'une régression multiple incluant les matrices d'interactions permettant de spécifier les associations entre les hôtes et les virus et ainsi calculer des tableaux de contingence virale. Ces deux tableaux de contingence sont ainsi utilisables afin de tester la méthode.

La méthode en elle-même est une analyse de co-inertie sur des résultats d'analyses factorielles des correspondances (AFC) effectuées sur les tableaux de contingence d'hôtes et de virus. L'AFC permet de décrire les relations entre les stations d'échantillonnage et les abondances relatives de chaque entité. Lors d'une AFC, l'ordination simultanée des entités et des sites permet de placer les entités à la moyenne des sites qu'elles occupent ainsi que de placer les sites à la moyenne des entités qu'ils contiennent. Cependant, dans une AFC, les sites possèdent des poids qui leur sont propres à l'échelle du tableau de contingence utilisé. Cet aspect est problématique lors d'une analyse de co-inertie car cette dernière requiert normalement un poids de site commun entre les tableaux utilisés.

L'originalité de cette approche en développement est d'adapter une analyse de co-inertie pour le couplage de 2 AFC. Dans ce cas, l'analyse de co-inertie utilise des poids qui maximisent la covariance entre les sites des deux tableaux d'AFC.

Cette analyse est décrite sous le terme d'analyse de correspondance. La significativité des co-structures entre les sites observées dans l'analyse de co-inertie est estimée à l'aide de tests par permutations.

De plus, dans la mesure où un seul échantillon est prélevé par date d'échantillonnage, nous testons la stabilité des relations hôtes-virus dans l'analyse de co-inertie par une méthode de bootstrap, en déterminant si les positions des hôtes et des virus dans l'espace de co-inertie sont robustes par rapport à des perturbations statistiques du jeu de données original. Suite à ces étapes de vérifications les associations hôte-virus seront identifiées en calculant la distance entre les points hôtes et les points virus sur le plan de la co-inertie. Cette distance est déterminée au travers de la distance du χ^2 à laquelle nous appliquons un test par permutations pour tester la significativité des associations observées.

Annexe 6 : *Curriculum vitae*

Informations personnelles

Prénom, Nom : Hugo Lebredonchel

Adresse : 16, rue du 14 juillet

66650, Banyuls-sur-Mer

Téléphone / Adresse e-mail : +33 (0)6 65 37 13 87 / Lebredonchel.hugo@gmail.com

Statut actuel : Doctorant à l'Observatoire Océanologique de Banyuls-sur-Mer, UPMC, UMR 7232 - Biologie Intégrative des Organismes Marins (BIOM)

Cursus

2012-2016 **Doctorat « Diversité des prasinovirus et contrôle par les facteurs environnementaux »** Université Pierre et Marie Curie, Observatoire Océanologique de Banyuls-sur-Mer.

2010-2012 **Master Sciences de l'Univers, environnement, écologie, spécialité Océanographie et Environnements Marins** à l'Université Pierre et Marie Curie, parcours environnement côtier.

2007-2010 **Licence Biologie, Spécialité biologie des organismes** à l'Université de Caen Basse-Normandie.

Expérience professionnelle

- 2012** **Stage de Master 2** à la Station Biologique de Roscoff : Caractérisation de prasinovirus infectant la picoalgue verte marine *Micromonas sp.*, Encadrants : Nathalie Simon et Anne-Claire Baudoux (5 mois). Roscoff, France.
- 2011** **Stage de Master 1** au CNRS INEE - FRE3484 BioMEA : Actualisation des données de taille et d'âge à maturité de Seiche en Manche en rapport avec le réchauffement climatique. Encadrants : Jean-Paul Robin et Michael Gras (2 mois). Université de Caen Basse-Normandie, France.
- 2010** **Stage volontaire** au CNRS-FRE3484 BioMEA : Impact de l'environnement sur le développement de la seiche *Sepia* (projet CRESH). Encadrants : Jean-Paul Robin et Georges Safi (2 mois). Université de Caen Basse-Normandie, France.
- 2009** **Stage volontaire** au CNRS-FRE3484 BioMEA : Biodiversité des microalgues sur l'estran. Encadrant : Benoit Veron (Mission Ponctuelle). Université de Caen Basse-Normandie, France.

