



**HAL**  
open science

# Diversité et structure spatio-temporelle des communautés protistes dans deux systèmes côtiers aux conditions trophiques contrastées : cas de la Manche Orientale et la Méditerranée Occidentale

Sara Rachik

► **To cite this version:**

Sara Rachik. Diversité et structure spatio-temporelle des communautés protistes dans deux systèmes côtiers aux conditions trophiques contrastées : cas de la Manche Orientale et la Méditerranée Occidentale. Ecosystèmes. Université du Littoral Côte d'Opale, 2018. Français. NNT : 2018DUNK0476 . tel-02178690v1

**HAL Id: tel-02178690**

**<https://theses.hal.science/tel-02178690v1>**

Submitted on 10 Jul 2019 (v1), last revised 23 Sep 2019 (v2)

**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.

**THESE DE DOCTORAT DE L'UNIVERSITE DU LITTORAL COTE D'OPALE**

**PRESENTEE ET SOUTENUE PAR**

**SARA RACHIK**

**POUR L'OBTENTION DU GRADE DE DOCTEUR DE L'UNIVERSITE DU  
LITTORAL COTE D'OPALE**

DOCTORAT : SCIENCES DE LA MATIERE, DU RAYONNEMENT ET DE  
L'ENVIRONNEMENT

DISCIPLINE : GEOLOGIE, ECOLOGIE, PALEONTOLOGIE, OCEANOLOGIE

---

**DIVERSITE ET STRUCTURE SPATIO-TEMPORELLE DES COMMUNAUTES  
PROTISTES DANS DEUX SYSTEMES COTIERS AUX CONDITIONS  
TROPHIQUES CONTRASTEES (*Cas de la Manche Orientale et la Méditerranée  
Occidentale*)**

---

SOUTENUE LE 08 FEVRIER 2018

Devant un jury composé de :

**Dr. Ingrid Obernosterer** (Dr CNRS, Laboratoire d'Océanographie Microbienne (LOMIC))  
**Rapporteuse**

**Dr. Ludwig Jardillier** (MCF HDR, Laboratoire Ecologie Systématique Evolution (ESE))  
**Rapporteur**

**Dr. Téléphore Sime-Ngando** (DR CNRS, Laboratoire Microorganismes: Génome et environnement (LMGE))  
**Examineur**

**Pr. Rachid Amara** (PR ULCO, Laboratoire d'Océanologie et de Géosciences (LOG))  
**Président du jury**

**Pr. Urania Christaki** (PR ULCO, Laboratoire d'Océanologie et de Géosciences (LOG))  
**Co-directrice de thèse**

**Pr. Sébastien Monchy** (PR ULCO, Laboratoire d'Océanologie et de Géosciences (LOG))  
**Directeur de thèse**





## AVANT PROPOS

Cette thèse a été réalisée grâce à l'allocation de recherche qui m'a été accordée à 50% par le Syndicat Mixte de la Côte d'Opale (SMCO) et à 50% par la région Hauts de France (anciennement nommée région Nord-Pas De Calais).

Ce projet de thèse est rattaché aux programmes:

ANR-ROME,

Nord-Pas-de Calais Région et projet FRB-Demo,

SOMLIT Service d'Observation en Milieu LITtoral (pour l'échantillonnage)

SPEciMed (Structures of Planktonic Ecosystems in the North-western Mediterranean)





I'm grateful to have people in my life who care about  
my growth as if it's part of their own.

I'm thankful to be loved and to be able to love.

You've always believed in me,

THANK YOU!

-----

J'aimerais tout d'abord remercier les membres du jury, Ingrid Obernosterer, Ludwig Jardillier, Télésphore Sime-Ngando et Rachid Amara d'avoir accepté d'évaluer mon travail.

Merci à M. François Schmitt, directeur du LOG, de m'avoir accueillie parmi eux et pour sa disponibilité.

Merci au comité de thèse, Laurent Seuront et les différents membres du jury pour leurs précieux conseils à mi-parcours,

Merci, merci, merci et encore merci à mon directeur de thèse Sébastien MONCHY et ma Co-directrice de thèse Urania CHRISTAKI pour TOUT, quoique je dise je ne peux exprimer ma reconnaissance. Vous étiez toujours là à m'épauler pendant mes pires moments. Vous m'aviez soutenu jusqu'au bout et je ne peux dire que juste j'étais honorée de travailler avec vous. Je vous remercie de m'avoir accueillie parmi vous, merci d'avoir toujours été présents pour me guider et m'aider à avancer ... Merci pour et pour ... Vraiment la liste est longue !

Un grand merci pour Elsa Breton, qui m'a toujours soutenu et qui m'a fait aimer davantage les statistiques (à point que je n'arrivais plus à m'arrêter☺).

Merci à Felipe Artigas pour son encadrement et ses conseils précieux pendant mes premiers pas en tant que « prof » lors de mon monitorat.

Merci à Eric Lecuyer et à l'équipe du Sépia II, qui ont rendu mes deux ans d'échantillonnage possibles et inoubliables (surtout quand il faisait moche et froid☺).

Merci à Sophie Merlin de GENES DIFFUSION de m'avoir accueillie dans leurs locaux, m'initier sur les amplifications et de porter réponse à toutes mes questions. Un merci à toute l'équipe pour leur bon accueil et la sacrée ambiance☺.

Merci à tous mes collègues du LOG de m'avoir accueillie parmi eux pendant ces trois ans de thèse.

Merci ensuite à mes 3 co-bureaux, Luen, Savvas et Solène pour les moments de détente et pour tout ce que j'ai pu apprendre avec eux pendant nos échanges.

Merci à mes collègues et amis qui ont rendu mes trois ans de thèse animés : Maria, Rym, Jess, Marie Anaïs, Corine, Mohamed, Kien, Date, Adam, Sundara, Benoit, Arnaud, Théo, Lionel, Khalef, Mahmoud, Mamadou, Ali, Ryma ...

Un merci spécial à ceux qui ont supporté mes sauts d'humeur et qui m'ont encouragé jusqu'au bout, à Hicham, Laury, Atallah, Waed, Amina, Nidale, Asmae, Firdaous, Sakina, Amine ... (La liste est bien longue, personne n'était à l'abri☺).

Les derniers remerciements, je les dois bien évidemment à ma famille. A mes parents qui m'ont soutenu financièrement et moralement depuis le début de mes études, à ma sœur et mon frère, mes deux âmes sœurs, à ma grand mère que sans elle je ne serai pas arriver là où j'en suis aujourd'hui, tout le mérite lui revient. A mes tantes, mes oncles et mes cousins et toute personne qui a participé de loin ou de près à l'aboutissement de ce travail.





You'll always be with me,  
Like a handprint on my heart...

A la mémoire de mon grand père  
Que ton âme repose en paix !



## RESUME

### **TITRE : DIVERSITE ET STRUCTURE SPATIO-TEMPORELLE DES COMMUNAUTES PROTISTES DANS DEUX SYSTEMES COTIERS AUX CONDITIONS TROPHIQUES CONTRASTEES (*Cas de la Manche Orientale et la Méditerranée Occidentale*)**

Dans les écosystèmes marins, les protistes jouent un rôle important en tant que producteurs primaires, prédateurs ou encore symbiontes. De par leur position clef, l'étude de leur diversité est fondamentale afin d'améliorer notre compréhension du fonctionnement des écosystèmes marins. Pour ce faire, des approches moléculaires couplées à la microscopie ont été utilisées afin d'examiner en profondeur la diversité des communautés microbiennes dans deux écosystèmes contrastés (Manche Orientale et Méditerranée). La première étude conduite en Manche Orientale a permis d'établir une vue d'ensemble de la diversité des protistes eucaryotes sur une période de 4 ans (de Mars 2011 à Juillet 2015) et d'examiner l'influence des variables environnementales sur leur distribution temporelle. L'approche novatrice utilisée consistait à séquencer simultanément l'ARNr et l'ADNr. Ce double ciblage a permis de calculer le rapport ARNr :ADNr et de démontrer qu'il pouvait être un paramètre utile et significatif pour évaluer l'activité cellulaire globale de la communauté microbienne. Nous avons également montré la faisabilité d'utiliser ce ratio ARNr :ADNr comme indicateur des transitions écologiques d'espèces microbiennes emblématiques (e.g. bloom de *Phaeocystis globosa*). Ainsi, le calcul individuel de ce rapport pour chaque OTU devrait permettre d'apporter une information supplémentaire indispensable pour une meilleure compréhension du fonctionnement de l'écosystème et de l'influence des paramètres biotiques et abiotiques sur la structuration des communautés microbiennes, en particulier dans un contexte de changement global. La seconde étude a également été menée en Manche Orientale sur les données de séquençage obtenues pendant la période 2011-2013. Elle avait pour objectif de mettre en évidence les relations existantes entre les parasites eucaryotes et les autres taxons ainsi que les paramètres environnementaux. Cette étude a montré la diversité insoupçonnée des symbiontes/décomposeurs en Manche Orientale, et leur importance pour la structuration des communautés et la succession saisonnière. En particulier, des analyses en réseaux de corrélations ont montré la prédominance des relations inter-taxa, plutôt qu'entre OTUs et paramètres abiotiques, et la position centrale des Symbiontes/décomposeurs dans ces relations. Cette étude a permis de souligner la complexité et l'importance des interactions microbiennes dans la structuration des communautés tout en apportant des informations cruciales permettant de mieux comprendre les mécanismes sous-jacents qui régissent cette structuration. La dernière étude a été menée dans la Méditerranée Occidentale, au niveau de quatre stations situées dans le Golfe du Lion, soumises à l'influence du fleuve Rhône, sur une période de deux ans (2012-2014). Elle avait pour objectif de réaliser la liste la plus exhaustive possible de la diversité des protistes eucaryotes et d'évaluer leur variabilité spatiale en relation avec les variables environnementales. Cette étude a confirmé l'effet prépondérant du Rhône sur la structuration des communautés, particulièrement dans la partie Est du Golfe du Lion. Globalement, ces travaux de thèse ont permis de renforcer, par le biais d'approches combinées de microscopie et de biologie moléculaire, nos connaissances sur le fonctionnement, la diversité taxonomique et la succession des espèces microbiennes, en relation avec les paramètres environnementaux, de deux écosystèmes marins contrastés.

## **ABSTRACT**

### **TITLE : DIVERSITY AND SPATIO-TEMPORAL STRUCTURE OF PROTIST COMMUNITIES IN TWO COASTAL SYSTEMS WITH CONTRASTED TROPHIC CONDITIONS (Case of the Eastern English Channel and the Western Mediterranean)**

In marine ecosystems, protists play an important role as primary producers, predators or symbionts. Therefore studying their diversity and metabolic activity is fundamental for understanding the functioning of marine ecosystems. In order to achieve this goal, molecular approaches were coupled with microscopy analysis for an in depth examination of microbial communities diversity in two contrasting ecosystems (Eastern English Channel (EEC) and Mediterranean). The first study, conducted in the EEC, allowed to establish a precise overview of eukaryotic protists diversity over 4 years (from March 2011 to July 2015) and to examine the influence of environmental variables on their temporal distribution. The innovative approach used during this study was to sequence simultaneously rRNA and rDNA. This double targeting allowed calculating the rRNA: rDNA ratio and demonstrate that it could be an useful and significant parameter for measuring the relative cellular activity of the microbial community. We also showed the feasibility of using this rRNA: rDNA ratio as an indicator of ecological transitions of iconic microbial species (e.g. bloom of *Phaeocystis globosa*). Thus, calculation of this ratio for each individual OTU provided additional information that are essential for a better understanding of the functioning of the ecosystem and the influence of biotic and abiotic parameters on the structuring of microbial communities, in particular in the context of global change. The second study was also conducted in the EEC on the sequencing data obtained during the period 2011-2013. It aimed to highlight the relationships between eukaryotic parasites and other taxa as well as environmental parameters. This study revealed an unsuspected diversity of symbionts / decomposers in the EEC, and their importance for structuring microbial community and influence seasonal succession. More specifically, correlation network analysis showed the predominance of inter-taxa relations, over those between OTUs and abiotic parameters, and the central position of symbionts / decomposers in these relationships. This study highlighted the complexity and importance of microbial interactions in the structuring of microbial communities while providing crucial information to better understand underlying mechanisms. The last study was conducted in the Western Mediterranean, at four stations located in the Golfe du Lion, subject to the influence of the Rhône River, over a period of two years. This study aimed to provide the most exhaustive list of eukaryotic protist diversity and to evaluate its spatial variability in relation with environmental variables. This study further underlined the effect of the Rhône on the structuring of microbial communities, particularly in the eastern part of the Golfe du Lion. Overall, thesis work has reinforced, through combined approaches of microscopy and molecular biology, our knowledge of the functioning, taxonomic diversity and succession of microbial species, in relation to environmental parameters, of two contrasted marine ecosystems.



## SOMMAIRE

Liste des abréviations .....	1
Liste des figures .....	2
Liste des tableaux .....	4
Liste des annexes .....	5
<b>CHAPITRE I : INTRODUCTION GENERALE .....</b>	<b>7</b>
<b>PARTIE I : GENERALITES .....</b>	<b>9</b>
I. IMPORTANCE DES ECOSYSTEMES MARINS .....	9
II. DIVERSITÉ DES EUKARYOTES UNICELLULAIRES .....	10
II.1. Contexte historique .....	10
II.2. Classification et évolution .....	13
II.3. L'apport des techniques moléculaires .....	17
II.4. Phylogénie et évolution (classification à l'heure actuelle) .....	26
II.4.1. Classification des organismes eucaryotes .....	26
III. PLACE ET ROLE DES PROTISTES DANS LES ECOSYSTEMES COTIERS .....	43
III.1. Places et rôles des protistes dans les écosystèmes marins .....	43
III.2. Focus sur le rôle des protistes dans les réseaux trophiques (particulièrement planctoniques .....	44
III.2.1. Préambule .....	44
III.2.2. Les réseaux trophiques planctoniques .....	45
III.2.3. Focus sur le rôle écologique des protistes eucaryotes .....	46
<b>PARTIE II : SITES D'ETUDES .....</b>	<b>49</b>
I. MANCHE ORIENTALE .....	49
I.1. Caractéristiques physiques .....	49
I.2. Caractéristiques biologiques et diversité des eucaryotes unicellulaires .....	52

II. MEDITERRANNEE : Méditerranée occidentale et Golf du Lion .....	56
II.1. Méditerranée occidentale .....	57
II.2. Golf du Lion .....	57
II.2.1. Climat .....	58
II.2.2. Caractéristiques hydrodynamiques .....	59
II.2.3. Apport fluviaux dans le GoL .....	61
II.2.4. Caractéristiques biologiques.....	62
<b>PARTIE III : OBJECTIFS .....</b>	<b>65</b>
<b>CHAPITRE II : MATERIELS ET METHODES .....</b>	<b>61</b>
I. APPROCHES MOLECULAIRES .....	68
II. APPROCHE MICROSCOPIQUE .....	85
III. MESURES DES PARAMETRES PHYSICO-CHIMIQUES .....	86
IV. ANALYSES STATISTIQUES .....	90
<b>CHAPITRE III : DIVERSITE ET ACTIVITE DES EUCARYOTES MICROBIENS DANS UN ECOSYSTEME COTIER MESO-EUTROPHE (MANCHE ORIENTALE)</b>	
I. Résumé du contexte et objectifs .....	100
II. Manuscrit de l'article: <b>Diversity and potential activity patterns of planktonic eukaryotic microbes in a mesoeutrophic coastal area (Eastern English Chanel)</b>	
II.1. Summary .....	101
II.2. Introduction .....	102
II.3. Results .....	108
II.4. Discussion .....	130
II.5. Conclusion .....	143
<b>CHAPITRE IV: LES PARASITES EUCARYOTES DANS UN SYSTEME COTIER MESO-EUTROPHE CARACTERISE PAR DES BLOOMS PHYTOPLANCTONIQUES MARQUES</b>	

I. Résumé du contexte et objectifs.....	147
II. Manuscrit de l'article: <b>Parasitic eukaryotes in a meso-eutrophic coastal system with marked phytoplankton blooms</b>	
II.1. Summary.....	148
II.2. Introduction.....	149
II.3. Results.....	152
II.4. Discussion.....	156
II.5. Conclusion.....	163
<b>CHAPITRE V : DIVERSITE DES PROTISTES DANS UN ECOSYSTRME COTIER OLIGOTROPHE (MEDITERRANEE OCCIDENTALE, GOLF DU LION)</b>	
I. Résumé du contexte et objectifs.....	174
II. Manuscrit de l'article: <b>Small and (sub) mesoscale variability of protistan plankton communities in oligotrophic coastal ecosystems (NW Mediterranean, Gulf of Lions) assessed by NGS.</b>	
II.1. Summary.....	175
II.2. Introduction.....	176
II.3. Results.....	184
II.4. Discussion.....	201
II.5. Conclusion.....	206
<b>CHAPITRE VI : DISCUSSION GENERALE ET PERSPECTIVES .....231</b>	
<b>REFERENCES BIBLIOGRAPHIQUES.....257</b>	
<b>ANNEXES.....280</b>	



## **LISTE DES ABREVIATIONS**

**ADN** : Acide DésoxyriboNucléique

**ADNr** : ADN ribosomal, séquence d'ADN codant les ribosomes

**ARNr** : Acide RiboNucléique ribosomique

**ARNr 16S**: petite sous-unité 16S de l'ARNr

**ARNr 18S**: petite sous-unité 18S de l'ARNr

**BLAST** : Basic Local Alignment Search Tool, logiciel qui permet la comparaison de la séquence de chaque OTU avec une base de donnée afin de trouver son équivalent le plus proche.

**CNRS** : Centre National de la Recherche Scientifique

**FISH**: Fluorescent in situ hybridization

**MALV**: Marine Alveolata

**MAST**: Marine Stramenopiles

**NCBI** : National Center for Biotechnology Information (USA), il s'agit d'une base de données qui permet de trouver de nombreux renseignements dans le domaine de la Biologie.

**NGS** : Nouvelles technologies de séquençage

**NGS**: New Generation Sequencing

**OTU** : Operational Taxonomic Unit, il s'agit de l'équivalent moléculaire de l'espèce

**pb** : Paires de bases

**PCR** : Polymerization Chain Reaction, il s'agit de la réplication d'une partie ciblée de l'ADN (ou de l'ARN en cas RT-PCR)

**PR2** : Protist Ribosomal Reference Database, base de données taxonomiques de référence spécialisée dans les régions hypervariables V2 et V3 de l'ADNr 18S.

**SOMLIT** : Service d'Observation en Milieu Littoral

**SPECiMed**: Structures of Planktonic Ecosystems in the North-western Mediterranean

**V2 et V3** : Régions hypervariables du gène 18S de l'ADNr permettant la différenciation entre espèces et l'affiliation des OTU à un taxon bien précis.