Publications scientifiques

H. Lebredonchel, M. Ros, N. Grimsley, Y. Desdevises. Green microalgae of the Mamiellophyceae dominate picoeukaryotic communities throughout the year in a coastal lagoon revealed by qPCR.

H. Lebredonchel, B. Alric, E. Derelle, N. Grimsley, S. Dray, Y. Desdevises. Impact of environmental factor on Prasinovirus diversity and distribution. En préparation.

B. Alric, **H. Lebredonchel**, Y. Desdevises, S. Dray. Statistical identification of the couples host-virus in massive sequencing dataset. En préparation.

A-C. Baudoux, **H. Lebredonchel**, H. Dehmer, M. Latimier, R. Edern, F. Rigaut-Jalabert, P. Ge, L. Guillou, E. Foulon, Y. Bozec, T. Cariou, Y. Desdevises, E. Derelle, N. Grimsley, H. Moreau, N. Simon. 2015. Interplay between the genetic clades of *Micromonas* and their viruses in the Western English Channel. Environmental Microbiology Report. DOI: 10.1111/1758-2229.12309

M. Gras, G. Safi, **H. Lebredonchel**, J. Quinquis, E. Foucher, N. Koueta, J-P. Robin. 2015. Stock structure of the English Channel common cuttlefish *Sepia officinalis* (Linnaeus, 1758) during the reproduction period. Journal of the Marine Biological Association of the UK. DOI: 10.1017/S0025315415001162

Congrès

Avr. 2015 **Rencontre Mer LR.** Présentation orale : « The impact of environmental variables on prasinovirus-host diversity and abundance measured by monthly metagenomic analyses over one year », Montpellier, France

Fev. 2015 **Association for the Sciences of Limnology & Oceanography (ASLO) :** Présentation orale : « the impact of environmental variables on

Prasinovirus-host diversity and abundance measured by monthly genomic analyses over one year ». Grenade, Espagne.

Nov. 2013 **Aquatic Virus Workshop 7 (AVW7)**. Présentation orale : « Temporal analysis of prasinovirus-host interactions in a trophic gradient ». St Petersburg, Floride, USA.

Juin. 2011 **Congrès de l'Association Française d'Halieutiques**. Poster : « Structure du stock de seiche, *Sepia officinalis*, en Manche en période de reproduction ». Boulogne-sur-Mer, France.

Formation Scientifique

Mai 2013 **Marine Ecological and Evolutionary Genomics, Marine Genomics for User (MG4U)**. (40 h). Bremen, Allemagne.

Avr. 2013 **Adaptation et phylogénie (NU950)**. (60 h), Master SDUEE/SEP. Observatoire Océanologique de Banyuls-sur-Mer, France.

Mars 2013 **Daniel Jouvance - UPMC International School in Marine Biology and Ecology**. (54 h). Rio de Janeiro, Brazil.

Enseignement

Mars 2013 **Travaux pratiques isolement et caractérisation des virus marins** (6 h). UPMC International School in Marine Biology and Ecology. Rio de Janeiro, Brazil.

Juillet 2012 **Moniteur pour les enseignements, Biodiversité marine : expertise FLORE, 4UM52**. Station Biologique de Roscoff, France.

Encadrement

Jan.-Fev. 2014 Anastasia Tsiola, étudiante EMBRC (Grèce), « étude des communautés virales naturelles », encadrement avec Nigel Grimsley. Observatoire Océanologique de Banyuls-sur-Mer, France.