## LISTE DES FIGURES

<b>Figure 1 :</b> Le microscope utilisé par Hooke et par lequel il est parvenu à découvrir les cellules. Il est de sa propre fabrication.....	10
<b>Figure 2 :</b> Le microscope utilisé par Antony Van Leeuwenhoeck .....	11
<b>Figure 3 :</b> L'arbre du vivant établi par Haeckel (1866) .....	15
<b>Figure 4 :</b> Classification du vivant en 5 règnes telle proposée par Robert H. Whittaker (1969). Cinq règnes ont été individualisés (Bactéries, Protistes, Plantes, Champignons et Animaux) et organisés dans le sens de la « complexité croissante » du bas vers le haut et de la diversification adaptative (mode de nutrition) de la gauche vers la droite.....	16
<b>Figure 5:</b> Arbre phylogénétique du vivant, modifié à partir de <i>Adl et al.</i> (2012) et <i>Burki</i> (2014).....	20
<b>Figure 6 :</b> Présentation schématique du processus de préparation et de séquençage des bibliothèques par les techniques de nouvelle génération les plus utilisées. Illustration empruntée à ( <i>Claudia Knief, 2014</i> ).....	25
<b>Figure 7 :</b> Schéma phylogénétique de l'arbre de vie basée sur les connaissances moléculaires actuelles (petite sous unité de l'ARNr et autres données moléculaires). Les triangles verts représentent les phyla, les divisions ou les groupes de haut rang taxonomique pour lesquels un membre a été cultivé et/ou décrit correctement. Les triangles rouges représentent les divisions ou les lignées fortement divergentes sans espèces cultivées/décrites (d'après López-García et Moreira, 2008).....	27
<b>Figure 8 :</b> Quelques organismes appartenant au phylum des Opisthokonta.....	27
<b>Figure 9 :</b> Quelques organismes et lignées appartenant au phylum des Amoebozoa.....	29
<b>Figure 10 :</b> Quelques organismes et lignées appartenant au phylum des Plantae.....	30
<b>Figure 11 :</b> Quelques organismes et lignées appartenant au phylum des Rhizaria.....	32
<b>Figure 12:</b> Quelques lignées appartenant au phylum des Straménopiles.....	34
<b>Figure 13 :</b> <i>Image à gauche :</i> Tubercule de pomme de terre atteint par le mildiou. <i>Image à droite :</i> Symptômes de mildiou sur feuilles de pomme de terre. Maladie causée par l'oomycète <i>Phytophthora infestans</i> .....	34
<b>Figure 14:</b> Quelques lignées appartenant au super-groupe des alvéolaires.....	36
<b>Figure 15:</b> Quelques espèces appartenant au phylum des Cryptophyta (1 et 2) et des Haptophyta (3).....	38
<b>Figure 16:</b> Quelques espèces appartenant au super-groupe des Excavés.....	40

<b>Figure 17:</b> Quelques espèces dont la position dans l'arbre phylogénétique eucaryote est toujours incertaine ( <i>Lecointre &amp; Le Guyader, 2006 ; Walker et al. 2011a</i> ).....	<b>42</b>
<b>Figure 18:</b> Représentation simplifiée de la boucle microbienne (en rouge) et la chaîne trophique classique (herbivore, en vert) ( <i>source : IRD</i> ).....	<b>46</b>
<b>Figure 19:</b> Site d'étude Manche Orientale – Situation géographique .....	<b>50</b>
<b>Figure 20 :</b> Représentation schématique du fleuve côtier et de la zone frontale de la Manche Orientale ( <i>Brylinski et al. 1991</i> ) .....	<b>51</b>
<b>Figure 21 :</b> Carte représentative de la Méditerranée et de la géolocalisation du Golfe du Lion.....	<b>58</b>
<b>Figure 22 :</b> Bathymétrie, circulation des masses d'eau et principaux forçages sur le plateau du GoL (d'après Millot, 1990). Les isobathes 160, 500, 1000 et 2000 sont représentées...	<b>59</b>
<b>Figure 23 :</b> Relation entre l'abondance (séquences de l'ADNr) et l'activité cellulaire relative (séquences de l'ARNr) pour les 100 OTU les plus abondantes classées par fonction trophique. Chaque OTU a été tracé en fonction de son nombre relatif d'ADNr (axe des abscisses) et d'ARNr (axe des y) sur le nombre total de séquences. Chaque point représente des pourcentages appariés (en échelles logarithmiques) de chaque groupe trophique de chaque échantillon.....	<b>244</b>
<b>Figure. 24 :</b> (a) En Manche Orientale les protistes hétérotrophes, en particulier les dinoflagellés hétérotrophes de grande taille, sont les principaux prédateurs des diatomées et les seuls consommateurs de colonies de <i>Phaeocystis globosa</i> . Notre hypothèse est que, l'accumulation de colonies de <i>Phaeocystis</i> pourrait être favorisée par le parasitisme sur les prédateurs et les concurrents potentiels de <i>Phaeocystis</i> . Les dinoflagellés peuvent alors subir un double contrôle descendant par les copépodes et les parasites. (b) En Méditerranée le phytoplancton de petite taille est consommé par du microzooplancton de petite taille à son tour consommé par le mesozooplancton. De même façon qu'en Manche Orientale, le rôle de parasites et leur potentiel de contrôle dans un contexte de changement global reste à élucider.....	<b>255</b>

## LISTE DES TABLEAUX

<b>Tableau 1 :</b> Les 45 phylums de protistes de Corliss (1984) regroupés en 18 « super-ensembles ».....	<b>16</b>
<b>Tableau 2 :</b> Nombre de séquences après chaque étape de traitement des jeux de données sous Mothur. MO : concerne les échantillons prélevés en Manche Orientale entre 2013-2015 et dont l'ADN et l'ARN ont été extraits. SPECIMED : concerne les échantillons prélevés en Golfe du Lion entre 2010-2014. Pour ces derniers, juste les échantillons prélevés entre 2012-2014 ont été intégrés dans l'étude. Séq.: Séquence, OTU : Unité taxonomique Opérationnelle .....	<b>84</b>
<b>Tableau 3 :</b> Statistiques relatives aux séquences présentes après compilation des 4 jeux de données issus du séquençage d'eucaryotes microbiens de Manche Orientale par NGS (pyroséquençage et Illumina) .....	<b>239</b>
<b>Tableau 4 :</b> Statistiques relatives aux séquences présentes après alignement des 4 jeux de données compilés issus du séquençage d'eucaryotes microbiens de Manche Orientale par NGS (pyroséquençage et Illumina).....	<b>239</b>
<b>Tableau 5 :</b> Statistiques relatives aux séquences retenues lors de différents tests de bornes d'alignement.....	<b>240</b>
<b>Tableau 6 :</b> (a) Bilan des caractéristiques physiques et biologiques en Manche Orientale et en Méditerranée Occidentale (b) Inventaire de la diversité obtenue via le séquençage à haut débit (Illumina) dans les deux écosystèmes : Manche Orientale et Méditerranée Occidentale à différentes périodes (2013-2015 et 2012-2014, respectivement). Les groupes taxonomiques soulignés ont été observés uniquement en Méditerranée Occidentale. Les groupes taxonomiques marqués en gras ont été observés uniquement en Manche Orientale. ND : groupe taxonomique non détecté.....	<b>251</b>

## **LISTE DES ANNEXES**

<b>ANNEXE 1 :</b> Les différentes étapes de traitement des 47 échantillons Pyro-séquencés telles que renseignées dans <i>Genitsaris et al., 2015</i> . Les échantillons ont été prélevés bimensuellement (sauf en cas de conditions environnementales difficiles) à la station SOMLIT en Manche Orientale entre Mars 2011 et Juillet 2013.....	<b>92</b>
<b>ANNEXE 2 :</b> Script élaboré pour la réalisation des différentes Analyses en Composante Principale (ACP) appliquées sur les variables environnementales des différentes stations étudiées. ....	<b>94</b>
<b>ANNEXE 3 :</b> Script des analyses de co-inertie appliquées sur les grands groupes taxonomiques et les variables environnementales des échantillons prélevés au niveau des quatre stations du Golfe du Lion (SOLA, MOLA, SOLEMIO et JULIO).....	<b>95</b>



# **CHAPITRE I – INTRODUCTION GENERALE**





## CHAPITRE I – INTRODUCTION GENERALE

### PARTIE I : GENERALITES

#### I. IMPORTANCE DES ECOSYSTEMES COTIERS MARINS

Les zones côtières sont des systèmes complexes situés à l'interface entre océans, continents et atmosphères, et de profondeur < 200m. Elles jouent un rôle considérable dans le déroulement des cycles biogéochimiques (carbone, azote et phosphore) et présentent une productivité biologique très importante (*Day et al., 1989 ; McLusky, 1989*). Leur forte productivité est liée à la présence d'un couplage étroit entre les compartiments benthique et pélagique et est associée d'une part aux apports terrigènes riches en matière organique et en nutriments (*Conley, 2000 ; Cloern, 2001*) et/ou aux phénomènes d'Upwelling (*Gonzales et al., 2004*) et/ou de diffusion (*Middelboe et al., 1998*). Leur localisation à l'interface océan/continent soumet leur fonctionnement à des forçages physico-chimiques de grande amplitude et à de forte variabilité spatio-temporelle. Ceci les expose à la pollution (anthropique ou naturelle) et également aux phénomènes d'eutrophisation (*Cloern, 2001 ; McManus et al., 2004*).

De par leur particularité, ces zones contribuent à environ 30% de la production primaire océanique et à 80% de la sédimentation de la matière organique et concentrent près de 90% des activités de la pêche mondiale (*Mantoura et al., 1991 ; Pernetta et Milliman, 1995 ; Pauly et Christensen, 1995*). La préservation de ces milieux sensibles et productifs nécessite une gestion durable qui passe par l'étude de la structure biologique des compartiments constitutifs et de leur fonctionnement (liens trophiques, *Algoni, 1998*). Dans les écosystèmes marins, les protistes jouent un rôle crucial en tant que producteurs primaires (phytoplancton), prédateurs (hétérotrophes et mixotrophes), et symbiotes (mutualistes et parasites). Les protistes apparaissent donc comme un compartiment clef, l'étude de leur diversité est fondamentale pour la compréhension des fonctionnements des écosystèmes marins (*Sherr et al., 2007; Caron et al., 2012*).

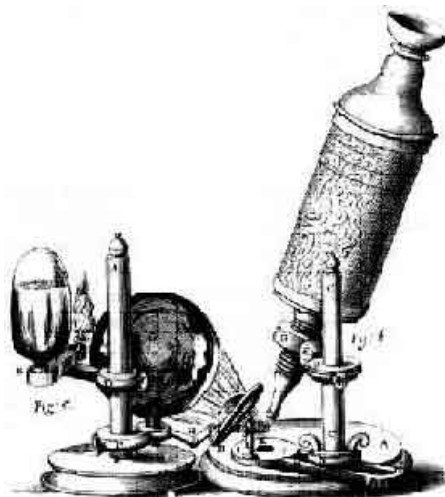
## II. DIVERSITÉ DES EUCARYOTES UNICELLULAIRES

### II.1. CONTEXTE HISTORIQUE

#### *Premières observations : Robert Hooke (1635-1703)*

C'est à Robert Hooke, un chimiste, mathématicien, physicien et inventeur anglais, que revient le mérite de l'invention du mot « cellule ». En 1665, au moyen d'un microscope rudimentaire de sa fabrication (**Fig. 1**) et en observant un morceau de liège, Hooke observa des cavités délimitées par des parois de cellulose et de suber. En 1667, il publia son ouvrage « *Micrographia* » où il nomma « cellules » les plus petites unités structurelles de la vie. Hooke marqua le début de la théorie cellulaire, selon laquelle tous les êtres vivants sont formés de cellules.

Son ouvrage est reconnu comme le premier de son genre dédié à des observations microscopiques. Il comporte 60 observations d'objets différents, accompagnés de dessins et des descriptions fines et détaillées et dont de nombreux objets biologiques. La première description certaine d'un organisme eucaryote microscopique fût celle d'un champignon du genre *Mucor*, préalablement décrit dans l'ouvrage de Hooke (*Hooke, 1665 ; Gest, 2004b*).

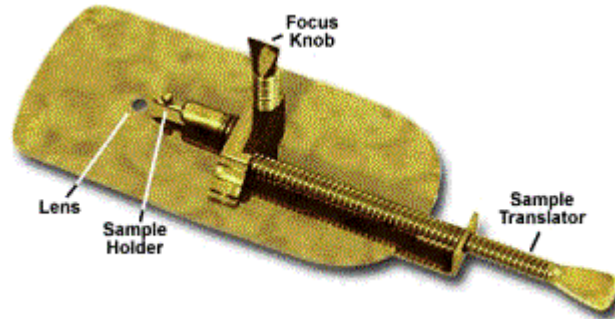


**Figure 1** : Le microscope utilisé par Hooke et par lequel il est parvenu à découvrir les cellules. Il est de sa propre fabrication.

#### *Descriptions à grossissement supérieur : Antony Van Leeuwenhoek (1632-1723)*

Une description plus détaillée des micro-organismes fût donnée par Antony Van Leeuwenhoek, marchand hollandais et chercheur amateur, inventeur du premier microscope (**Fig. 2**). Ce matériel permettait l'ajustement précis de la focalisation d'un objet

agrandi avec une résolution encore meilleure que celle des lentilles primitives. Depuis ses 41 ans (1673), Leeuwenhoek a consacré tout son temps à dessiner de façon très détaillée de nombreux micro-organismes qu'il avait observés dans l'eau de pluie ou dans d'autres échantillons.



**Figure 2 :** Le microscope utilisé par Antony Van Leeuwenhoek

### ***La théorie cellulaire : Matthias Schleiden (1804-1881) et Theodor Schwann (1810-1882)***

Après l'invention du microscope, plusieurs scientifiques ont commencé à s'en servir pour étudier davantage les cellules mises en évidence par Hooke et Van Leeuwenhoek. En 1836, le botaniste *Matthias Schleiden (1804-1881)* et le zoologiste *Theodor Schwann (1810-1882)* confirmèrent les hypothèses avancées par leurs collègues et établirent la théorie cellulaire, selon laquelle tous les organismes (bactéries, animaux et plantes), sont constitués de cellules et de produits cellulaires. Certains d'entre eux sont unicellulaires (formés d'une seule cellule autonome) ou pluricellulaires (formés de plusieurs cellules différenciés).

### ***Le second axiome de la théorie cellulaire : Robert Remak (1815-1865) et Rudolf Virchow (1821-1902)***

La découverte des microorganismes a suscité l'intérêt des scientifiques à découvrir l'origine de ces êtres. A cette époque, les microorganismes étaient supposés apparaître spontanément de la matière inorganique. Ceux tenants de cette théorie croyaient également que les êtres vivants de grandes tailles (souris, grenouilles, etc.) pouvaient se former d'eux-mêmes. Ce n'est qu'en 1855, que la théorie de la génération spontanée a été démentie, suite aux travaux de Robert Remak (1815-1865) et Rudolf Virchow (1821-1902). D'autres également avaient démenti l'hypothèse précédente comme Lazzaro Spallanzani (1729-1799), Charles Cagniard de La Tour (1838), Theodor Schwann (1837) et Friedrich Kützing

(1837). Le second axiome de la théorie cellulaire a été ainsi conçu : toute cellule provient d'une cellule préexistante.

***Rien ne provient du hasard : Louis Pasteur (1822-1895)***

De nombreuses expériences sont venues démentir la théorie cellulaire élaborée par Remak et Virchow, mais en 1861, Louis Pasteur (1822-1895) trancha et approuva que rien ne puisse croître spontanément. Les bactéries sont dans l'air, et leur croissance dans des solutions considérées comme stériles montre qu'il s'est produit une contamination. Les recherches de Pasteur se poursuivirent et permirent à la science de faire un bond considérable.

Les années qui suivirent représentèrent l'âge d'or de la microbiologie. Plusieurs microorganismes furent découverts. Avec le développement des méthodes de microscopies, de nombreux microorganismes ont pu être classifiés, essentiellement ceux connus néfastes à la santé humaine.

***De plus en plus d'organismes décrits : du 18<sup>ème</sup> au 20<sup>ème</sup> siècle***

Depuis le 18<sup>ème</sup> siècle, de nombreux organismes ont été décrits, parmi lesquels on trouve les ciliés (*Paramecium* ; Joblet, 1718), les dinoflagellés (*Noctiluca* ; Baker, 1753) et d'autres. Le zoologiste danois Otto Friedrich Müller (1730-1784) est considéré comme le premier à fournir une description claire et précise et à proposer une classification des êtres microscopiques, sur des critères morphologiques (*Animacula infusoria fluviatilia et marina*, 1786), dans laquelle on peut voir un début de distinction des bactéries. Il a construit la 1<sup>ère</sup> drague moderne, en 1779, lui permettant de découvrir plusieurs groupes marins tels que des protistes, des ciliés, des mollusques et des crustacés. Il a décrit et nommé, dans un ouvrage posthume (1786), plus de 300 organismes de micro-organismes eucaryotes. Certains noms sont toujours utilisés de nos jours (e.g. *Paramecium*, *Colpoda* et *Cyclidium*) (Corliss, 1991). Depuis Müller, plusieurs organismes ont été décrits grâce au microscope qui devenait de plus en plus performant. Suite à l'augmentation du nombre des espèces décrites, les biologistes se sont intéressés à les classer dans des groupes taxonomiques. Plusieurs ouvrages concernant la classification des protozoaires ont été publiés à cette époque (e.g. W.S. Kent 1880-1882 et O ; Bütschli 1880-1889 (Corliss,

1991 ; Fokin, 2013)). Les ciliés ont fait également l'objet de plusieurs publications (Cf. travaux de Balbiani, Chatton, Fauré-Fremiet (*Corliss, 1991*)).

## II.2. CLASSIFICATIONS ET EVOLUTION

### *Premières classifications du vivant*

Depuis l'antiquité, les scientifiques ont tenté de faire une classification des êtres vivants. Le premier à proposer une classification du monde vivant est Hippocrate 400 ans avant Jésus Christ, mais c'est Aristote qui parvient à un découpage cohérent présenté dans son œuvre "L'Histoire des animaux". Il s'est concentré particulièrement à l'étude des animaux (*Duris et Gohau, 2011*) quant à son disciple Théophraste s'est penché exclusivement sur la classification des végétaux. D'après Duris et Gohau (2011), les premières classifications ont été modestes et platoniques. Ce n'est qu'à l'arrivée de Carl Von Linné (1707-1778), un naturaliste suédois, que la classification du vivant prendra plus d'ampleur avec la fondation des bases du système moderne de nomenclature binomiale. De son point de vue, « *la connaissance des corps naturels nécessite de les nommer* » (« *Nomina si nescis, perit cognito rerum* », formule d'Edward Coke). Il a déterminé plus de 6000 espèces végétales et 4400 animales. En ce qui concerne les micro-organismes « les animalcules », il les a classés pour la plupart dans un genre *Chaos infusoria* (*Giordan, 1987*) sans prendre en considération les genres précédemment définis (e.g. *Paramecium*, *Colpoda* et *Cyclidium*). Néanmoins, sa méthodologie de classification reposant sur l'étude des traits morphologiques sera, comme pour les macro-organismes, également utilisée pour les micro-organismes.

### *Création d'un 3<sup>ème</sup> règne du vivant : Protista (Haeckel, 1866)*

Au 19<sup>ème</sup> siècle, la classification traditionnelle de Linné (1735) en deux groupes (végétal/animal) a évolué pour aboutir à la constitution d'un troisième règne (ou 4<sup>ème</sup> si on prend en compte le règne minéral). Sa constitution est venue du besoin de classer les animalcules ne partageant pas des critères communs avec les autres règnes existants (végétal/animal) et dont certains sont réclamés par les botanistes comme par les zoologistes (e.g. *Euglena*). Ce groupe généré supportait tous les micro-organismes et certains micro-

organismes. Plusieurs propositions ont été établies pour ce nouveau règne (par exemple : Protozoa d'Owen (1858), Primigenum de Hogg (1860) et Primalia de Cassin et Wilson (1863)) mais la plus marquante est celle d'Ernst Haeckel (1866) : le règne Protista. Il a repris la théorie de l'évolution de Darwin et a formulé la loi de la biogénétique fondamentale : « *L'ontogénèse récapitule la phylogénèse* » par observation du développement des embryons de vertébrés. Haeckel semble être le premier à proposer un arbre du vivant, montrant la relation entre les différents organismes. Dans son arbre (**Fig. 3**), il montrait que les Protista sont dérivés des animaux et des végétaux quoiqu'il fût convaincu que ce règne forme une part entière distincte des autres règnes.

### ***Découverte d'autres règnes : Edouard Chatton (1883-1947)***

Edouard Chatton fut le premier à observer les différents systèmes d'organisation cellulaire existant entre les Procaryotes (qui ont un chromosome unique sans membrane nucléaire et sans appareil de mitose, et une structure cellulaire élémentaire (pas de mitochondries)) et les Eucaryotes (qui possèdent un noyau entouré d'une membrane, des chromosomes, un appareil de mitose et une structure cellulaire complexe (mitochondries notamment)) et proposa en 1925 dans l'article *Pansporella perplexa* : Réflexions sur la biologie et la phylogénie des Protozoaires. Il a décrit 60 genres nouveaux et 150 espèces inédites de Protozoaires (1937) et a contribué également à la connaissance des dinoflagellés et des ciliés (*Christophe Charle et Eva Telkes, 1989*). Cette distinction cellulaire est restée longtemps négligée jusqu'à 1937 où Herbert Faulkner Copeland (1902-1968) l'a intégré dans la classification du vivant. Ceci lui avait permis de séparer les bactéries des autres règnes existants. Il créa ainsi un 4<sup>ème</sup> règne distinct : Les Monera (un règne contenant des organismes unicellulaires avec une organisation cellulaire procaryote (sans membrane nucléaire)). En 1969, Robert H. Whittaker a décrit une classification à cinq règnes : quatre règnes eucaryotes (Animal, Végétal, Champignons et Protistes) et les protistes qui se regroupent dans le règne des monères (**fig. 4**). La création du règne des champignons « Fungi » a résulté des observations des modes de nutrition et sont définis par une nutrition par absorption.

L'apparition du microscope électronique au milieu du 20<sup>ème</sup> siècle a permis d'avoir une meilleure caractérisation de nombreux caractères ultra-structuraux (e.g. forme des replis de la membrane interne mitochondriale). Par ceci, la séparation entre les espèces vues

semblables et l'identification de nouvelles espèces sont devenues possible. En 1984, Corliss découvre 45 grands taxons (**Tableau 1**), dont nombreux sont toujours d'actualité tels que les 3 groupes appartenant aux alvéolés (ciliés, Apicomplexa et dinoflagellés), les diatomées et les chlorophytes. D'autres caractères ont été identifiés à la fin du 20<sup>ème</sup> siècle permettant à nouveau une connaissance améliorée de la taxonomie des eucaryotes et particulièrement la détermination des liens de parentés entre les taxons (*Philippe et al., 2000*).

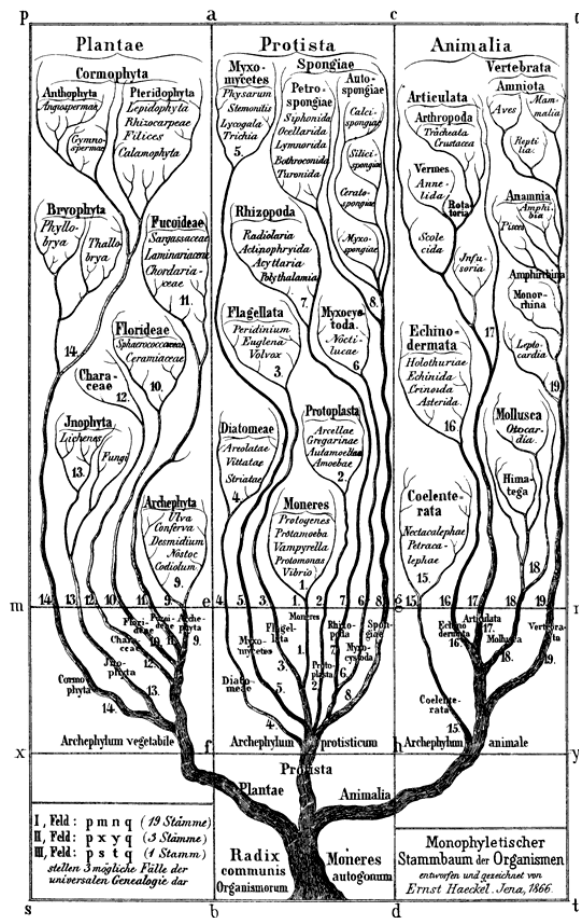
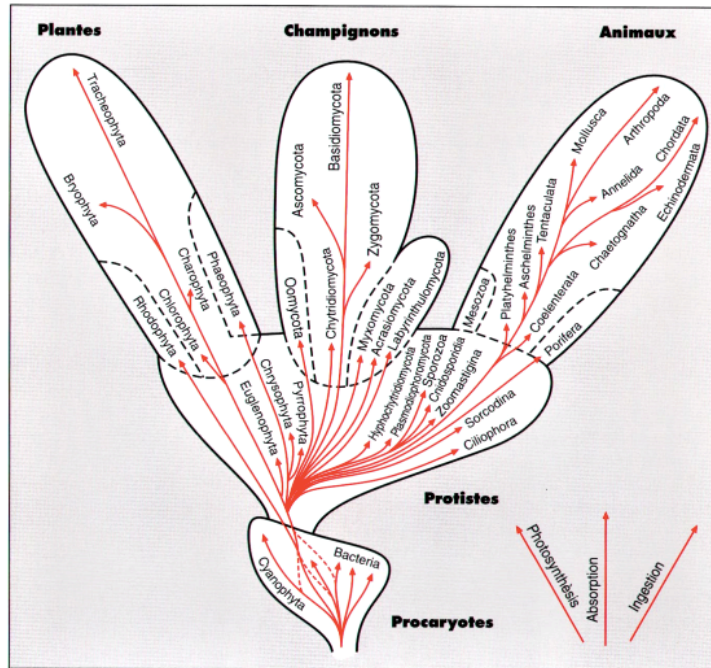


Figure 3 : L'arbre du vivant établi par Haeckel (1866)



**Figure 4 :** Classification du vivant en 5 règnes telle proposée par Robert H. Whittaker (1969). Cinq règnes ont été individualisés (Bactéries, Protistes, Plantes, Champignons et Animaux) et organisés dans le sens de la « complexité croissante » du bas vers le haut et de la diversification adaptative (mode de nutrition) de la gauche vers la droite.

<b>Rhizopodes</b>	<b>Chromobiontes</b>
Karyoblastea	Chrysophyta
Amoebozoa	Haptophyta
Acrasia	Bacillariophyta
Eumycetozoa	Xanthophyta
Plasmodiophorea	Eustigmatophyta
Granuloreticulosea	Phaeophyta
(incert. sed. : Xenophyophora)	(Incert. sed. : Proteromonadea
<b>Mastigomycètes</b>	Bicosoecida
Hyphochytridiomycota	Heterochloride
Oomycota	Raphidophyceae)
(Incert. sed. :	<b>Actinopodes</b>
Chytridiomycota)	Heliozoa
<b>Chlorobiontes</b>	Taxopoda
Chlorophyta	Acantharia
Prasinophyta	Polycystina
Conjugatophyta	Phaeodaria
Charophyta	<b>Dinoflagellés</b>
(Incert. sed. : Glaucophyta)	Peridinia
<b>Euglenozoa</b>	Syndinea
Euglenophyta	(Incert. sed. : Ebridea
Kinetoplastidea	Ellobiophyceae
(Incert. sed. : Pseudociliata)	Acritarcha)
<b>Rhodophytes</b>	<b>Ciliés</b>
Rhodophyta	Ciliophora
<b>Cryptomonadines</b>	<b>Sporozoaires</b>
Cryptophyta	Sporozoa
<b>Choanoflagellés</b>	(Incert. sed. : Perkinsida)
Choanoflagellata	<b>Microsporidies</b>
Labyrinthomorphes	Microsporidia
Labyrinthulea	<b>Haplosporidies</b>
Thraustochytriacea	Haplosporidia
<b>Polymastigotes</b>	<b>Myxosporidies</b>
Metamonadea	Myxosporidia
Parabasalia	(Incert. sed. : Actinomyxidea
<b>Paraflagellés</b>	Marteiliidea
Opalinata	Paramyxidea)

**Tableau 1 :** Les 45 phylums de protistes de Corliss (1984) regroupés en 18 « super-ensembles ».



### II.3. L'APPORT DES TECHNIQUES MOLECULAIRES

#### *Le lancement de la biologie moléculaire : Zuckerland et Pauling (1965), Woese et Fox (1977)*

Zuckerland et Pauling (1965) furent les premiers à considérer les molécules comme des marqueurs de l'histoire évolutive et à développer le concept d'horloge moléculaire. Ils ont pu identifier les polymères macromoléculaires comme étant les seuls portant directement dans leur séquence de l'information génétique et que leur utilisation puisse donner des informations fiables sur les relations entre les organismes. Les deux chercheurs supposèrent que parmi les différentes molécules sémantides (I<sup>aire</sup> : ADN, II<sup>aire</sup> : ARN<sub>m</sub>, III<sup>aire</sup> : les polypeptides synthétisés à partir d'un ARNm sans modification post-traditionnelle), les sémantides primaires comportent plus d'information en raison de la présence de séquences non transcrites et également des mutations iso-sémantiques rendues invisibles dans la séquence des protéines.

Près de dix ans après les travaux de Zuckerland et Pauling, Woese et ses collaborateurs (Fox et al., 1977; Woese et Fox, 1977) appliquent cette théorie en comparant différents taxons et chloroplastes en utilisant l'ARN ribosomique comme marqueur de l'évolution. Les premières études portent sur l'analyse de fragments de séquences et définissent des signatures spécifiques qui divisent le monde du vivant en trois grands domaines. La dichotomie du monde vivant en termes de Procaryote et Eucaryote était abandonnée au profit d'une division tripartite constituée d'Eucarya, de Bacteria et d'Archaea. Les études menées par la suite sur des séquences entières confirmèrent ces résultats (Woese, 1987; Woese et al., 1990) et proposèrent ainsi pour la première fois l'utilisation du marqueur moléculaire, le plus utilisé de nos jours, le gène codant l'ARN de la petite sous-unité du ribosome (16S ou 18S).

Dans les années qui suivent, Sogin et ses collègues (1986) s'intéressent à l'étude des liens de parenté entre les organismes eucaryotes et utilisent ainsi le gène codant l'ARN18S de la petite sous-unité du ribosome. Ils comparent les séquences des ADN18S de plusieurs eucaryotes unicellulaires (e.g. les amibes *Acanthamoeba castellanii* et *Dictyostelium discoideum*, la levure *Saccharomyces cerevisiae*, les ciliés *Paramecium tetraurelia* et *Tetrahymena thermophila*, le trypanosome *Trypanosoma brucei* et l'euglène *Euglena gracilis*), de plusieurs macro-organismes et également l'ADN16S de deux bactéries et deux archées.

Ces progrès ont encouragé les chercheurs à étudier davantage les liens de parenté entre les micro-organismes afin de déterminer une classification minutieuse, plus fiable que celle précédemment définie basée uniquement sur l'observation des traits phénotypiques.

### *L'arbre évolutif entre le passé et aujourd'hui*

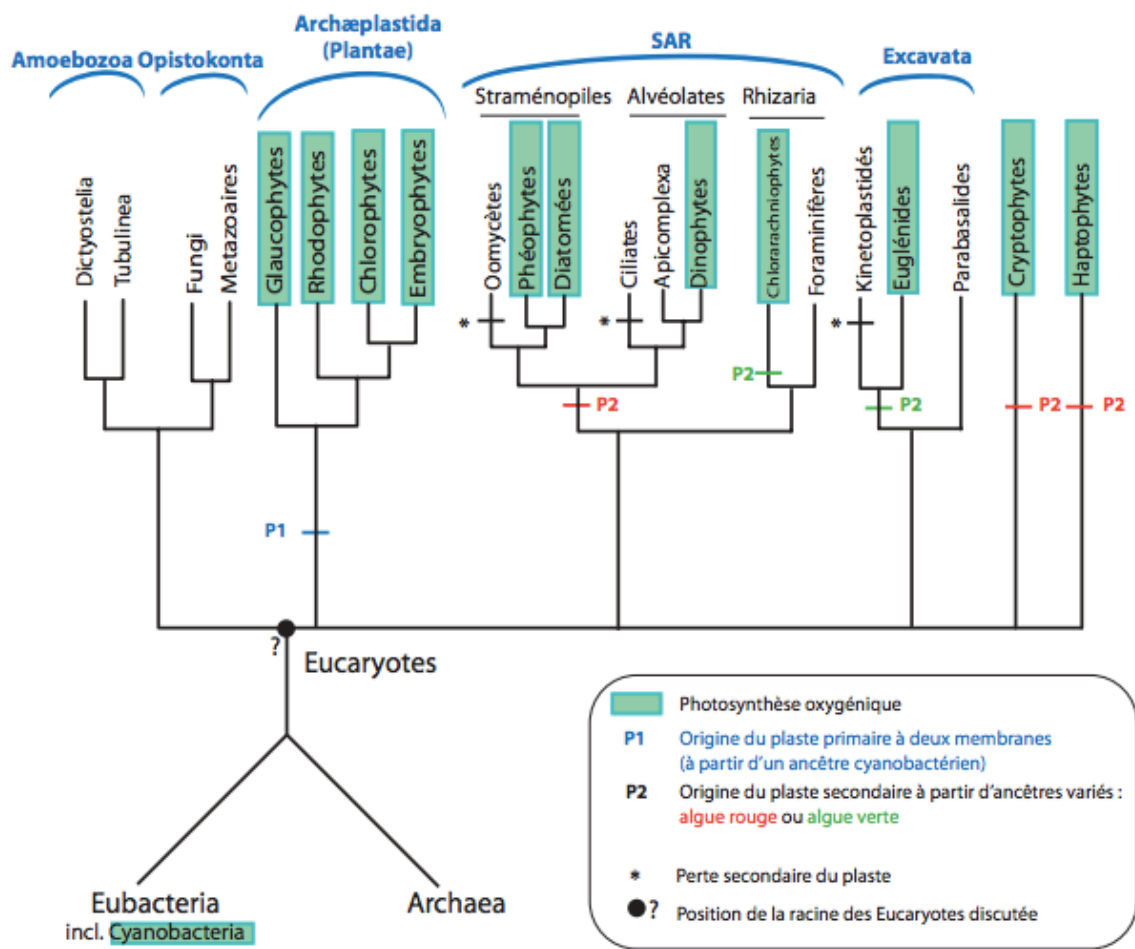
L'analyse comparative de séquences d'ARN ribosomal (ARNr18S dans le cas des eucaryotes), montre trois ensembles : les eucaryotes qui comportent les plus grands groupes (les animaux, les champignons et les plantes) ; les archées (ou Archaeobactéries) et les bactéries (ou Eubactéries). En 1989, Thomas Cavalier-Smith propose un nouveau taxon les *Archezoa*. Il regroupe diverses lignées de protistes supposés primitifs car dépourvus de mitochondries (apparues avant l'endosymbiose mitochondriale), telles les *Microsporidia*, les *Trichomonada* et les *Diplomonada* et qui forment un sous-règne au sein des *Protozoa*. Depuis, l'hypothèse des *Archezoa* est remise en cause. La découverte des gènes d'origine mitochondriale dans toutes les lignées d'eucaryotes amitochondriales (en particulier, celles d'émergence précoce comme les diplomonadines et les microsporides) a fortement rejeté l'hypothèse des *Archezoa* (Embley et Hirt, 1998). De nombreuses analyses ont démontré que toutes les branches se trouvant à la base du cladogramme basé sur l'ARNr étaient artefactuellement mal placées (Artefact de l'attraction des longues branches : **ALB**). L'artefact ALB a comme conséquence de grouper les plus longues branches qui correspondent aux taxons évoluant le plus rapidement, indépendamment de leurs véritables affiliations génétiques (Felsenstein, 1978). Les microsporides, supposées appartenir aux premières lignées eucaryotes (Vossbrinck et al., 1987), représentent un exemple spectaculaire d'ALB en raison de leur appartenance réelle à des champignons très évolués (Keeling, 2001b).

Depuis, les analyses continuent et les chercheurs essaient de trouver les positions phylogénétiques correctes de toutes les lignées pour lesquelles les séquences d'ARNr évoluent rapidement. Suite à l'utilisation des gènes autres que l'ADNr (les gènes codant des chaperonnes Hsp70 et Hsp90, ou les monomères de microtubules  $\alpha$  et  $\beta$ -tubulines notamment), seuls ou en concaténations, ainsi que de meilleurs modèles d'évolution (en prenant en considération les différences de vitesses d'évolution entre sites d'une séquence) réorganisent le cladogramme des eucaryotes. A nos jours, les eucaryotes se regroupent en

sept grands groupes (**Fig. 5**) : Alveolata, Excavata, Amoebozoa, Opisthokonta, Rhizaria, Archaeplastida (Plantae) et Stramenopiles.

### ***L'intérêt des outils moléculaires***

Classiquement, l'identification des micro-organismes se fait par microscopie sur des critères morphologiques ou de coloration, ou encore par tests biochimiques après isolement et culture. Or, l'identification via le microscope nécessite du temps et un grand savoir-faire afin de pouvoir différencier entre les micro-organismes ayant la même allure externe mais qui peuvent être phylogénétiquement différents. En outre, les outils moléculaires offrent un intérêt tout particulier pour les microorganismes environnementaux car la très grande majorité de ces micro-organismes est incultivable (*Zinger et al., 2011*) et sont donc complètement ignorés lors d'études utilisant des méthodes de microbiologie classiques basées sur la culture sur milieux synthétiques, mais également pour ceux cultivés, permettant une analyse phylogénétique, complémentaire et souvent plus précise que l'étude basée sur la morphologie. De plus, ces méthodes classiques requièrent l'isolement du micro-organisme de son milieu naturel et de sa communauté, ce qui biaise forcément l'évaluation de ses activités, ainsi que pour la plupart des approches *in situ* nécessitant des étapes d'incubation et soumettant ainsi les micro-organismes à des contraintes de confinement. Les méthodes moléculaires permettent de passer outre ces limitations. Elles permettent de caractériser les micro-organismes d'un échantillon, qu'ils soient cultivables ou non, dans leur environnement naturel (*Huybens et al., 2009*).



**Figure 5:** Arbre phylogénétique du vivant, modifié à partir de *Adl et al.* (2012) et *Burki* (2014).

### *Principes, objectifs et applications de quelques techniques de biologie moléculaire*

L'intérêt dans cette partie n'est pas de lister toutes les techniques moléculaires permettant d'identifier la diversité des protistes eucaryotes mais nous allons nous concentrer juste sur les techniques les plus utilisées et qui se basent sur l'amplification de marqueurs moléculaires, directement extraits d'un échantillon environnemental sans passage par une phase de mise en culture.

D'un point de vue pratique, ces techniques utilisent des séquences d'ADN spécifiques des microorganismes cibles afin de les identifier et d'estimer leur abondance relative. La séquence ciblée est choisie dans une région très conservée des gènes marqueurs, et est amplifiée par réaction de polymérisation en chaîne (PCR). Les gènes codants les ARN ribosomiques sont les plus employés, particulièrement le gène codant l'ARN de la petite sous-unité du ribosome (ADNr 18S chez les eucaryotes). Ce marqueur est de plus en plus informatif en raison de la richesse des bases de données de séquences qui en résulte. Il

existe également d'autres marqueurs utilisés comme l'ADNr 28S, plus long, ou l'ADNr 16S chloroplastidial en cas d'étude des eucaryotes photosynthétiques. L'amplification de ces marqueurs est assurée par utilisation d'amorces générales des eucaryotes. A partir du pool d'amplicons obtenus, plusieurs méthodes non mutuellement exclusives peuvent être mises en œuvre pour étudier la diversité moléculaire présente dans le pool d'amplicons et par extension dans l'échantillon environnemental.

⇒ *Quantification de la diversité et comparaison des échantillons*

Des techniques d'empreinte moléculaire peuvent être utilisées mais elles ne permettent pas d'identifier les organismes qui se trouvent dans l'échantillon environnemental. Par contre, elles peuvent permettre de comparer semi-quantitativement la diversité présente dans différents échantillons, ou d'estimer la part de la diversité moléculaire partagée par ces échantillons. Les techniques d'empreinte moléculaire les plus courantes sont l'électrophorèse sur gel avec gradient de dénaturation (DGGE) et le polymorphisme de longueur du fragment de restriction terminal (T-RFLP). Au cours de DGGE, les produits de PCR migrent dans le gel et rencontrent des concentrations en dénaturants de plus en plus importantes, entraînant la dénaturation progressive de leurs brins d'ADN selon leur composition en bases GC et AT. L'abondance des bandes sur le profil électrophorétique obtenu reflète la diversité de la communauté pour les micro-organismes considérés, chaque bande correspondant au moins à une espèce cible. Le T-RFLP repose sur la digestion de l'ADN par une ou plusieurs enzymes de restriction (enzymes coupant l'ADN). La taille des fragments est dépendante des enzymes utilisées. Après migration électrophorétique des fragments obtenus et révélation de l'extrémité marquée, les différentes bandes correspondent à des amplicons de séquences différentes. Dans les deux cas, comparer le nombre et la position des bandes sur le gel d'électrophorèse pour différents échantillons permet de comparer la diversité présente dans ces échantillons. Il est également possible de connaître l'identité des microorganismes par couplage de ces techniques au séquençage.