Mar.-Juin 2014 Mickaël Ros Master 1, Université de Perpignan, « Développement d'un protocole de PCR quantitative pour estimer l'abondance de picoeucaryotes de la classe des Mamiellophyceae (Chlorophyta) », encadrement avec Yves Desdevises. Observatoire Océanologique de Banyuls-sur-Mer, France.

Animation scientifique

2013-2015 Mission doctorale de médiation scientifique. BIODIVERSARIUM. Conception d'ateliers et de contenus muséographiques, animations d'atelier. Observatoire Océanologique de Banyuls-sur-Mer, France.

Oct. 2014 Atelier « Les virus marins ». Fête de la science. Observatoire Océanologique de Banyuls-sur-Mer, France.

Oct. 2014 Atelier « Au Cœur des Océans ». Fête de la science. Observatoire Océanologique de Banyuls-sur-Mer, France.

Oct. 2013 Atelier « Les virus Marins ». Fête de la science. Observatoire Océanologique de Banyuls-sur-Mer, France.

Responsabilités

2014 - 2016 Représentant étudiant au Conseil d'Administration de l'Observatoire Océanologique de Banyuls-sur-Mer, France.

2014 - 2016 Représentant étudiant au Conseil Hygiène et Sécurité de l'Observatoire Océanologique de Banyuls-sur-Mer, France.

Expérience sur le terrain

2013 - 2014 Echantillonnage mensuel à bord du bateau scientifique Néréis II. Observatoire Océanologique de Banyuls-sur-Mer, France.

2013 - 2014 Echantillonnage mensuel dans la lagune de Salse-Leucate. Observatoire Océanologique de Banyuls-sur-Mer, France.

Avril 2012 Echantillonnage ponctuel à bord du bateau scientifique Neomysis. Station Biologique de Roscoff, France

Mars 2011 Prélèvement et dissection de céphalopode, Station IFREMER de Port en Bessin, France.

2010 Trois évaluations du stock de céphalopode à partir des débarques en criée. Port en Bessin, France

Mai 2009 Echantillonnage sur estran pour l'évaluation de la biodiversité algale le long d'un transect. Université de Caen Basse-Normandie, France.

Compétences techniques et personnelles
--

Langue : Français, Anglais

Plongé sous-marine : Niveau 2, CMAS

Techniques : Culture d'algue, culture et isolation de virus, extraction et purification d'ADN, PCR, qPCR, microscopie électronique, PFGE, purification de génome, dosage d'oxygène (méthode de Winkler), cytométrie en flux, tri cellulaire, purification par ultracentrifugation, SDS-Page.

Informatique : Script Unix, suite Microsoft office, GIMP, Inkscape, Illustrator, Photoshop, R, XLstat, MrBayes, Seaview, Bioedit, Mega, Geneious, Usearch, Uchime, Mothur, Logiciel de cytométrie Diva6.

Annexe 7 : Congrès

Congrès #1 : Aquatic Virus Workshop 7 (AVW7)**Date et Lieu** : 3-7 Novembre 2013, Saint Petersburg, Floride, USA**Contribution** : Présentation Orale**Résumé** : Temporal analysis of prasinovirus-host interactions in a trophic gradient

H. Lebretonchel*, N. Grimsley, H. Moreau, C. Hemon, Y. Desdevises

Univ. Pierre et Marie Curie (UPMC) and Centre National de la Recherche Scientifique (CNRS), Banyuls-sur-Mer, France

Viruses are known to be widespread and abundant in the marine environment, but how they are controlled by environmental factors, such as temperature or nutrients, is still poorly understood. Photosynthetic picoeukaryotes from the class Mamiellophyceae play an important role in marine ecosystems, and are regulated in part by Prasinoviruses. Several complete genome sequences are available for both prasinoviruses infecting their host genera *Micromonas*, *Bathycoccus* and *Ostreococcus*, and for their host algae themselves. Recently, we isolated a new Prasinovirus infecting *Mantoniella squamata*. We are currently monitoring three contrasted environments in the Gulf of Lion (NW Mediterranean sea), where we are measuring 18 environmental descriptors during one year, as well as investigating the diversity and abundance of viruses and photosynthetic picoeukaryotes. Virus and putative host diversity will subsequently be assessed using new generation sequencing of PCR-amplified partial DNA polymerase and major capsid protein markers (for viruses) and 18SrDNA sequences (for hosts). This molecular dataset will be related to environmental factors using numerical techniques to investigate how viral assemblages are controlled by the environment.