⇒ *Détermination de la composition des communautés microbiennes présentes*

Les amplicons qui découlent de la PCR sur l'ADN total extraits des échantillons environnementaux peuvent être séquencés, via plusieurs techniques, telles que le clonage et séquençage Sanger ou encore par séquençage direct des amplicons par des techniques de séquençage à haut débit (ex : pyroséquençage, Illumina) (**Fig. 6**). Les séquences obtenues sont supposées représenter la diversité génétique des amplicons et donc celle des

organismes présents dans chaque échantillon. Les séquences disposant d'un certain pourcentage d'identité partagé (ex : 98% pour les eucaryotes ; *Caron and Countway, 2009*), sont groupées en unités taxonomiques opérationnelles (OTU), considérées comme un proxy des espèces. Les séquences de chaque OTU détectée sont comparées à des organismes préalablement identifiés par analyses BLAST et/ou reconstructions phylogénétiques. Elles sont par la suite affiliées au même groupe taxonomique que les séquences de références proches, si celles-ci sont disponibles. Il existe certains organismes dont leurs séquences ne sont proches avec aucune séquence d'organisme cultivée et représentent donc une part encore non décrite de la diversité microbienne (*López-García & Moreira 2008*). Les proportions de chaque OTU au sein de l'ensemble des séquences issues de l'échantillon représentent de façon semi-quantitative l'abondance relative de l'espèce correspondante dans cet échantillon.

⇒ *Distribution et abondance relative*

Il est également possible de déterminer la distribution et l'abondance relative des communautés microbiennes par usage du FISH (technique d'hybridation *in situ* en fluorescence). Dans cette approche, les cellules sont marquées à l'aide d'un marqueur fluorescent visant les gènes de l'ADNr 18S ou 16S des gènes d'intérêt. Suite à un comptage au microscope, l'abondance relative du groupe cible, parmi l'ensemble des cellules de l'échantillon, peut être déterminée. Cette technique peut permettre également d'observer les cellules dont seulement le marqueur phylogénétique est connu (*Kolodziej et Stoeck, 2007 ; Jones et al., 2011*). D'un autre côté, il est possible de doser les pigments propres aux organismes photosynthétiques pour évaluer l'abondance relative des groupes concernés, comme pour les haptophytes où le 19'-hexanoxyloxyfuxoxanthine (19-Hex) est utilisé afin d'observer leur prédominance dans le phytoplancton marin (*Andersen, 2004 ; Liu et al., 2009*).

⇒ *Détermination de l'activité relative*

La détermination de l'activité physiologique des groupes planctoniques est une étape importante vers la répartition des rôles fonctionnels des différents organismes et ainsi à une meilleure compréhension du fonctionnement des écosystèmes. Chez certains groupes de procaryotes (bactéries et archées), le Card-Fish couplé à la micro-autoradiographie (MAR) a été souvent utilisé pour déterminer à la fois la composition et l'activité procaryotique (*Lee*

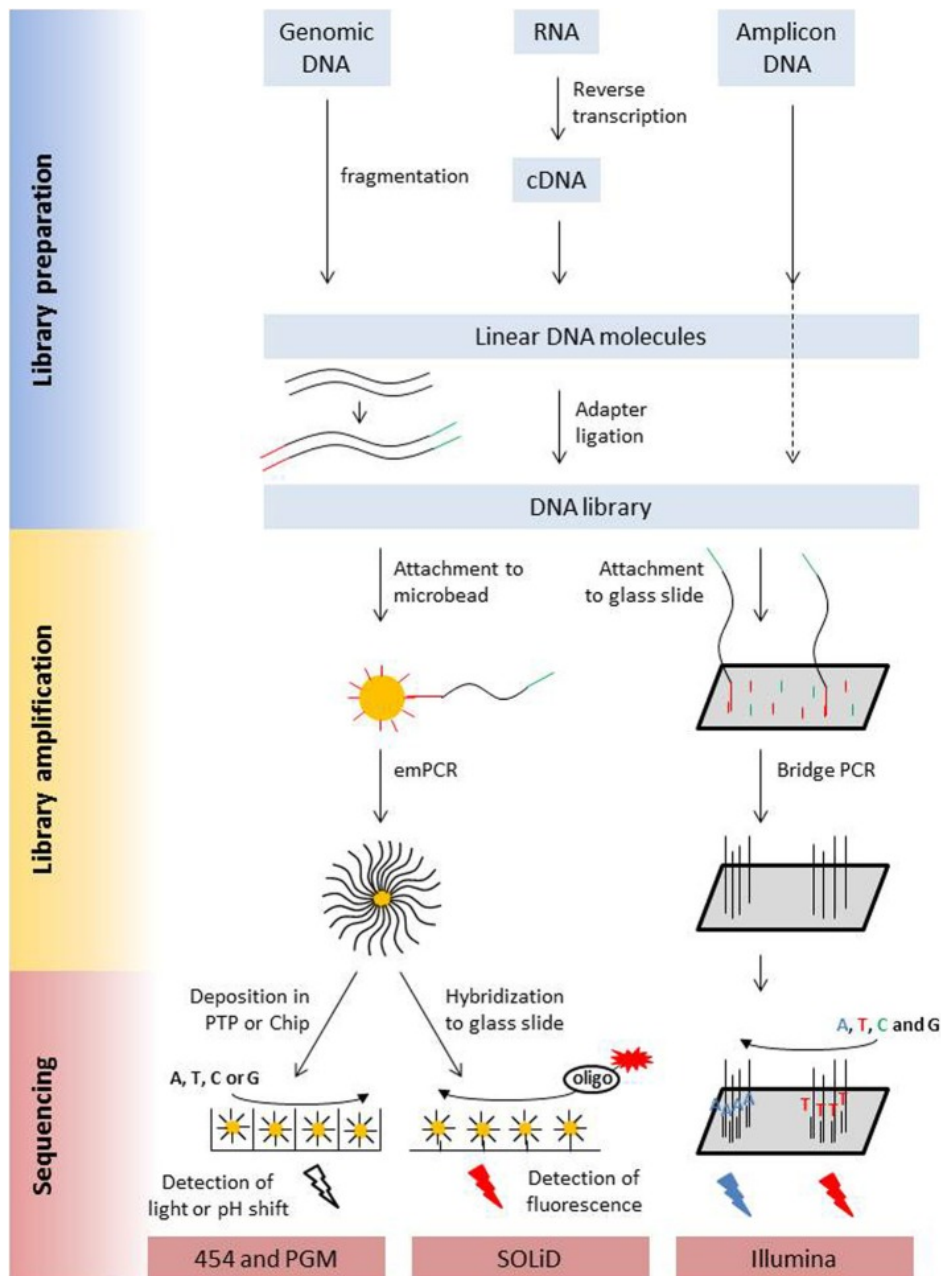
*et al.*, 1999 ; *Ouverney et Fuhrman*, 1999 ; *Cottrell et Kirchman*, 2000 ; *Elifantz et al.* 2005 ; *Alonso-Saez et Gasol*, 2007). La micro-autoradiographie consiste à utiliser un radioélément comme traceur de l'activité microbienne, elle permet de visualiser au niveau de la microscopie des grains d'argent ou des traces développées dans les radiations ionisantes émises par un radioligand incorporé à l'échelle de la cellule, quant au Card-Fish permet l'amplification d'un signal suffisamment fluorescent (*Trebesius et al.*, 1994 ; *DeLong et al.*, 1999) permettant ainsi d'identifier et de quantifier les populations microbiennes (*Wagner et al.*, 2003). Les données des deux techniques sont récupérées par l'utilisation d'un microscope à épi-fluorescence à l'aide d'un système d'analyses d'image et d'un logiciel de comptage.

Les techniques de séquençage dites de nouvelle génération (e.g. pyroséquençage, Illumina) permettent également de déterminer le potentiel d'activité des communautés planctoniques en plus de la diversité spatio-temporelle. Ceci est réalisé par ciblage simultané des ADN<sub>r</sub>16S/18S et ARN<sub>r</sub>16S/18S. Plusieurs études récentes ont proposées cette méthode chez les bactéries marines (*Mills et al.*, 2005; *Moeseneder et al.*, 2005; *Gentile et al.*, 2006; *Campbell et al.*, 2011), les archées (*Hugoni et al.*, 2013), et les eucaryotes (*Stoeck et al.*, 2007; *Not et al.*, 2009; *Orsi et al.*, 2013; *Logares et al.*, 2014; *De Vargas et al.*, 2015; *Charvet et Al.*, 2014; *Massana et al.*, 2015; *Hu et al.*, 2016). En effet, ce double ciblage de l'ADN<sub>r</sub> et de l'ARN<sub>r</sub> permet d'avoir une estimation de l'activité des cellules planctoniques via le calcul du ratio ARN<sub>r</sub>/ADN<sub>r</sub>. En effet, la quantité d'ARN<sub>r</sub> est, dans une certaine mesure, corrélée à la traduction des protéines et donc à l'activité métabolique des cellules. A l'heure actuelle, il est ainsi possible de considérer à la fois la diversité microbienne d'un échantillon, tout en évaluant l'activité métabolique de chaque espèce ou de chaque groupe taxonomique. Toutefois, certaines limitations demeurent quant à l'utilisation de ce ratio, liées essentiellement aux différentes stratégies de vie microbienne ainsi qu'aux activités non liées à la croissance (*Blazewicz et al.*, 2013). Ce rapport ARN<sub>r</sub>/ADN<sub>r</sub> pourrait éventuellement ne pas prendre en compte certains événements associés au cycle de vie des espèces (*Blazewicz et al.*, 2013). Par exemple, il a été démontré que certains procaryotes développent une stratégie de vie anticipée (*Tagkopoulos et al.*, 2008; *Mitchell et al.*, 2009; *Barnard et al.*, 2013) consistant à accumuler de l'ARN<sub>r</sub> dans des conditions environnementales défavorables, en période de faible activité, juste avant que les conditions redeviennent adéquates (*Zinser et al.*, 2009; *Lepp et al.*, 1998). Ceci n'a pas été signalé pour les protistes eucaryotes dans les milieux marins.

## ***Quelques limites des méthodes moléculaires et la nécessité des observations traditionnelles***

Dès leur apparition, les méthodes moléculaires ont permis de s'affranchir de l'étape de la mise en culture, et elles étaient souvent utilisées dans la détermination de la communauté bactérienne. Ce n'est que quelques années plus tard que ces techniques ont été développées afin d'étudier la diversité eucaryote. Leur utilisation a permis de découvrir un large pan d'organismes sans représentants cultivés (*Moreira et López-García, 2002*). Néanmoins, ces techniques présentent certaines limites. En termes de qualité, certaines OTU considérées comme des lignées divergentes peuvent être éventuellement artefactuelles (e.g. chimères, erreurs de séquençage (*Kunin et al., 2010*)). En termes de quantité, pour certains organismes, l'extraction d'ADN est plus difficile que chez d'autres, les amorces universelles utilisées n'amplifient pas aussi bien les marqueurs moléculaires de tous les organismes (*Zinger et al., 2011*) et également, le nombre de copies d'ADN18S dans les génomes varie d'une espèce à une autre. Ainsi, la proportion de séquences affiliées à un groupe donné dans une base de données de séquences ne correspond pas toujours exactement à la proportion de ce groupe dans l'échantillon environnemental de départ. Finalement, cette approche moléculaire ne peut pas remplacer les observations classiques des espèces. En effet, sans une description préalable des espèces, l'affiliation des séquences environnementales d'un taxon ne peut avoir lieu. De plus, elle ne fournit aucun renseignement sur la biologie, le métabolisme, ou encore l'écologie des organismes détectés. Comme indiqué précédemment, le FISH permet de visualiser des organismes décrits uniquement par une séquence, toutefois, cette technique présente également des biais. Elle ne permet d'observer simultanément qu'un nombre limité d'espèces différentes sur un même échantillon et ces observations ne sont pas toujours fiables. Comme par exemple, en se basant sur des observations réalisées grâce au FISH, Not et collègues (2007) ont pu identifier des organismes photosynthétiques qu'ils ont nommés « Picobiliphytes ». Des études génomiques réalisées quelques années plus tard ont montré qu'il s'agissait plutôt d'un groupe d'organismes hétérotrophes, bactérivores, renommé « Picozoa » (*Seenivasan et al., 2013 ; Moreira et López-García, 2002*). Les approches moléculaires et classiques sont ainsi complémentaires.





**Figure 6 :** Présentation schématique du processus de préparation et de séquençage des bibliothèques par les techniques de nouvelle génération les plus utilisées. Illustration empruntée à (Claudia Knief, 2014). Toutes les différentes molécules de départ sont transformées en molécules d'ADN à double brin reliées par des adaptateurs. Ces derniers sont spécifiques à la plate-forme de séquençage et permettent de lier les molécules des bibliothèques aux surfaces, soit des billes, soit une cellule d'écoulement, où elles sont amplifiées avant séquençage. Les amplicons clonaux sont séparés spatialement sur des lames de verre, des puces ou des plaques en fibre optique (PicoTiterPlate). Le séquençage est soit un processus de séquençage par ligation avec des oligonucléotides marqués par fluorescence de séquence connue, soit un processus de séquençage par synthèse. Pendant le séquençage par Illumina quatre nucléotides marqués différemment sont déposés sur la cellule d'écoulement en cycles multiples, selon la longueur de lecture souhaitée. Au cours du 454-pyroséquençage et de l'Ion PGM, les nucléotides non marqués sont rincés dans un ordre séquentiel sur la cellule d'écoulement. L'incorporation est détectée par une réaction de lumière couplée (454) ou par détection des protons libérés lors de l'incorporation des nucléotides.

## II.4. PHYLOGENIE ET EVOLUTION (classification à l'heure actuelle)

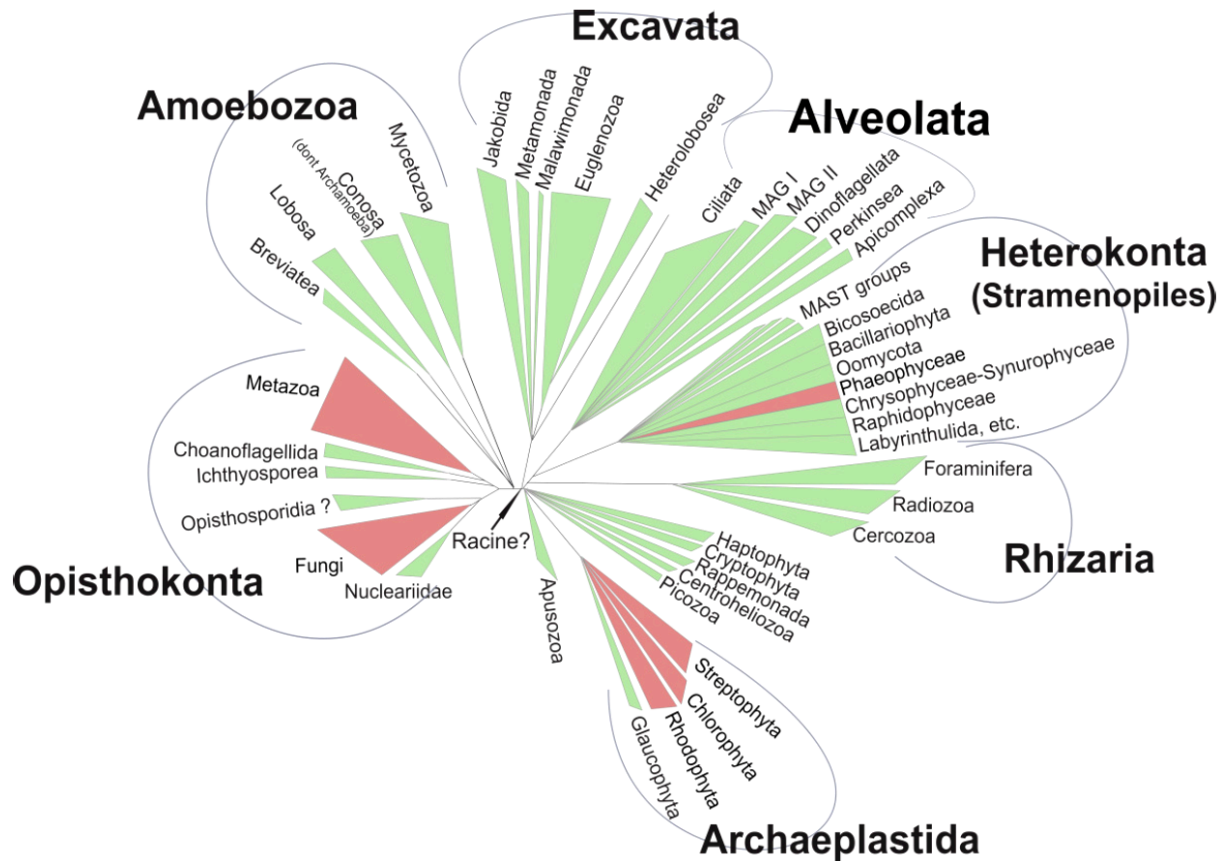
La phylogénie et donc la classification des eucaryotes connaît aujourd'hui une grande évolution. Les phylogénies moléculaires permettent de révéler des parentés insoupçonnées, dès lors que les artefacts de reconstruction phylogénétique sont considérés et corrigés.

Notre compréhension de la biologie des eucaryotes est basée essentiellement sur l'étude des plantes terrestres, des animaux et des champignons. Or, la majeure partie des eucaryotes est composée principalement de phylums unicellulaires dont un nombre important est encore mal caractérisé.

### II.4.1. Classification des organismes eucaryotes

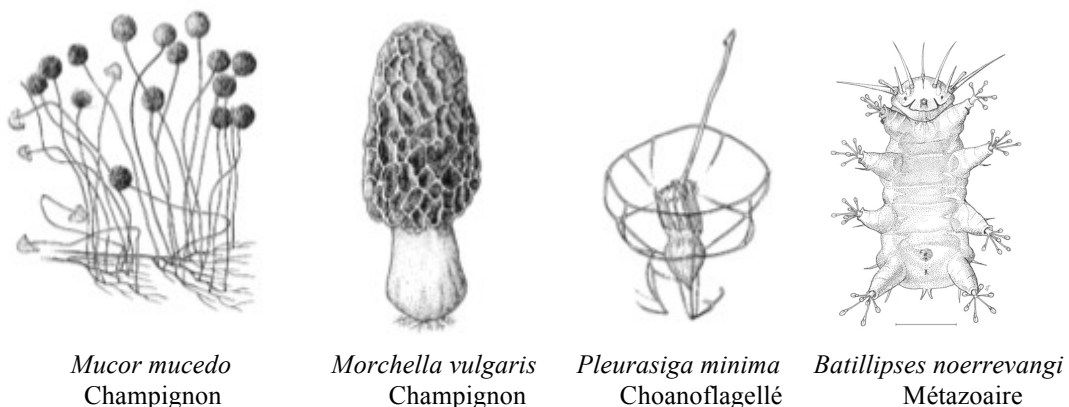
Une bonne compréhension de la phylogénie des eucaryotes a commencé à émerger autour des données de projets de séquençage à grande échelle (*Burki et al., 2007 ; Hackett et al., 2007; Rodriguez-Ezpeleta et al., 2005*). Toutefois, il reste encore une diversité non négligeable d'eucaryotes peu connue et dont sa position dans l'arbre phylogénétique eucaryote reste encore mal définie (*Adl et al., 2005*). De plus, la connaissance de la diversité cryptique des eucaryotes, supposée énorme, reste très faible (*Slapeta et al., 2005*). Enfin, une énorme diversité des pico- et nano-eucaryotes vient d'être découverte dans des banques de clones obtenues par amplification par PCR d'ADN environnemental (*Moreira et López-García, 2002*).

A l'heure d'aujourd'hui, sept grands super-groupes eucaryotes sont les mieux étudiés (définis en particulier par López-García et Moreira, en 2008) (**Fig. 7**).



**Figure 7 :** Schéma phylogénétique de l'arbre de vie basée sur les connaissances moléculaires actuelles (petite sous unité de l'ARNr et autres données moléculaires). Les triangles verts représentent les phyla, les divisions ou les groupes de haut rang taxonomique pour lesquels un membre a été cultivé et/ou décrit correctement. Les triangles rouges représentent les divisions ou les lignées fortement divergentes sans espèces cultivées/décrites (d'après López-García et Moreira, 2008).

### Les Opisthokonta

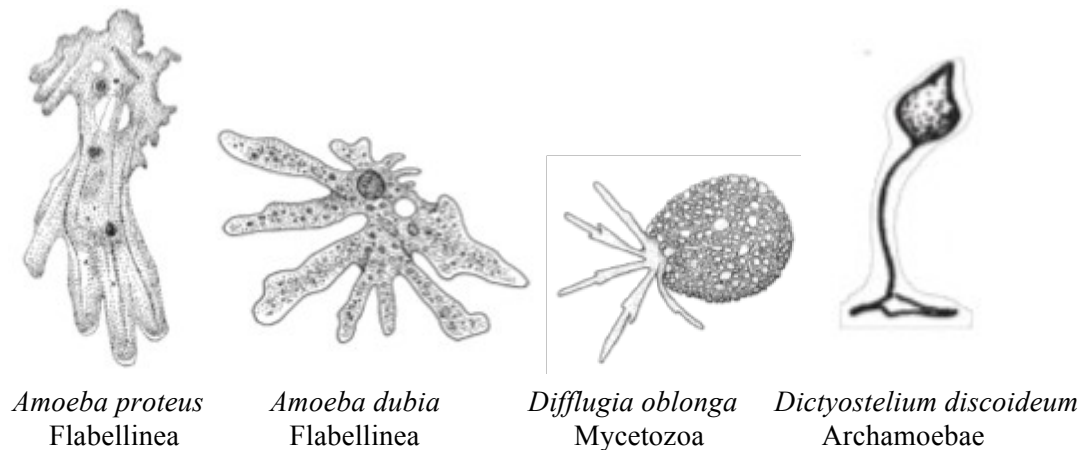


**Figure 8 :** Quelques organismes appartenant au phylum des Opisthokonta (Lecointre & Le Guyader, 2006).

Les opisthocontes regroupent des organismes hétérotrophes, unicellulaires ou pluricellulaires. Ils présentent une large diversité de morphologies, habitats et modes de vie. On y trouve, les animaux et les champignons qui représentent les groupes taxinomiques les mieux explorés (*Steenkamp et al., 2006*). Chacune des lignées possèdent au moins deux taxons unicellulaires frères ; les ichtyospores et les Choanoflagellés dans le cas des animaux, et les nucléariides et les rozellida dans le cas des champignons (*Lara et al., 2009*). Ces taxons regroupent des organismes dont le caractère ancestral commun est un flagelle unique postérieur. Les premières branches des champignons sont celles du groupe paraphylétique des chytridiomycètes (*James et al., 2006*). Ce sont les seuls champignons avec des flagelles. Les autres se constituent d'organismes multicellulaires à hyphes ou unicellulaires (levures). Les champignons se nourrissent par osmotrophie, ils sont essentiellement saprotrophes, parfois parasites ou symbiotes, et jouent ainsi un rôle important en tant que décomposeurs ou symbiotes d'embryophytes (e.g. mycorhizes), notamment dans les écosystèmes terrestres. Bien que l'arbre phylogénétique fongique soit maintenant clair, le nombre d'espèces non découvertes est estimé à 95% des espèces existantes (*Vandenkoornhuyse et al., 2002 ; James et al., 2006*). Les Choanoflagellés et les nucléariides sont des prédateurs libres qui se déplacent respectivement à l'aide d'un flagelle ou par déplacement amiboïde (cellule déformant et émettant des pseudopodes). Ils sont connus par leur ressemblance avec les cellules à collerettes (choanocytes) des éponges (Porifera). Les ichtyosporés quant à eux regroupaient des organismes parasites ou symbiotes. Récemment un nouveau groupe a été proposé parmi les opisthocontes. Les Opisthosporidia (*Karpov et al., 2014*) rassembleraient les microsporidies, les Cryptomycota-Rozellida et aphélides, tous parasites intracellulaires, osmotrophes ou phagotrophes.

Ce super groupe a été constitué au début des années 1990, sur la base de phylogénies moléculaires. La possession d'un flagelle unique et propulseur, est comptée comme une synapomorphie du groupe. Les Choanoflagellés, les ichtyosporés et les métazoaires d'une part, et les champignons, les nucléariides et les opisthosporides d'autre part sembleraient former deux clades différents.

## Les Amoebozoa



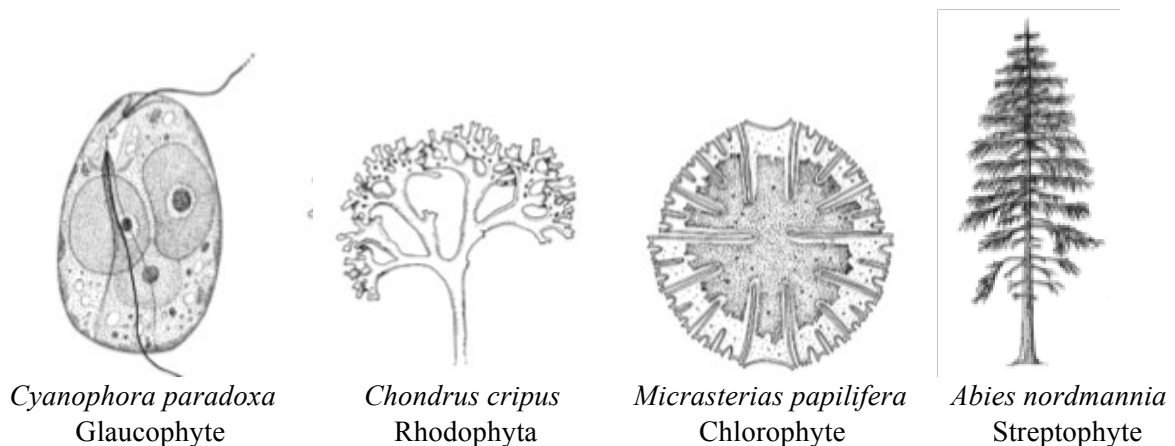
**Figure 9 :** Quelques organismes et lignées appartenant au phylum des Amoebozoa (Lecointre et Le Guyader, 2006 ; Walker et al., 2011a).

Les amibozoaires constituent un grand groupe incluant plusieurs divisions d'amibes (e.g. Archamoebae et les Mycetozoa) (Bapeste et al., 2002). Ils regroupent des organismes amiboïdes ou amiboflagellés (formant aujourd'hui le groupe monophylétique des percolozoaires) et ont tendance à avoir des pseudopodes en forme de tube (appelés également lopodia), un noyau simple et des mitochondries tubulaires à sillon (Adl et al., 2005). Très peu de caractéristiques permettent d'identifier ce groupe, construit essentiellement sur des bases moléculaires. La présence des pseudopodes peut être considérée comme un synapomorphie. Néanmoins, ce critère est partagé avec d'autres taxons (e.g. Chlorarachniophytes, Foraminifères, Mycétozoaires et Percolozoaires) (Parfey et al., 2006). Ces pseudopodes leur permettent le déplacement généralement provoqué par l'existence d'une molécule dans le milieu extracellulaire (chimiotactisme) ainsi qu'à la nutrition. La plupart sont sous forme libres se nourrissant par bactériovorie en ingérant leurs proies à l'aide de leurs pseudopodes. D'autres s'alimentent des débris dans les flaques d'eaux ou même d'autres parasites. Certains taxons des amibozoaires sont parasitaires telles que des archéamibes. Elles comptent parmi elles le genre *Entamoeba* pathogène pour les humains, provoquant des maladies telles que la dysenterie amibienne. Les organismes y appartenant sont anaérobiques ou microaérophiles et ont perdu leur mitochondrie. Ils sont assez proches des mycétozoaires. Certaines amibes (dites sociales), telle que *Dictyostelium discoideum*, forment par agrégation des structures pluricellulaires afin d'assurer la dispersion des spores, formes de résistance. Les amibes communes du sol, appartenant à la

famille des *Arcellinidae* (division des Lobosea, genre *Tabulinea*), sont les seules qui possèdent une thèque, construite à partir de matériel organique.

Ce super groupe est relativement récent. Les affiliations sont incertaines en raison du faible nombre d'organismes isolés et le peu de données moléculaires disponibles sur les organismes. Les Amoebozoa sont supposés d'être le groupe frère des opisthocontes, formant à eux deux les unicontes (eucaryotes uniflagellés) (*Cavalier-Smith, 2002*).

### *Les Plantae* (selon Cavalier-Smith)



**Figure 10 :** Quelques organismes et lignées appartenant au phylum des Plantae (*Lecointre & Le Guyader, 2006*).

Egalement appelé Archaeplastida (selon le Guyader et Lecointre), c'est à la base de ce groupe que serait apparue la photosynthèse d'eucaryotes en premier (*Adl et al., 2005*). La plupart des espèces comprises dans ce super-groupe sont photosynthétiques à l'exception de quelques lignées parasites ou hétérotrophes mineurs. Les organismes photosynthétiques colonisent tous les écosystèmes, terrestres et aquatiques. On distingue trois lignées principales, les Rhodophytes, les Glaucophytes et les Chloroplastides (ou Viridiplantae). Ces derniers regroupent les streptophytes et les chlorophytes et sont répartis dans tous les milieux, marins, eaux douces et terrestres. Le reste des eucaryotes photosynthétiques ont procuré leurs plastes par absorption d'un organisme appartenant à ce groupe (endosymbiose secondaire), sauf une exception récente (*Nowack et al., 2008*), et de nombreuses preuves affirment que les chloroplastes des Plantae sont monophylétiques (*Reys-Prieto et al., 2007*). La monophylie de leurs génomes nucléaires est incertaine, possiblement en raison de l'antériorité du groupe et/ou du prélèvement moléculaire rares des Rhodophytes et des Glaucophytes.

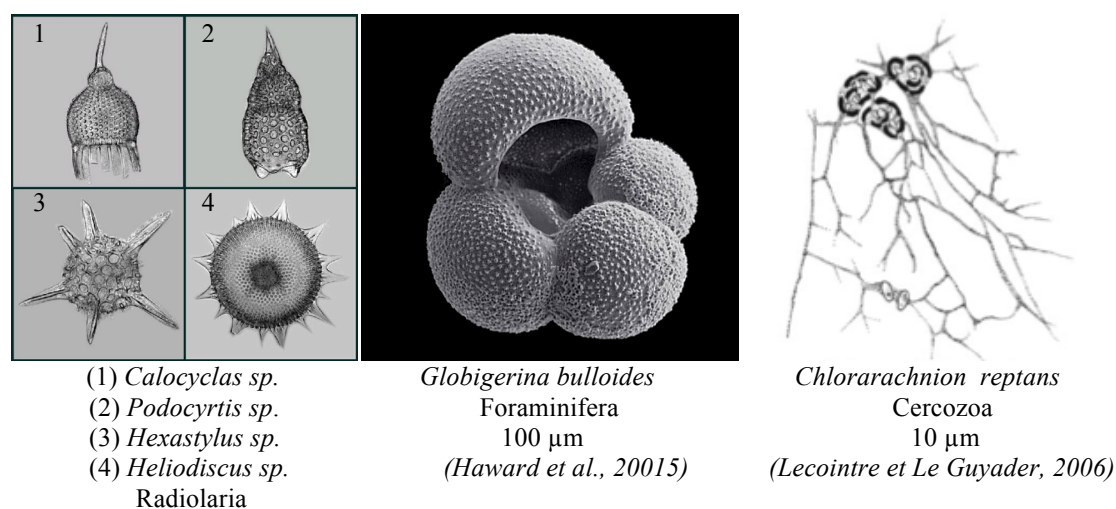
Les Glaucophytes constituent le groupe le moins abondant, nageant dans les mares douces des régions tempérées et les milieux terrestres. Ils regroupent les organismes unicellulaires structurés dorso-ventralement, au dos arrondi et au ventre et présentant deux flagelles de longueurs inégales. Les organismes regroupés sous ce taxon se caractérisent par des chloroplastes (ou cyanelles), de couleur bleu-vert contenant les pigments accessoires, phycocyanine et l'allophycocyanine, ainsi que par la présence d'alvéoles sous la membrane plasmique sous-tendues par des microtubules.

Les Rhodophytes, ou algues rouges, sont pour la plupart marines et multicellulaires, benthiques ou pélagiques, mais certains occupent les eaux douces. La plupart sont sessiles, se développant fixés sur un substrat (e.g. *Chondrus crispus* et *Mastocarpus stellatus* (*Gigartina stellata*)). Ils sont distingués des autres groupes par leur composition pigmentaire avec un type de chlorophylle unique, la chlorophylle *a*, des caroténoïdes et des pigments caractéristiques, les phycobiliprotéines. Ce groupe possède des thylacoïdes libres, ne formant pas de grana (ceci est lié à la présence des phycobilisomes à la surface des thylacoïdes), ils sont répartis concentriquement dans le chloroplaste (ou rhodoplaste), mais occupant tout l'espace, contrairement à ceux des Glaucophytes et des Cyanobactéries. Ce groupe se caractérise également par sa capacité de réaliser la photosynthèse même à faible luminosité. Certaines algues rouges sont abondantes dans les récifs de coraux. Les corallines produisent une enveloppe extracellulaire de carbonate de calcium qui, leur confère l'aspect d'une roche. D'autres sont très résistantes à des conditions extrêmes (extrémophiles) (e.g. *Cyanidium caldarium*, vivant à un pH inférieur à 1 dans les sources acides). Deux sous-groupes majeurs sont reconnus, *Bangiophyta* et *Florideophyta*, qui ont possiblement tous deux inventé la multicellularité séparément.

Les Chloroplastides regroupent les plantes terrestres et une grande partie des algues vertes. Ils comprennent les *Chlorophyta*, les *Ulvophyta*, les *Prasinophyta* et les *Strepsytera*. On retrouve des organismes multicellulaires mais également unicellulaires, représentant les principaux constituants du phytoplancton. Chez les chlorophycées, *Chlamydomonas sp.* est considérée comme l'espèce la plus utilisée dans les travaux de recherches, essentiellement dans l'étude de la photosynthèse eucaryote. Ceci en raison de son développement dans l'obscurité sans besoin d'un apport en carbone organique, tout en maintenant sa particularité photosynthétique (Harris et al., 2001). C'est un des genres retenus pour la production éventuelle de biocarburants (hydrogène ou hydrocarbure) par des algues

(Ghirardi et al., 2007). En ce qui concerne les prasinophytes, leur diversité taxonomique massive a été dévoilée au travers de récentes études moléculaires (Guillou et al., 2004 ; Worden et al., 2006). Ces organismes vivent en eau douce ou salée et peuvent épisodiquement présenter des abondances transcendantes, tel le cas chez *Ostreococcus tauri*. Cet organisme unicellulaire domine par son abondance dans l'étang de Thau dans le sud de la France, atteignant jusqu'à 30% de la biomasse photosynthétique (Vaquer et al., 1996). *O. tauri* est compté parmi les espèces pouvant mieux s'adapter aux conditions défavorables. En 2012, une équipe du laboratoire d'Océanographie microbienne de l'observatoire océanologique de Banyuls sur mer (LOMIC) a réalisé les premiers tests sur *O. tauri* comme bio-senseur luminescent afin de tester la toxicité de l'environnement. *Micromonas pusilla*, un picoeucaryote photosynthétique appartenant également aux prasinophytes, est largement distribuée dans certains écosystèmes marins. Son abondance peut atteindre entre  $10^3$  et  $10^4$  cell.mL<sup>-1</sup> dans les régions côtières, telles les côtes norvégiennes ou dans le golf de Naples (Harrison et al., 1991 ; Throndsen, 1976).

### Les Rhizariens (Rhizaria)



**Figure 11 :** Quelques organismes et lignées appartenant au phylum des Rhizaria.

Les Rhizaires aussi appelés Rhizariens (Lecointre et al., 2006) (ou Rhizaria) sont des organismes le plus souvent unicellulaires. Ils peuvent vivre dans des eaux oligotrophes en s'associant symbiotiquement à des algues. Leur biomasse a été pour longtemps sous-estimée en raison de leur fragilité (facilement détruits par le filet à plancton). L'évolution des techniques d'échantillonnage non destructives a permis de réévaluer l'importance de ce super-groupe. Ces protistes représentent 33% de l'abondance totale du plancton



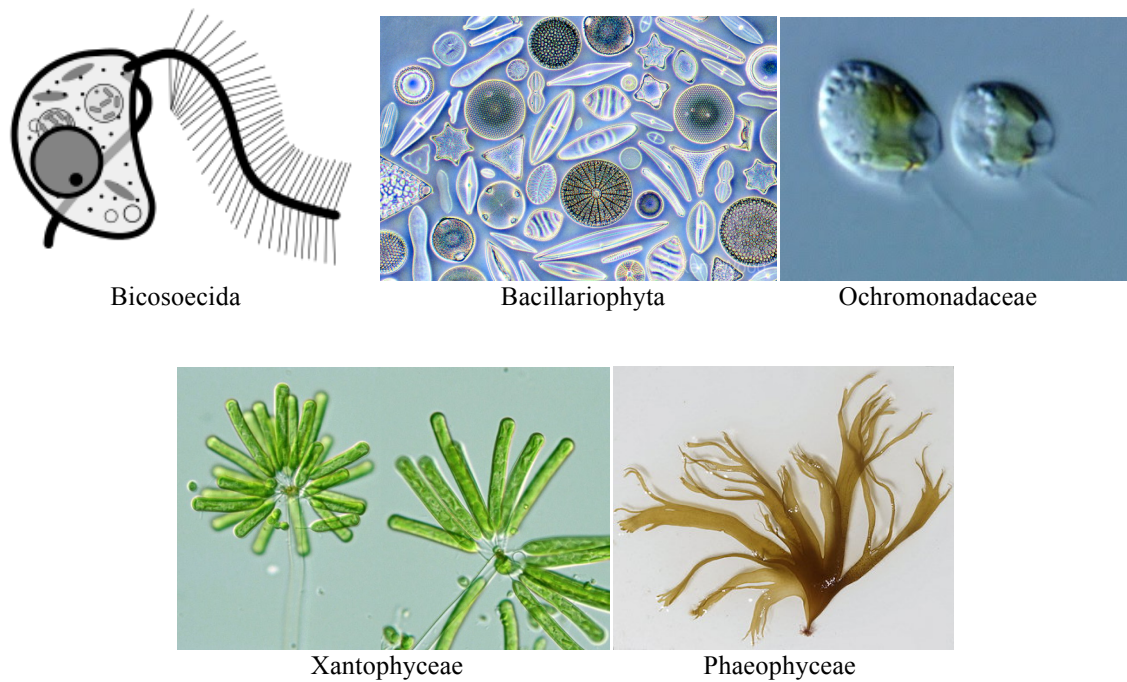
hétérotrophe de grande taille à l'échelle de l'océan mondiale et contribuent à 5% de la biomasse marine globale (Caron, 2016). La plupart sont de forme amiboïde, ils faisaient historiquement partie des Amoebozoa. Leur positionnement a été incertain, considéré auparavant également comme groupe frère des Excavata (Cavalier-Smith, 2002) puis des Straménopiles (Burki et al., 2007) (exactement aux Halvaria). Ce super-groupe se divise en trois groupes principaux : les Radiolaires, les Foraminifères et les Cercozoaires (Pawlowski et Burki, 2009).

Les Radiolaires sont exclusivement marins possédant un squelette à base de dioxyde de silicium autour duquel rayonnent de fins pseudopodes (axopodia). Ils vivent à l'état isolé mais peuvent également exister sous formes coloniales (Caron et al., 2012). Ils se caractérisent par la présence d'une capsule centrale séparant l'ectoplasme et l'endoplasme. Les foraminifères sont également caractérisés par la présence d'un squelette minéral perforé. Ils sont très abondants et absolument ubiquitaires en eau salée et parfois douce. Ceci leur confère une importance écologique mais également scientifique, du fait qu'ils sont utilisés comme des marqueurs biostratigraphiques en micropaléontologie et comme des indicateurs de l'âge et de la profondeur des océans et des anciennes températures de l'eau en paléo-océanographie (Habura et al., 2004b). Ils ont un mode de vie soit benthique (vivant sur ou dans le sédiment, sur le fond des lagunes et des mers) soit planctonique (en suspension dans la colonne d'eau, et particulièrement dans la zone euphotique mais parfois jusqu'aux abysses).

Enfin les Cercozoaires sont comptés pour la majorité comme des parasites dans la colonne d'eau (Chantangsi et al., 2010). Ils sont pour la plupart hétérotrophes (ne disposant pas de vrai cytostome) mais certains sont mixotrophes tel le cas des chlorarachniophytes. Ces organismes marins ont acquis leur capacité photosynthétique par endosymbiose secondaire d'une algue verte (Archibald, 2005). Les Cercozoaires comportent également des espèces communes en eau douces ou dans les sols, comme les euglyphydes, les haplosporides (endoparasites d'invertébrés d'eau douce et marine) et les plasmodiophorides (endoparasites de plantes ou d'algues straménopiles) (Cavalier-Smith et Chao, 2003).