Congrès #2 : ASLO Aquatic Sciences Meeting**Date et Lieu :** 22-27 Février 2015, Granada, Espagne**Contribution :** Présentation Orale**Résumé :** The impact of environmental variables on prasinovirus-host diversity and abundance measured by monthly genomic analyses over one year

H. Lebretonchel*, N. Grimsley, H. Moreau, Y. Desdevises

Univ. Pierre et Marie Curie (UPMC) and Centre National de la Recherche Scientifique (CNRS), Banyuls-sur-Mer, France

Viruses are known to be widespread and abundant in the oceans, but while the influence of environmental factors such as temperature or concentrations of nutrients are known to influence viral assemblages, detailed studies on this are lacking. This study focuses on Prasinovirus that infect an important class of photosynthetic picoeukaryotes, the Mamiellophyceae. Several prasinovirus strains are already well characterized. The complete genomes of both prasinoviruses and their host algae have been characterized in the genera *Micromonas*, *Bathycoccus* and *Ostreococcus*. We monitored prasinoviruses and their hosts in three contrasted environments in the Gulf of Lion (Northwestern Mediterranean sea) and measured 18 environmental descriptors monthly for one year at the 4 sampling sites. Viruses were isolated using iron-chloride flocculation and their putative picoeukaryote hosts were collected by flow cytometry coupled with cell sorting. We are assessing viral and microalgae diversities using Illumina sequencing, based on partial DNA polymerase and major capsid protein sequences for viruses, and 18S rDNA sequences for picoeukaryotes. This dataset was used to explore the control of viral diversity and abundance by environmental factors.





AU CŒUR DES OCÉANS

PLONGEZ AU SEIN DE L'INFINIMENT PETIT

Le plancton

Du grec *planktós*, ce terme désigne les organismes aquatiques en suspension, inaptes à lutter contre le courant. Ce groupe est un ensemble très diversifié, comprenant à la fois des animaux ou **Zooplancton** tel que les méduses, des petits crustacés ainsi que des larves, des végétaux ou **Phytoplancton** représentés largement par des microalgues, des bactéries ou **bactérioplancton** et des virus ou **viroplancton**.



Le phytoplancton

Base essentielle des chaînes alimentaires, il constitue à lui seul 50% de la matière organique produite sur terre. Il est aussi considéré comme l'une des plus **importantes sources d'oxygène** de la planète. Leur impact est tel, qu'après leur mort, la sédimentation des algues possédant un squelette calcaire (les Coccolithophoridés) est capable de former de véritables falaises. La plus petite algue connue s'appelle *Ostreococcus tauri*, elle mesure 1µm, et fut isolée dans l'étang de Thau (Languedoc-Roussillon).

Les virus marins

Depuis leur découverte dans les années 1950, de nombreuses études ont révélé **l'extrême abondance et diversité** de ces entités biologiques parmi les plus infimes. Tous les organismes ont potentiellement un virus qui leur est associé. Les virus sont des **parasites obligatoires**, composés d'ADN ou ARN qui est protégé dans une enveloppe : la capsid. Après s'être développés au sein de la cellule hôte, ils sont libérés dans le milieu par un **mécanisme de lyse**, ayant pour conséquence la destruction de la cellule.