La monophilie est bien appuyée par les phylogénies moléculaires quoiqu'aucune synapomorphie n'ait été déterminée pour l'ensemble des Rhizariens (Moreira et al., 2007). Par contre, les liens entre les groupes sont mal caractérisés. Ce super groupe constitue un groupe monophylétique (« SAR ») avec les straménopiles et les alvéolaires.

*Les Straménopiles (ou Heterokonta)*



**Figure 12:** Quelques lignées appartenant au phylum des Straménopiles.

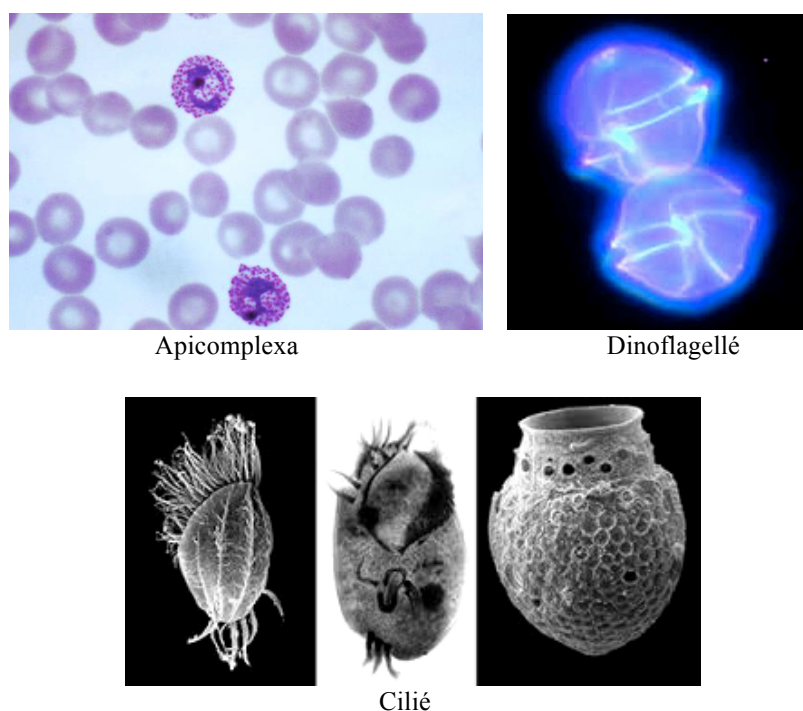


**Figure 13 :** Image à gauche : Tubercule de pomme de terre atteint par le mildiou. Image à droite : Symptômes de mildiou sur feuilles de pomme de terre. Maladie causée par l'oomycète *Phytophthora infestans*

Les straménopiles ou encore appelés hétérocontes regroupent des organismes caractérisés par l'existence, durant leur cycle, d'une cellule biflagellée avec deux flagelles différents (d'où la nomination alternative « hétéroconte »), un flagelle lisse et un autre plumeux à mastigonèmes (appendices protéiques) tubulaires tripartites. Ce super groupe est très divers et classé en troisième position, en termes de description détaillée des organismes, après les opisthocontes et les archaeplastides. Il regroupe plusieurs lignées, comme les bacillariophycées, les xanthophycées, les phaeophyceae ou encore les raphidophyceae. Il comprend des organismes unicellulaires hétérotrophes (Bicosoecida), des parasites

plasmodiales (oomycètes) comme *Phytophthora infestans*, qui cause le mildiou de la pomme de terre et affecte également les cultures de tomates et d'autres Solanacées (**Fig. 13**), et d'autres pluricellulaires (xanthophytes et phaeophytes). Les straménopiles forment un important réservoir de producteurs primaires mais sont également abondants en eau douce (chrysophytes) ou saumâtre (xanthophytes). Certains sont autotrophes tel le cas des diatomées. Ces dernières regroupent les organismes photosynthétiques les plus diversifiés. Elles comprennent au moins 12.000 espèces (*Kooistra et al., 2007*) et peuvent représenter jusqu'à 40% de la production primaire totale des océans (*Nelson et al., 1995 ; Mann, 199 ; Trèguer & Pondaven, 2000*) et jouent un rôle très important dans les cycles biogéochimiques, essentiellement dans le cycle du carbone (pompe biologique) et de la silice (*Sarthou et al., 2005*). D'autres sont hétérotrophes dont les oomycètes, majoritairement parasites (e.g. *Phytophthora infestans, Blastocystis sp.*), les Bicosoecida prédateurs ou encore les labyrinthulomycètes. Ces derniers sont essentiellement marins, généralement sous forme de parasites sur des algues ou des décomposeurs de matériel végétal mort (*Raghukumar, 2002*). Leur groupe comporte également des parasites d'invertébrés marins. Ce super groupe comprend également des organismes « les MAST : straménopiles marins ». Ces derniers sont abondants en milieu marin mais également en eau douce (*Massana et al., 2013*) et regroupent des organismes autotrophes et hétérotrophes, comprenant en grande partie des espèces de petite taille (pico et nano-eucaryotes). Ils sont considérés comme des bactériophages, représentant jusqu'à 20% des flagellés hétérotrophes (*Massana et al., 2011*).

La caractéristique principale de ce super groupe, le démarquant des autres, se rallie aux particularités des flagelles et aux phylogénies multigènes.



**Figure 14:** Quelques lignées appartenant au super-groupe des alvéolaires.

Le super groupe des alvéolés est formé d'organismes unicellulaires de formes diversifiées. Sa monophylie repose essentiellement sur les études moléculaires des ARNr18S. Il est distingué par l'existence de vésicules sous-membranaires, les alvéoles (d'où ce taxon tire son nom), amplement distribuées et pouvant servir de réservoir calcique. Il regroupe plusieurs groupes, parmi lesquels on compte les dinoflagellés, les ciliés, les apicomplexés, les perkonsozoa, les colpodelliles et les syndiniales (anciennement appelés « *Marine Alveolates* », MALVs). Les dinoflagellés, encore appelés Dinophytes ou Péridiniens sont très abondants et vivants dans les milieux marins et dulçaquicoles. Ils sont très diversifiés, essentiellement par leur comportement alimentaire. Certains sont hétérotrophes, d'autres mixotrophes ou encore photosynthétiques (Stoecker, 1987, 1999 ; Jones, 2000). L'hétérotrophie représente le régime alimentaire le plus commun (environ 50% des dinoflagellés sont hétérotrophes). Comme leur nom l'indique, ils sont dotés de deux flagelles insérés latéralement dans la cellule (Taylor, 1987), mais ils existent également sous forme benthique, vivant en épiphytes sur des macro-algues ou encore dans les interstices des substrats sableux (tout en pouvant migrer verticalement pour se reproduire). Ces micro-organismes jouent un rôle trophique important, plus particulièrement ceux

hétérotrophes herbivores, régulant les populations de micro-algues et contribuant au transfert des nutriments vers les niveaux trophiques supérieurs (métazoaires, poissons, mammifères marins, ...). On compte aujourd'hui environ 1500 espèces de dinoflagellés répertoriées, dont 160 sont d'origine benthique (Gomez, 2005). Certaines possèdent quelques uns des plus grands génomes nucléaires connus (entre 3000 et 215000 Mb), avec de grandes quantités de séquences d'ADN répétées (Hackett et al., 2005). Chez beaucoup d'entre elles, la cellule est protégée par une thèque formée de plaques rigides de cellulose incrustée de silice. Ces espèces dites thécates sont majoritairement représentées par les genres *Protoperdinium*, *Ceratium* ou *Gonyaulax*. Quant à celles dites athécates (absence de thèque) sont amplement représentées par les genres *Gymnodinium* et *Gyrodinium*. Ce super-groupe reste encore assez mal connu. Récemment lors d'une étude des Dinophyta du sédiment en Bretagne, plusieurs espèces, jusqu'alors non décrites, appartenant aux genres des *Prorocentrum*, *Sinophysis* et *Cabra* ont été trouvées. Les ciliés sont présents également dans les eaux douces, saumâtres et marines, où ils existent sous différentes formes, libres, nageuses, pédonculées, coloniales, parasitaires non pathogènes ou encore symbiotiques. Ils se caractérisent par la présence de cils vibratiles à leur surface et la possession de deux noyaux, un micronoyau reproducteur servant à la transmission du patrimoine génétique pendant la mitose et un macronoyau végétatif permettant la synthèse des protéines (Prescott, 2000). Ces micro-organismes sont hétérotrophes et prédateurs, ils peuvent se nourrir de particules organiques, de bactéries, d'autres ciliés, de flagellés ou même d'animaux microscopiques. En termes de nombre et de biomasse, les ciliés dominent les communautés microzooplanctoniques (Dolan et al., 2000; Dolan et Mckeon, 2005; Vladstein et al., 2004). Quant aux dinoflagellés, ils sont plus abondants, représentant plus de la moitié de la biomasse totale des alvéolaires hétérotrophes (ciliés et dinoflagellés), dans les zones où les blooms de diatomées se produisent (forte concentration en chlorophylle *a*) (Hansen et al., 1991; Verity et al., 2002; Horner et al., 2005). Contrairement en hiver ou dans les zones peu productives, ils sont peu présents représentant un faible pourcentage de la biomasse totale des protistes hétérotrophes (~1% en hiver en mer de Wadden (Tillman & Hesse, 1998), 3% dans le golf de Trieste (Umani & Beran, 2003)).

Les apicomplexés ou encore appelés sporozoaires, sont des organismes presque tous parasites intracellulaires obligatoires des métazoaires. Les agents pathogènes les plus connus sont les *Plasmodium falcifarum* (responsable de la malaria) et *Toxoplasma sp.*

(responsable de la toxoplasmose). Certains organismes possèdent un ADN chloroplastique vestigial laissant penser qu'ils dérivent d'un ancêtre photosynthétique, ou qu'ils ont acquis et recyclé certains gènes d'un organisme photosynthétique. Ils sont caractérisés par leur complexe apical, leur permettant l'attachement et la pénétration de la cellule hôte. Les colpodellides possèdent également un complexe apical servant à l'attachement à d'autres organismes mais non à la pénétration. Il est plutôt utilisé pour l'aspiration du cytoplasme des proies. Ce sont des flagellés libres et prédateurs. Les Perkinsea sont composés de protistes parasites, particulièrement des mollusques et des dinoflagellés. Le plus connu est *Perkinsus sp.*, responsable de la «Dermo-disease» causant des mortalités massives de bivalves. Ces organismes sont très peu détectables (au niveau de la séquence) dans l'écosystème (marin ou lacustre). Il a été démontré récemment que la diversité des Perkinsea en milieu marin est largement plus importante mais qu'ils passaient inaperçus aux séquençages environnementaux à cause de leur localisation principalement dans le sédiment (Chambouvet et al., 2014).

Enfin, les syndiniales qui se montrent très abondants dans tous les milieux marins (1/3 des séquences totales des alvéolaires, correspondant aux groupes des séquences environnementales MALV I et II (Lopez-Garcia et al., 2001)). En plus de leur large abondance, ils sont caractérisés par un large spectre d'hôtes potentiels (radiolaires, ciliés, dinoflagellés et copépodes) (Massana et al., 2011). Pour l'instant, ces parasites intracellulaires n'ont été observés qu'en milieu marin (Guillou et al., 2008).

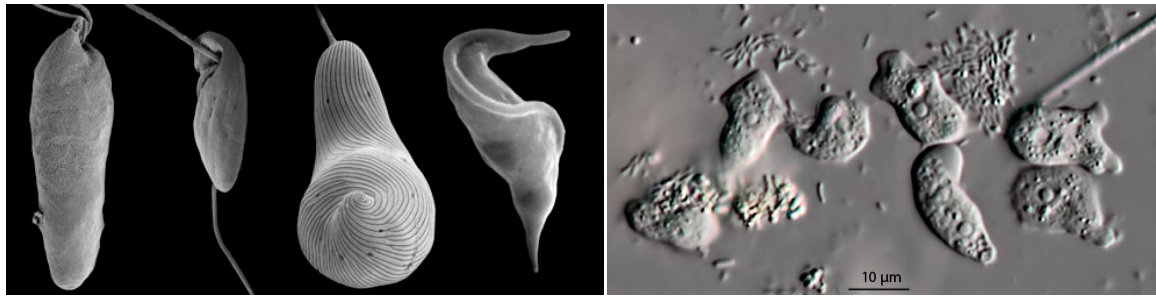
### *Les Haptophytes et les Cryptophytes*



**Figure 15:** Quelques espèces appartenant au phylum des Cryptophyta (1 et 2) et des Haptophyta (3).

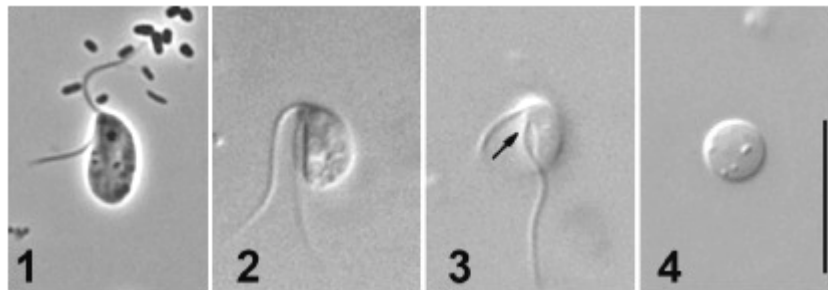
Ce sont principalement des algues à chlorophylle c. Quoique ces deux groupes (s'il s'agit vraiment de groupes) partagent avec les Straménopiles une origine commune, il existe peu de preuves génétiques affirmant que leurs génomes nucléaires sont les mêmes (*Burki et al., 2007 ; Hackett et al., 2007 ; Burki et al., 2014*). Les Cryptophytes sont des organismes généralement de petite taille (de 2 à 10 µm de diamètre) et se rencontrent dans tous les milieux aquatiques, allant des milieux océaniques aux eaux douces en passant par les eaux interstitielles des milieux terrestres humides. Certains sont parasites intestinaux de métazoaires, endosymbiotes de Dinophytes ou vivant en symbiose dans l'ectoplasme des radiolaires (*Patron et al., 2007 ; Fehling et al., 2007*). Les Haptophytes sont caractérisés par leur haptone d'où ils tirent leur nom, appendice filiforme contenant des microtubules, dont la taille varie selon l'espèce, permettant l'adhésion à un substrat, le déplacement de particules voire la capture de proies. On compte aujourd'hui plus que 300 espèces (*Gran-Stadniczenko et al., 2017*). Nombreuses sont tropicales, certaines vivent en eau douce et d'autres groupes fossiles existent. Ils incluent les coccolithophoridés, connus également sous le nom des coccolithes, se sont des organismes unicellulaires exclusivement marins qui vivent principalement dans les régions froides comme l'Atlantique nord. Ils font partis du nanoplancton, leur taille est comprise entre 5 et 50 microns et se caractérisent par leur squelette calcaire comme chez les diatomées. *Emiliana huxleyi* est comptée parmi les espèces les plus étudiées en raison de sa capacité de produire des efflorescences importantes affectant significativement les paramètres physiques (température et conditions optiques) du milieu où elles se trouvent (*Buitenhuis et al., 1996*). Au sein des haptophytes, on trouve également les Prymnésiales englobant le genre *Phaeocystis*. Ce dernier largement distribué dans l'environnement marin (*Schoemann et al., 2005*) est très reconnu, particulièrement par la mousse blanche qu'il produit sur le littoral lors des efflorescences algales quasi-monospécifiques atteignant jusqu'à  $10^7$  cells.L<sup>-1</sup> et 10 mg.C.L<sup>-1</sup> en masse carbonique (*Cadée & Hegeman, 1986 ; Schoemann et al., 2005*). Ces blooms sont récurrents au printemps (de mars à juin) sur le littoral européen de la Manche et de la mer du Nord.

*Les Excavés*

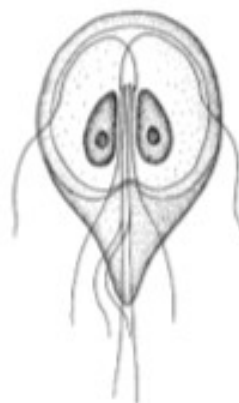


Euglenozoa  
(Leander and Alastair, 2008)

*Naegleria gruber*  
Heterolobosea  
(<https://www.arcella.nl/naegleria-gruberi>)



*Andaluca godoyi* (1. cellule fixe, 2. Cellule active, 3. Cellule active avec une flèche indiquant la rainure, 4. Cyste)  
Jakobida  
10µm  
(Lara et al., 2006)



*Giardia Lamblia*  
Métamonada  
20 µm  
(Lecointre & Le Guyader, 2006)

**Figure 16:** Quelques espèces appartenant au super-groupe des Excavés.

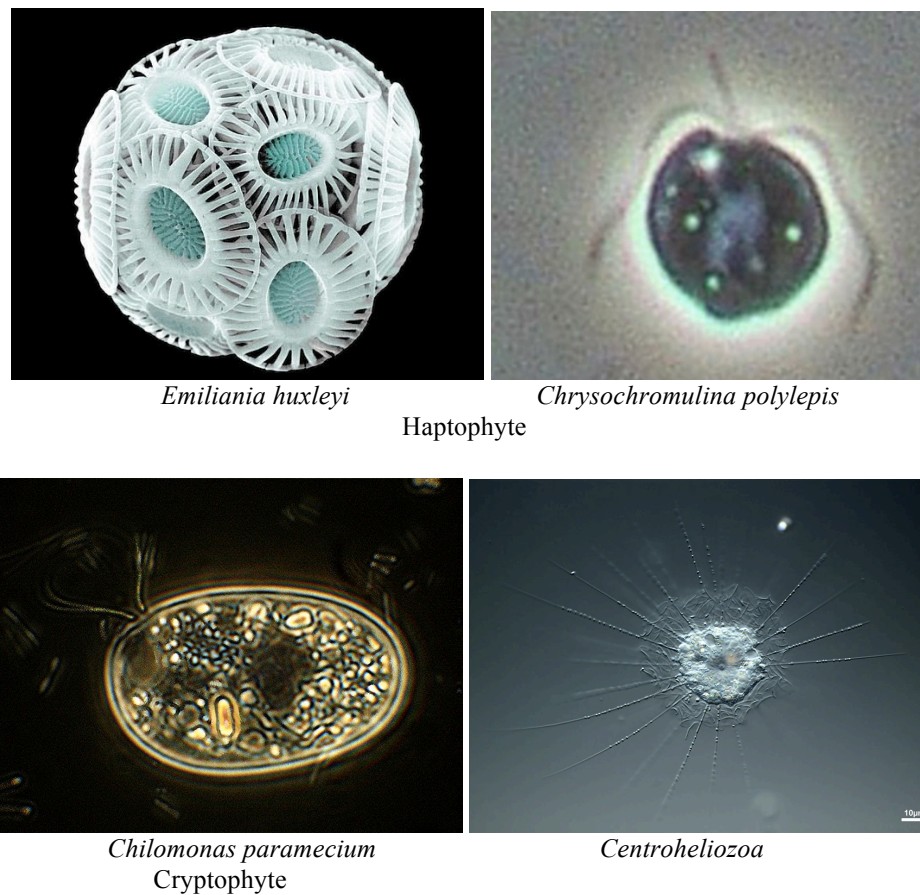
Les excavés (Excavata) ou encore appelés Excavobionta sont des organismes unicellulaires hétérotrophes, possédant généralement un flagelle (Simpson, 2003). Ils sont caractérisés par



la présence du cytostome (zone creusée) leur permettant l'ingestion de fines particules alimentaires. Ils sont pour la plupart anaérobies ou microaérophiles, en possession d'organites particuliers et spécialisés : les mitosomes et les hydrogénosomes (e.g. *Diplomonade*). Certains sont amiboïdes tels que les Heterolobosea. Ces derniers sont abondants, ubiquistes mais leur importance écologique est toujours mal assimilée. Ce sont des prédateurs bactériens du sol ou d'eau douce, et pathogènes facultatifs de l'homme (*Naeglerie fowleri*) (Maclean et al., 2004). On trouve également des excavés commensaux, symbiontes, possédant une mitochondrie (e.g. *Trypanosoma brucei*) ou des mitosomes au lieu de mitochondrie (organismes dits amitochondriaux, ex. *Giardia intestinalis*). Quelques excavés possèdent des plastes et sont photosynthétiques suite à une endosymbiose secondaire avec une algue verte, comme le cas de certains organismes appartenant aux Euglénoides. Ils sont également hétérotrophes et mixotrophes, se rencontrant dans les écosystèmes pélagiques et benthiques (Buck et al., 2000). En plus des Euglénoides, les excavés mitochondriaux incluent également les Heterolobosea (précédemment cités) et les Jakobida (Simpson et al., 2006). Ce sont des petits prédateurs de bactéries vivants sous forme libre et sont démarqués par leur morphologie mitochondriale qui varie (Lara et al., 2006).