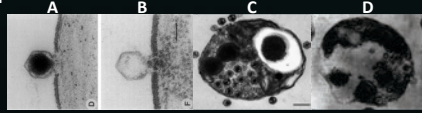
Phycodnaviridés

(Virus d'algue eucaryote)



Réplication d'un virus algue

A : Adsorption
B : Injection du génome
C : Réplication
D : Lyse



Virus géant

(aussi gros qu'une bactérie)



Virophage

(virus de virus)



Bactériophages

(Virus de bactérie)



Le saviez vous ?

Les virus sont les entités biologiques les plus abondantes sur terre. Il y a plus de virus dans les océans que d'étoiles dans notre galaxie.

Hugo Lebedonchel
Équipe GENOPHY - UMR7232 – Laboratoire Océanologique de Banyuls sur Mer
Université Pierre et Marie Curie - CNRS



LES VIRUS MARINS

La découverte des virus

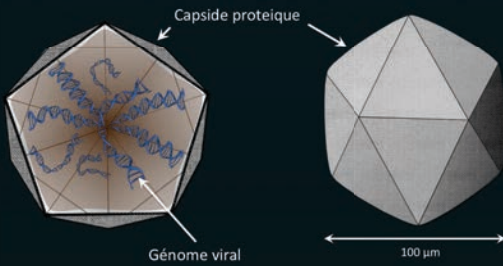
Le terme « Virus » provient du latin « **Poison** ». Ces entités affectent les organismes depuis la nuit des temps. On trouve des traces de l'impact des virus sur l'homme dans les hiéroglyphes égyptiens (3700 AC), où la **Poliomyélite** (virus de la polio) est représenté. Mais il faut attendre la fin du XVIII^{ème} siècle, pour que les virus soient étudiés et décrits. Le **virus de la mosaïque du tabac** fut le premier.



Mais concrètement : Qu'est ce qu'un virus ?

Les virus sont des entités de **très petite taille**, 1/10^{ème} de millimètre. Ils possèdent une enveloppe, appelée « **capside protéique** ». Cette « boîte » contient et protège le **génom viral** (pouvant être de l'ADN ou de l'ARN).

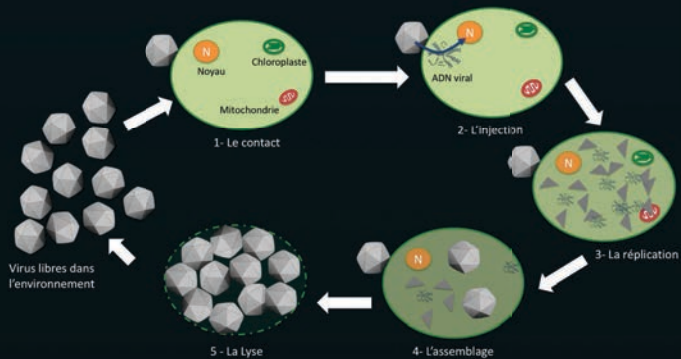
Ce sont des parasites **obligatoires**, ils ont besoin d'une cellule hôte pour leur **réplication**. Les virus sont en général **spécifiques** d'un « type » cellulaire. La transmission entre les espèces est extrêmement limitée.



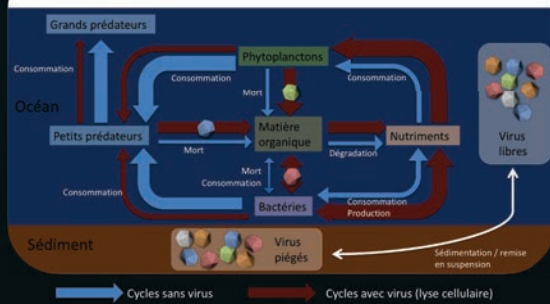
La réplication des virus par la lyse

5 étapes sont nécessaires pour leur réplication :

- 1 - **Le contact** : Les virus ne peuvent pas se déplacer, l'hôte doit être « au bon endroit, au bon moment ».
- 2 - **L'injection** : le génome viral entre dans la cellule.
- 3 - **La réplication** : la machinerie cellulaire de l'hôte est détournée pour produire des parties du virus.
- 4 - **L'assemblage** : les différentes parties s'assemblent pour créer des virus.
- 5 - **La lyse** : l'enveloppe de la cellule est détruite pour relâcher les virus.



Impact des virus sur les cycles trophiques



Hugo Lebretonchel
Équipe GENOPHY - UMR7232 – Laboratoire Océanologique de Banyuls sur Mer
Université Pierre et Marie Curie - CNRS

Les virus dans l'océan

Ce sont les entités biologiques **les plus abondantes** (1 à 100 millions de virus par millilitre). La grande majorité de ces virus infectent des **organismes unicellulaires** comme les bactéries et le phytoplancton. Ils jouent un rôle important dans la régulation des populations planctoniques. Les virus peuvent être piégés longtemps dans les sédiments.

Ce sédiment peut être remis en suspension, relâchant les virus piégés, pouvant ainsi de **nouveau se répliquer**. Ce gigantesque réservoir viral constitue pour les scientifiques un important stock de gènes pouvant avoir **des applications révolutionnaires** (médecine, OGM,...).

Résumé - Abstract

Les virus sont les entités les plus abondantes dans les océans et ces parasitoïdes interagissent avec chaque composante de la biosphère marine, soit par la sélection des communautés d'hôte, soit en influençant les cycles biogéochimiques. Cependant, l'impact des conditions environnementales sur ces assemblages viraux complexes est encore mal compris. L'objectif de ce travail est de comprendre comment les assemblages viraux sont influencés par les facteurs environnementaux. Durant un suivi mensuel des communautés de mars 2013 à avril 2014, nous nous sommes intéressés à un système hôte-virus abondant et largement répandu, les communautés de microalgues Mamiellophyceae et leurs *Prasinovirus*. Durant ces études, des approches de PCR quantitative et de séquençage massif ont été développées afin de décrire les populations environnementales présentes au sein du Golfe du Lion. Les populations de Mamiellophyceae dominent le compartiment des picoeucaryotes photosynthétiques, avec des dynamiques fortes au cours de l'année. Les communautés de *Prasinovirus* reflètent les dynamiques de leurs hôtes, elles-mêmes en lien avec les facteurs environnementaux. En revanche, la dispersion ainsi que le maintien des virions semblent directement impacté par les conditions environnementales, comme la température et l'hydrodynamisme. Certains groupes viraux, comme les virus d'*Ostreococcus*, sont inféodés aux lagunes et aux zones côtières, suggérant ainsi que la dispersion des populations virales est limitée par la présence des hôtes.

Mot-clés : Virus, Mamiellophyceae, distribution, dynamiques, environnements

Viruses are the most abundant biological entities in oceans and they every member of the marine biosphere is affected by them, they influence the composition of communities and influence biogeochemical cycles. However, the influence of environmental conditions on complex viral populations is still poorly understood. The aim of this study is to understand how environmental factors influence viral communities. We followed viral communities monthly from March 2013 to April 2014 and investigated the host-virus system of Mamiellophyceae communities and their prasinoviruses, a model sytem that is abundant and widespread. *Prasinovirus*-host populations from the Gulf of Lion were quantified by PCR amplification and analysis of genetic marker genes and high throughput sequencing. Photosynthetic populations of picoeukaryotes were dominated by populations of Mamiellophyceae showing high levels of quantitative and qualitative annual variations that were related to environmental factors. *Prasinovirus* communities mimed host dynamics, but viral dispersion and persistence appeared to be impacted directly by environmental conditions such as temperature and hydrodynamics. Several viral groups, such as *Ostreococcus* viruses, were specific to lagoonal and coastal areas, suggesting that their dispersal is limited by host occurrence.

Key words : Viruses, Mamiellophyceae, Distribution, Dynamics, Environments.