En ce qui concerne les amitochondriaux, ont été considérés auparavant comme des Archaezoa, suggérant que ceux-ci dériveraient des premières lignées d'eucaryotes prémitochondriées (Cavalier-Smith et Chao, 1996). Cette hypothèse a été rejetée après la découverte de gènes ancestraux mitochondriaux dans leurs génomes et des empreintes de mitochondries très réduites (Embrey et Martin, 2006). Or, ces organismes apparaissent toujours comme les toutes premières branches dans les arbres moléculaires enracinés (Arisue et al., 2005) quoique ce positionnement soit souvent considéré comme un artefact d'attraction des longues branches. (Philippe et Germot, 2000).

*Incertae sedis* (positions incertaines)



**Figure 17:** Quelques espèces dont la position dans l'arbre phylogénétique eucaryote est toujours incertaine (Leconte & Le Guyader, 2006 ; Walker et al., 2011a).

Plusieurs protistes cultivés, plus ou moins abondants ou variés sont encore mal positionnés dans l'arbre phylogénétique eucaryote : 230 ont été répertoriés en 1999, ce nombre a diminué à 204 en 2005 (Adl et al., 2005). Ils sont pour la plupart des organismes hétérotrophes (e.g. Katablepharida, Telonemida), ou des amibes (e.g. Centroheliozoa), ou encore des parasites ou endosymbiontes de dinoflagellés (e.g. Cryptophytes). Les études de diversité eucaryote basées sur des techniques moléculaires n'ont pas souligné de nouveau super-groupe, contrairement aux bactéries et archées (Lopez-Garcia & Moreira, 2008 ; Massana & Pedros-Alio, 2008). Par contre, l'exploration de la diversité via ces approches moléculaires a montré une grande part de diversité reste encore méconnue, appartenant pour la plupart aux super- groupes précédemment décrits. Il a été démontré qu'une grande partie de séquences environnementales ne ressemblaient à aucune séquence préalablement

identifiée ou même observée dans des travaux de recherche précédents (*Shi et al., 2009 ; Triado-Margarit & Casamayor, 2012*), prouvant qu'il reste encore des lignées majeures d'eucaryotes à explorer. Comme le cas par exemple des Syndiniales (MALV I et II) qui sont représentés par peu d'organismes cultivés (*Diez et al., 2001 ; Moon-van der Staay et al., 2001 ; Lopez-Garcia et al., 2001 ; Guillou et al., 2008*) et des MAST (*Massana et al., 2013*). Il existe également une grande partie de la diversité pouvant être formée de nano- et pico-eucaryotes à explorer. En raison de leur petite taille (<1%), leur identification via des observations microscopiques et leur distinction des bactéries s'avèrent très compliquées voire impossibles.

### **III. PLACE ET ROLE DES PROTISTES DANS LES ECOSYSTEMES COTIERS**

#### **III.1. PLACES ET ROLES DES PROTISTES DANS LES ECOSYSTEMES MARINS**

Dans ce projet de thèse, je me suis focalisée essentiellement sur l'étude des protistes eucaryotes. Ces derniers sont très abondants et très courants des communautés planctoniques. Ils présentent une grande diversité de types trophiques (phototrophie, hétérotrophie et mixotrophie) et de mode de vie (libres, parasites ou symbiotes mutualistes). Ils existent dans tous les milieux et habitats (marins, dulcicoles ou terrestres) dont les milieux extrêmes (océan profond, *Lopez-Garcia et al., 2001* ; les sources hydrothermales sous marines, *Lopez-Garcia et al., 2001* ; des environnements acidophiles, *Amaral Zettler et al., 2002* ou la glace, *Bachy et al., 2011 ; Piwosz et al., 2013*). Certains possèdent la capacité de former du carbonate de calcium (CaCO<sub>3</sub>, Foraminifères et les Coccolithophoridés) ou de la silice (diatomées, radiolaires). D'autres produisent également des toxines qui peuvent entraîner, chez l'homme, la mort par paralysie et créent des troubles digestifs et des hépatites (*Dinophysis, Alexandrium, Gymnodinium, Gonyaulax, Gambierdiscus*, etc.). Ces toxines agissent également sur les autres organismes du milieu (e.g. neurotoxines de dinoflagellés toxiques pour les métazoaires, acide acrylique bactéricide de certains haptophytes).

Constituant les bases du réseau trophique, les protistes jouent un rôle majeur dans les cycles biogéochimiques (par exemple, leur participation aux cycles de carbone ; qu'ils piègent

dans la matière organique ou en précipitant des carbonates ; et du soufre) et les réseaux trophiques. Ils constituent une part importante de la biodiversité et peut-être plus encore de la « biodiversité fonctionnelle » : sur environ 300 000 espèces estimées existantes, les 2/3 pourraient être assez largement distribuées, dans les océans notamment et 100 000 environ pourraient avoir une distribution plus restreinte (*Wilhelm Foissner, 2008*). Ils rendent d'importants services écologiques, en particulier pour le climat, par exemple via leur production de diméthyl-sulfoniopropionate (DSMP) transformé par les bactéries en diméthylsulfide (DMAS) qui, libéré dans l'atmosphère, entraîne la nucléation des nuages, et de là augmente l'albédo terrestre (*Jordan & Chamberlain, 1997 ; Simó, 2001*), pour également l'épuration de l'eau, la régulation du CO<sub>2</sub> et des minéraux dans l'eau, la pêche, etc.

## **III.2. FOCUS SUR LE ROLE DES PROTISTES DANS LES RESEAUX TROPHIQUES** (particulièrement planctoniques)

### **III.2.1. Préambule**

Pour tout organisme vivant (autotrophes ou hétérotrophes), l'assimilation des matières énergétiques est d'ordre vital. Elle lui permet de maintenir sa survie et ses capacités de reproduction (Stearns 1992). D'après Elton (1927), les interactions entre les organismes d'une communauté sont principalement d'ordre trophique. En 1942, Lindeman précise que l'organisation des écosystèmes repose principalement sur les flux d'énergie entre les niveaux trophiques d'une chaîne alimentaire. Le mot « écosystème » a été utilisé pour la première fois en 1935 par Transley, et en 1969, Odum définit ainsi un écosystème comme *"une unité d'organisation biologique composée de tous les organismes dans une zone donnée (c'est-à-dire, la « communauté »), interagissant avec l'environnement physique, de sorte qu'un flux d'énergie conduit à une structure trophique caractéristique et aux cycles de matières dans le système"*. Dans un premier temps, les écologistes se sont intéressés particulièrement aux liens trophiques et à l'importance relative des différents transferts d'énergie pouvant exister entre les compartiments biologiques de ces écosystèmes (*Paine, 1980*) afin de décrypter leur structure et leur fonctionnement général.

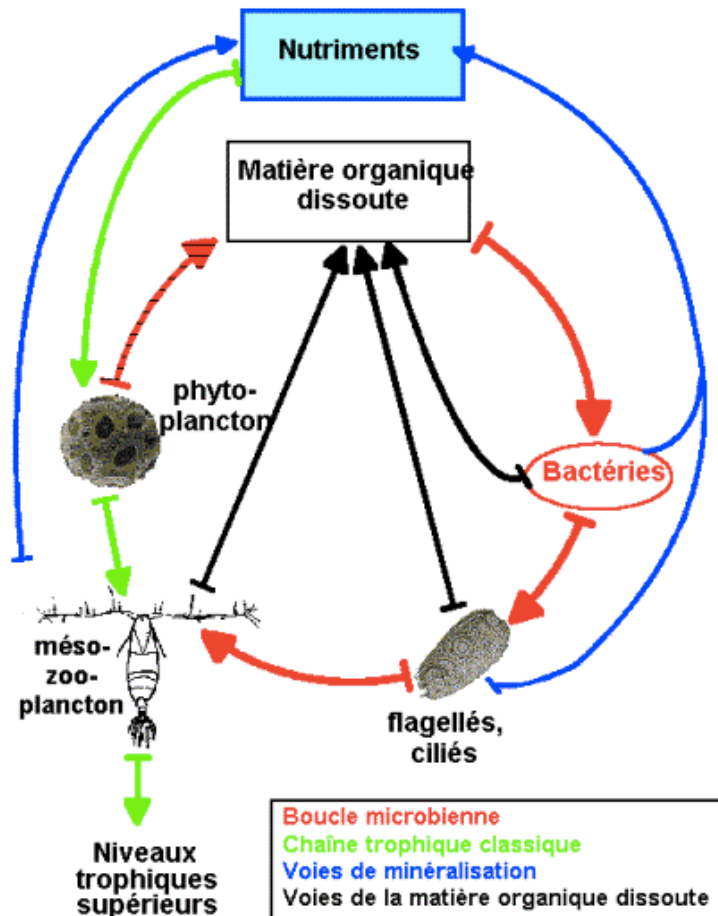
### III.2.2. Les réseaux trophiques planctoniques

L'ensemble des écosystèmes est caractérisé par une production primaire et un transfert de matière organique des organismes autotrophes vers les organismes hétérotrophes. Ceux-ci dépendent essentiellement des paramètres environnementaux du milieu ainsi que la nature des eaux. Au niveau de la chaîne trophique pélagique, le plancton et les flux de matières régissent et conditionnent majoritairement le fonctionnement de l'écosystème : source de matière organique, exportation vers le benthos ou les écosystèmes connexes, efficacité trophique vers les maillons supérieurs (*Legendre & Rassoulzadegan, 1995; Sintes et al., 2004 ; Leguerrier, 2005 ; Marquis et al., 2007*). La compréhension du fonctionnement du réseau trophique peut ainsi donner une idée des caractéristiques émergentes d'un écosystème (efficacité éco-trophique, exportation, recyclage, etc.).

Deux modèles de transferts de matière organique ont été définis à la base du réseau trophique planctonique (**Fig. 18**)

- La chaîne herbivore (*Cushing, 1989*) où les producteurs primaires (phytoplancton), essentiellement de grande taille ( $>5\mu\text{m}$ ), constituent une source d'alimentation aux consommateurs primaires zooplanctoniques également de grande taille. Ces derniers sont à leur tour consommés par des juvéniles de poissons.
- La boucle microbienne (*Rassoulzadegan, 1993*) où la production primaire est caractérisée par du phytoplancton de petite taille ( $<5\mu\text{m}$ ) consommé par du zooplancton de petite taille (flagellés). Ces microorganismes relarguent de la matière organique dissoute et de l'ammonium exploités par les bactéries. Ces bactéries sont consommées également par le microzooplancton, à son tour consommé par le zooplancton de grande taille, lui-même exploité par les juvéniles de poisson. Dans ce système presque clos, le phytoplancton et les bactéries rentrent en compétition pour l'utilisation de l'ammonium.

Dans les écosystèmes marins, il peut y avoir la dominance de la chaîne herbivore comme de la boucle microbienne. La structure du réseau trophique dépend principalement des conditions hydrologiques et la disponibilité des nutriments dans le milieu.



**Figure 18** : Représentation simplifiée de la boucle microbienne (en rouge) et la chaîne trophique classique (herbivores, en vert). (Source : Institut de Recherche pour le Développement, <http://www.atolls-polynesie.ird.fr/commatoll/bactpl.htm>)

### III.2.3. Focus sur le rôle écologique des protistes eucaryotes

A l'exception de la chimiolithotrophie, uniquement réalisable par les protéobactéries, les Aquificales, les archaea méthanogènes et les bactéries neutrophiles oxydant le fer, le reste des voies métaboliques existent chez les protistes eucaryotes marins. Les phototrophes sont considérés comme les premiers contributeurs à la production primaire globale (*Field et al., 1998*) et les seuls pouvant interagir à plusieurs niveaux du réseau trophique en raison de leurs diverses tailles.

Les picoeucaryotes représentent la fraction du phytoplancton dont la taille est comprise entre 0.2 et 3  $\mu\text{m}$  (*Vaulot et al., 2008*). Ils sont très abondants et particulièrement présents dans les régions oligotrophiques et contribuent significativement à la biomasse, à la production primaire océanique, dans le réseau trophique et les cycles biogéochimiques (*Massana et al., 2011*). On compte aujourd'hui une quarantaine d'organismes décrits. Leur

répartition ainsi que leurs contributions à la biomasse chlorophylliennes sont toujours mal connues. Ils sont composés entre 70 et 80% d'organismes autotrophes (*Jurgens & Massana, 2008*), leur abondance dans les zones oligotrophiques est de l'ordre de  $1 \text{ à } 3 \cdot 10^3 \text{ cell.ml}^{-1}$  et pouvant atteindre jusqu'à  $10^5 \text{ cell.ml}^{-1}$  dans les zones côtières (*Sanders et al., 2000 ; Li et al., 2009*). Les picoeucaryotes peuvent donc représenter une large portion de la biomasse phytoplanctonique et de la production primaire, comme par exemple dans l'Océan Indien où ils peuvent contribuer jusqu'à 50% de la biomasse phytoplanctonique (*Not et al., 2007*) ou dans le nord de l'Atlantique où ils peuvent représenter jusqu'à 34% de la production primaire (*Jardillier et al., 2010*).

Les protistes phytoplanctoniques dont la taille est supérieure à  $20\mu\text{m}$  habitent majoritairement les eaux eutrophes, généralement représentés par les bacillariophycées (diatomées) et les dinoflagellés autotrophes (*Li, 2002*). Les diatomées représentent le constituant majeur du phytoplancton et contribuent à 50% de la production primaire océanique globale (*Mann, 1999*). Ces organismes jouent un rôle primordial dans la vie des écosystèmes marins en raison de leur positionnement à la base des réseaux trophiques alimentaires de nombreuses espèces et également au sein de la pompe biologique et ainsi dans le cycle du carbone (*Falkowski et al., 1998; Ducklow et al., 2001*). Plus de 200.000 organismes sont répertoriés (*Round et al, 2007*).

Les protistes hétérotrophes jouent un rôle essentiel dans le transfert du carbone vers les réseaux trophiques supérieurs en étant les principaux prédateurs des cyanobactéries (*Pomeroy et al., 1974*), des bactéries (*Azam et al., 1983*) ainsi que d'autres protistes (*Jeong et al., 1999*). Ils s'alimentent également du phytoplancton (*Landry et al., 1997*) et du méso-microplancton (*Hansen & Calado, 1999*).

Les pico- et nanoflagellés hétérotrophes sont les principaux prédateurs du picoplancton dans les milieux marins et lacustres. Leurs abondances sont d'environ  $3 \cdot 10^2 - 3 \cdot 10^3 \text{ cell.ml}^{-1}$ . Ils représentent environ 30% de la biomasse hétérotrophe ( $<20 \mu\text{m}$ ). Dans le nanoplancton, les nanoflagellés hétérotrophes occupent le premier rang en termes d'espèces et de biomasse (*Fuduka et al., 2007*). Les picoflagellés bactérivores (*Jurgens & Massana, 2008*) contribuent significativement dans le processus de la boucle microbienne (stabilité du stock bactérien, transfert de la matière organique vers les niveaux supérieurs, reminéralisation du phosphore et de l'azote (*Pomeroy et al., 2007*)).

Le microzooplancton hétérotrophe est un groupe dominé largement par les alvéolaires, essentiellement les dinoflagellés et les ciliés hétérotrophes (*Sherr & Sherr, 1994*). Ces derniers présentent un large spectre de proies partant des bactéries (*Azam et al., 1983*), aux microphytoplanctons (*Calbet & Landry, 2004*), jusqu'au microzooplancton (*Jeong, 1999*). Par leur capacité d'ingérer différents types de proies de diverses tailles, les dinoflagellés et les ciliés occupent une place importante dans le réseau trophique aussi bien en tant que prédateurs, que de proies de niveaux supérieurs. Leur positionnement clé dans le réseau trophique leur attribue un rôle important dans les cycles biogéochimiques, plus particulièrement dans le cycle du carbone et de la silice (*Sherr & Sherr, 1994*).

Il existe également un autre mode trophique chez les protistes assez complexe, la mixotrophie. Ce mode s'établit généralement selon un gradient des conditions trophiques (disponibilité des proies et des nutriments) et physiques (présence d'énergie lumineuse) (*Jones et al., 2000*). La mixotrophie peut s'établir via plusieurs processus : endosymbiose secondaire ou tertiaire (*Stoecker, 1999*) ; ectosymbiose (*Gas et al., 2009*) ; mutualisme (*Caron & Swamberg, 1990*) ; parasitisme ou commensalisme (*Caron, 2000*). Le couplage des deux modes trophiques hétérotrophie et autotrophie participe à l'amélioration de la compétitivité des organismes dans les milieux pauvres en nutriments par synthèse de la matière organique à partir de matière minérale comme le dioxyde de carbone. Ceci n'empêche pas que dans certaines régions eutrophes, les mixotrophes se trouvent en abondance comme le cas dans les îles de Kerguelen (*Christaki et al., 2008*).

Dans les écosystèmes marins, les protistes, par leur large diversité et leur contribution à différents niveaux du réseau trophique et dans les cycles biogéochimiques, occupent une grande place dans les recherches en écologie.



## **PARTIE 2 : SITES D'ETUDES**

### **I. MANCHE ORIENTALE**

La majeure partie des résultats de ce travail provient du site SOMLIT en Manche Orientale. La Manche Orientale, est un écosystème méso-eutrophique qui constitue un site d'étude privilégié de par ses particularités hydrologiques et biologiques, et un endroit idéal pour travailler sur la compréhension de la diversité, en raison du contexte environnemental, qui est à la fois relativement complexe et assez bien compris via plusieurs projets d'observations classiques (voir section I.3).

#### **I.1. Caractéristiques physiques**

##### ***Géographie et climat***

La Manche Orientale est une mer épicontinentale dont la profondeur ne dépasse pas les 175m. Elle s'étend de la Baie de Seine jusqu'au détroit du Pas-de-Calais. En Manche Orientale les masses d'eau transitent entre l'Océan Atlantique et la Mer du Nord et se caractérisent par un fort régime marégraphique semi-diurne (2 marées/jour) et mégatidal (marnage > 5m). La Manche se divise en Manche Orientale et Occidentale. Leur séparation est fixée à l'isobathe 60m, de Start Point au Cap de la Hague. La zone d'étude est localisée en Manche Orientale, au sud de Boulogne sur mer (**Fig. 19**).

Cette région a un climat tempéré, est influencé par les courants atmosphériques humides et frais provenant de l'Atlantique ou plus sporadiquement de la Mer du Nord. Ce type de climat est caractérisé par des températures moyennes, comprises entre 20 et 25 °C en été et 0 et 5°C en hiver, ainsi que par des précipitations mensuelles importantes de l'ordre de 40 à 60 mm de février à août et de 70 à 90 mm entre septembre et janvier. Les précipitations moyennes annuelles sont de l'ordre de 800-1000 mm/an. Ces précipitations influencent fortement les apports fluviaux. La répartition des vents se caractérise par une nette dominance des directions de secteur Sud-ouest, avec la présence marquée au printemps de vents de Nord-est. Leur régime est constant à l'échelle pluriannuelle.



**Figure 19:** Site d'étude Manche Orientale – Situation géographique. ● Station de prélèvements.

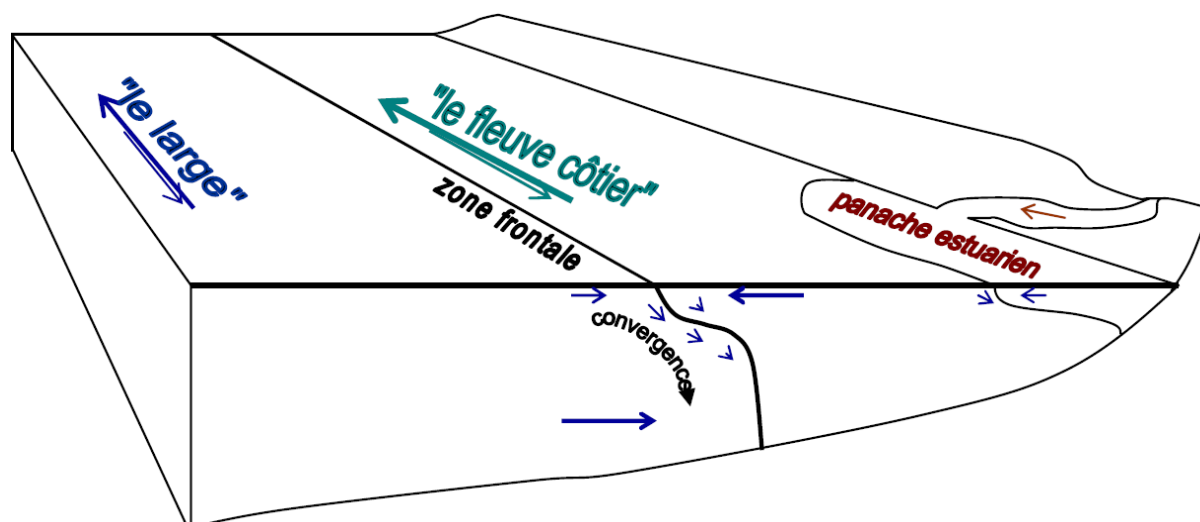
### *Caractéristiques marégraphiques et hydrologiques*

La Manche Orientale est une mer semi-fermée où l'hydrodynamisme est sous l'influence des marées, le vent et les vagues. Le marnage dans le Pas-de-Calais varie de 7m en vives eaux et à moins de 5m en mortes eaux. Il est considéré parmi les plus importants du monde. L'intensité des courants instantanés varie de 120 à 150 cm s<sup>-1</sup> en marée de vive-eau moyenne. Ces courants sont parallèles aux côtes (Brylinski, 1993), portant globalement vers le nord-est au moment du flot et vers le sud-ouest au cours du jusant. Il en résulte une dérive résiduelle d'une vitesse moyenne de 2.7 milles nautiques par jour (Salomon et Breton, 1993), permettant le déplacement de la masse d'eau de la Manche vers la Mer du Nord.

Ce contexte marégraphique favorise la création d'une masse d'eau côtière permanente bien individualisée, appelée « le fleuve côtier » (Fig. 20 ; Brylinski et al., 1991). Cette masse d'eau est alimentée par les apports fluviaux entre la baie de Seine et le détroit du Pas-de-Calais (et notamment à ceux de la Somme, de la Canche et de l'Authie). Sa circulation est régie par le courant résiduel de marée dirigé vers la Mer du Nord. Ce dernier est riche en éléments transportés (matières en suspension : MES (Guéguéniat et al., 1995; Lafite et al., 2000 ; cellules algales, Hault et al., 1994 ; sels nutritifs, Bentley et al., 1993), ce qui pourrait le rendre responsable d'une partie non négligeable des flux de matières en suspension (e.g. nutriments, polluants, etc.) vers la Mer du Nord (Brylinski et al., 1991 ;

Salomon et Breton, 1993). La Manche Orientale constitue donc une zone de transition entre celle occidentale et la région sud de la Mer du Nord (Velegrakis et al., 1999) et est considérée globalement comme une zone de transfert entre les eaux atlantiques et celles de la mer du nord (Salomon et Breton, 1993).

Il existe en revanche, dans la zone d'étude, un gradient côte-large prononcé pour l'ensemble des paramètres hydrologiques, tels que la salinité, la masse volumique, la turbidité (Brylinski et al., 1984 ; Bentley, 1985 ; Dupont et al., 1991 ; Genitsaris et al., 2016), la concentration en éléments nutritifs (Gentilhomme et Lizon, 1998), l'abondance phytoplanctonique (Brunet et al., 1996 ; Hoch et Garreau, 1998 ; Lizon et al. 1998 ; Breton et al., 2000), zooplanctonique (Brylinski et al., 1988 ; Breton, 2000) et bactérienne (Maillard-Quitshoudt, 1988) lié au fleuve côtier. Les masses d'eaux côtières se caractérisent par des biomasses phytoplanctoniques et production primaire plus importantes par rapport à celles du large (Brylinski et al., 1984 ; Brunet et al., 1992 ; Gentilhomme et Lizon, 1998) mais également par une composition taxinomique phytoplanctonique (Brunet, 1994 ; Breton et al., 2000 ; Shapira, 2005) et zooplanctonique (Brylinski et al., 1988 ; Breton, 2000 ; Cottonnec, 2001) différente.



**Figure 20** : Représentation schématique du fleuve côtier et de la zone frontale de la Manche Orientale (Brylinski et al., 1991)

## I.2. Caractéristiques biologiques et diversité des eucaryotes unicellulaires

Le site de la Manche Orientale a été bien étudié à la fois en termes de recherches sur les composantes phytoplanctoniques (*Brunet et al., 1996; Lizon et al., 1998; Gentilhomme et Lizon, 1998; Breton et al., 2000; Lizon, 2002; Vantrepotte, 2002; Brunet et Lizon, 2003; Seuront et al., 2005*), zooplanctoniques (*Brylinski et al., 1988; Breton et al., 1999; Cotonnec et al., 2001, 2003; Devreker et al., 2005*) en relation avec l'hydrodynamisme (*Brylinski et Lagadeuc, 1990; Brylinski et al., 1991; Salomon et Breton, 1993; Lagadeuc et al., 1997*) et les conditions climatiques (*Seuront et Souissi, 2002*), ainsi que bactériennes (*Hubas et al., 2006; Lamy et al., 2009*).

Cet écosystème méso-eutrophe est connu par son importante productivité (*Brylinski et al., 1996*) avec des productivités photosynthétiques pouvant aller de  $5\text{mg C.m}^{-3}.\text{h}^{-1}$  au large à  $40\text{mg C.m}^{-3}.\text{h}^{-1}$  à la côte et des stocks phytoplanctoniques se situant entre 20 et  $50\ \mu\text{g Chl-a.l}^{-1}$  (*Brunet et al., 1996; Gentilhomme et Lizon, 1998; Vantrepotte, 2003; Seuront et al., 2006*). Les abondances en phytoflagellés se situent entre 11 et 98% des abondances phytoplanctoniques totales (57% en moyenne par an). Les diatomées représentent 85% de la biomasse phytoplanctonique totale (*Breton et al., 2000; Grattepanche et al., 2011; Lefebvre et al., 2011, Hernández Fariñas et al., 2013*). Une saisonnalité de type « diatomées ; bloom ; diatomées » est remarqué à peu près annuellement (*Breton et al., 2000; Gómez et Souissi, 2007; Lamy et al., 2009; Monchy et al., 2012; Houliez et al., 2013*).

Cette région est annuellement sujette à une prolifération de *Phaeocystis globosa* au printemps (e.g. *Brylinski et al., 1984; Brunet et al., 1992, Brunet et al., 1996; Hoch et Garreau, 1998; Breton et al., 2000; Monchy et al., 2012*). Cette espèce constitue à elle seule 90% de l'abondance cellulaire totale du microphytoplancton et environ 50% de sa biomasse (par exemple : *Grattepanche et al., 2011 a, b; Breton et al., 2017*).

Les observations conduites en Manche Orientale, et majoritairement réalisées par des outils classiques (par exemple : *Gomez and Souissi, 2007; Schapira et al., 2008; Lefebvre et al., 2011; Grattepanche et al., 2011a, b; Hernandez-Farinas et al., 2013*), ont permis de mettre en évidence l'existence d'un lien entre les proliférations massives de phytoplancton observées de façon récurrente et les teneurs en sels nutritifs. En effet, ces proliférations sont initiées par un déséquilibre dans les sels nutritifs, caractérisé par un excès d'azote et une

limitation en silicate (par exemple : *Peperzak et al., 1998* ; *Schapira et al., 2008*). Ces proliférations sont caractérisées par la dominance de l'espèce coloniale *Phaeocystis globosa* (prymnesiophyte) et par un consortium de diatomées, dont certaines potentiellement toxiques (*Sazhin et al., 2007*). Lamy et al. (2009) ont montré l'impact significatif de ces augmentations subites de biomasse algale sur la structure et l'activité des communautés procaryotes hétérotrophes. Cet impact est traduit par une augmentation significative de l'abondance et de l'activité de certains groupes bactériens en réponse à l'accumulation de la matière organique pendant et après le bloom phytoplanctonique.

Ces proliférations ont également un impact significatif sur toute la communauté planctonique et benthique ainsi que sur les transferts de carbone (*Grattepanche et al. 2011a, b*). Ces études ont permis d'identifier le rôle des dinoflagellés comme consommateurs principaux du phytoplancton, en particulier de diatomées et de colonies de *P. globosa*, et ont également suggéré le rôle des copépodes en tant que prédateurs des protistes hétérotrophes (top-down control).

En Manche Orientale, la structure et la succession saisonnière des communautés microbiennes ont fait l'objet de plusieurs études par séquençage à haut débit basé sur l'ADN (*Monchy et al., 2012* ; *Christaki et al., 2014* ; *Genitsaris et al., 2015, 2016*). Ces mêmes approches ont été utilisées pour étudier la structure des communautés bactériennes marines et son évolution en fonction des variables environnementales. Elles ont montré que les paramètres environnementaux expliquent la plus grande partie de la variabilité saisonnière des bactéries, suggérant que les changements temporels des paramètres abiotiques influencent davantage la structure des communautés bactériennes que les relations entre taxons (*Gilbert et al., 2009* ; *Gilbert et al., 2012*). Le résultat d'autres travaux conduits dans d'autres écosystèmes rejoignaient cette conclusion (par exemple : *Fuhrman et al., (2006)* et *Salter et al., 2014*). *Genitsaris et al. (2015)*, a démontré via des analyses canoniques, que les variables environnementales expliquaient seulement ~30% de la structure des communautés eucaryotes microbiennes et leur succession en Manche Orientale. De plus, des analyses en réseau et de cooccurrences ont montré peu de corrélations significatives entre les OTUs (prises individuellement) et les variables environnementales (*Christaki et al. 2017*). Ainsi, contrairement aux bactéries, les relations entre-taxa (interactions trophiques) semblent être le principal facteur régissant la structure des communautés eucaryotes et leurs successions temporelles (*Genitsaris et al., 2015, 2016*). La complexité

des traits trophiques entre les différents groupes taxonomiques chez les protistes pourrait être une explication plausible de l'importance accrue des relations inter-taxa chez les eucaryotes. En effet, la similarité fonctionnelle entre groupes taxonomiquement distincts est répandue chez les bactéries, tandis que pour les eucaryotes, la proximité phylogénétique ne correspond pas forcément à une fonction commune écologique. Ainsi, en plus d'étudier l'impact des paramètres abiotiques sur la structure et la succession des communautés microbiennes, la particularité des travaux conduits en Manche Orientale est de s'être focalisé sur l'influence d'un facteur biotique : l'efflorescence naturelle de *P. globosa*. Dans ce cadre-là, l'utilisation d'approches moléculaires a permis dans un premier temps d'accéder à la diversité des eucaryotes microbiens, y compris la composante rare et, dans un second temps, de définir une succession dans la structuration des communautés. Dans cet écosystème, trois périodes distinctes (post-bloom, bloom et pré-bloom de *P. globosa*) ont pu être identifiées dans la structuration des communautés eucaryotes microbiennes, ce qui est conforme aux résultats précédemment obtenus via les observations au microscope (*Grattepanche et al., 2011a*). Au sein de ces périodes, il a été observé que, dans la plupart des cas, la richesse des OTUs augmentait considérablement au début du printemps (Mars) et durant la période estivale et diminuaient pendant le bloom de *P. globosa*. Ainsi, l'OTU affiliée à *P. globosa* décrivait annuellement une augmentation considérable du nombre de séquences durant la période printanière (fin Mars-Avril), et contribuant ainsi à la forte abondance relative du groupe des Haptophyta.

○ *Intérêt de l'approche adoptée dans l'étude*

Plus globalement, l'approche adoptée dans ce projet (i.e séquençage massif) de thèse a déjà été appliquée sur des échantillons environnementaux de différents écosystèmes marins pour explorer la diversité spatiale et temporelle des procaryotes et des eucaryotes dans différentes régions du monde (*Amaral-Zettler et al., 2009* ; *Cheung et al., 2010* ; *Edgcomb et al., 2011* ; *Logares et al., 2014* ; *Massana et al., 2008* ; *Massana et al., 2015*) et a permis de révéler une diversité autrefois négligée ou difficilement identifiable par les méthodes conventionnelles (par exemple la microscopie) (*Massana et al., 2008* ; *Guillou et al., 2008* ; *Brate et al., 2012*). D'un point de vue pratique, le gène codant la petite sous-unité ribosomique des eucaryotes (18S) est retrouvé chez tous les eucaryotes. Il possède des régions très conservées entre espèces, et d'autres régions plus variables permettant l'assignation phylogénétique. Dans cette thèse, l'identification taxonomique des protistes a

été réalisée par amplification PCR des régions variables V2-V3 du 18S à partir de l'ADN total extrait des échantillons environnementaux. Les amplicons obtenus sont ensuite séquencés par des approches de séquençage à haut débit (i.e 454-pyrosequençage, Illumina). Les séquences obtenues, qui représentent la diversité génétique des amplicons et donc celle des organismes se trouvant dans chaque échantillon, sont groupées en unités taxonomiques opérationnelles (OTU, équivalent d'une espèce en biologie moléculaire). Les séquences de chaque OTU identifiée sont comparées à des organismes préalablement identifiés puis affiliées au même groupe taxonomique que les séquences de références. De cette manière, une majorité des organismes eucaryotes présents dans les échantillons sont identifiés. Il est toutefois important de noter que cette diversité peut inclure des organismes morts ou vivants et / ou comprenant de l'ADN provenant des pools extracellulaires dissous (*Karl et Bailiff, 2012*).

Dans cette thèse, nous avons également étudié la diversité des protistes par l'intermédiaire de l'ARNr du 18S (**Chapitre 3**). Tout comme l'ADNr, la séquence de l'ARNr est spécifique à chaque espèce, mais à l'avantage de n'être présent que chez les organismes métaboliquement actifs (et vivants). En outre, plus un organisme sera métaboliquement actif, plus il devra synthétiser des ribosomes (indispensable à la synthèse protéique) et donc plus il devra transcrire son ADNr en ARNr. Ainsi, le double ciblage de l'ADNr et de l'ARNr permet, en plus de l'identification de la diversité spatio-temporelle des espèces, d'estimer l'activité potentielle relative de chacune d'entre-elles par calcul du rapport ARNr : ADNr. Plusieurs études ont appliqué cette stratégie sur les communautés bactériennes marines (*Mills et al., 2005 ; Moeseneder et al., 2005 ; Gentile et al., 2006 ; Campbell et al., 2011*), les archées (*Hugoni et al., 2013*) et les eucaryotes (*Stoeck et al., 2007 ; Not et al., 2009 ; Orsi et al., 2013 ; Logares et al., 2014 ; Charvet et al., 2014 ; Massana et al., 2015 ; Hu et al., 2016*). Et malgré le fait qu'il existe plusieurs limitations à cette stratégie (voir **Chapitre 1, Partie I, section II.3**), son utilisation appliquée à chaque OTU individuellement devrait permettre d'apporter des informations supplémentaires indispensable pour une meilleure compréhension du fonctionnement de l'écosystème et de l'influence des paramètres biotiques et abiotiques sur la structuration des communautés microbiennes. Il devrait aboutir à mieux comprendre la structuration future des communautés face aux changements globaux.

## II. MEDITERRANEE : Méditerranée occidentale et Golf du Lion

Un chapitre de cette thèse est dédié à des données de diversité de protistes planctoniques provenant d'une région contrastée à la Manche orientale, la mer Méditerranée Occidentale (**Fig. 21**). Cette étude a été effectuée dans le cadre du projet SPECiMed (Structures of Planktonic Ecosystems in the North-western Mediterranean) (PI Bernard Queguiner MIO). SPECiMed visait à établir une plate-forme d'observation de trois ans (2012-2014) des communautés planctoniques, incluant les bactéries, microzoo- et mésozooplancton tout en associant les cycles biogéochimiques des éléments majeurs (C, N, P, et Si). Dans un contexte de gestion des écosystèmes marins, ce projet avait pour but de développer une connaissance prédictive de la manière dont les cycles biogéochimiques et les écosystèmes marins répondent à l'évolution des forçages, y compris comment les variations climatiques à grande échelle impactent le fonctionnement des écosystèmes régionaux, suite à l'évolution de la dynamique physique et de l'altération des cycles biogéochimiques. SPECiMed s'est appuyé sur l'expertise conjointe d'océanographes dans les domaines de la physique, de chimie et de biologie avec pour but ultime de comprendre le continuum réactionnel entier contrôlant la réponse des écosystèmes.

La Méditerranée est considérée comme un modèle réduit des océans. Elle est semi-fermée et couvre une superficie d'environ 2 510 000 km<sup>2</sup>, sa longueur d'est en ouest est d'environ 3860 km et sa largeur maximale de près de 1600 km. Sa profondeur moyenne est de 1500 m et elle atteint au maximum 5150 m au large de la côte méridionale de la Grèce. Elle est divisée en deux secteurs principaux, la Méditerranée Occidentale et la Méditerranée Orientale.

La Méditerranée ouverte est une mer oligotrophe voir très oligotrophe surtout pour sa partie Est alors que les zones côtières sont soumises à une forte pression anthropique et présentent souvent de conditions d'eutrophisation. En conséquence, la dynamique des populations des organismes marins est également fragilisée et les changements globaux entraînent une réponse rapide de son écosystème (*Turley, 1999 ; Licandro et Ibanez, 2000*), d'où l'importance d'évaluer la diversité des communautés y habitant et d'examiner leur réponse aux variations environnementales.



## **II.1. Méditerranée occidentale**

### *Caractéristiques générales*

Le bassin occidental méditerranéen se divise en six compartiments, on distingue : la mer d'Alboran, le bassin Algérien, la mer des Baléares, le bassin nord-occidental, la mer Ligure et la mer Tyrrhénienne. Il est délimité par cinq îles : les trois îles Baléares à l'ouest et la Corse et la Sardaigne à l'est. Le bassin occidental est peu profond environ 2600m au maximum et comporte de larges régions peu profondes comme c'est le cas pour le Golfe du Lion. Il est relié à l'ouest, par le détroit de Gibraltar (~14km de large et ~300m de profondeur) à l'Océan Atlantique et à l'est à la Méditerranée Orientale par le détroit de Sicile (~150km de large et ~400m de profondeur).

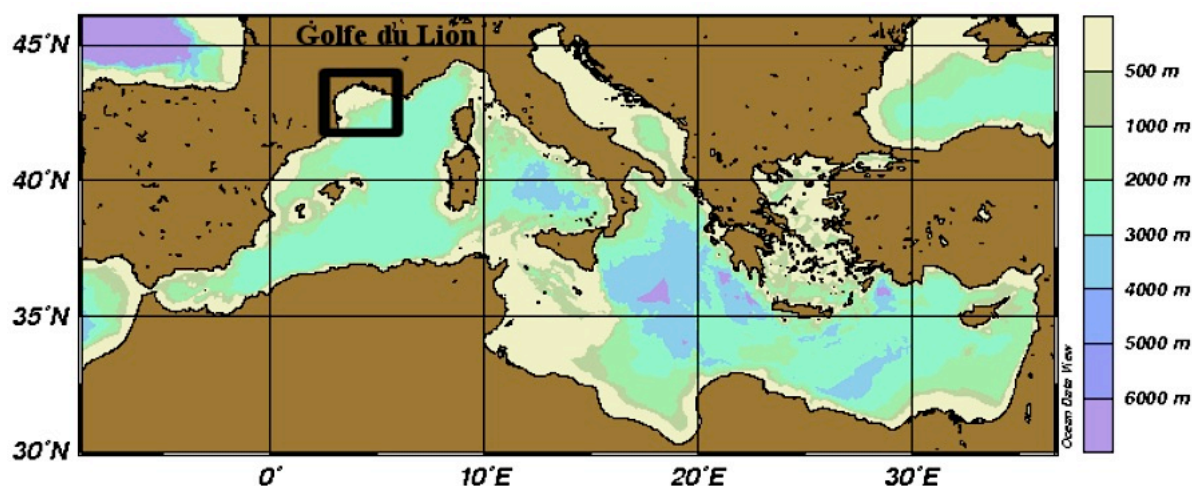
Le climat en Méditerranée est de type aride, caractérisé par un hiver humide et doux et un été chaud et sec. Néanmoins, les intersaisons donnent lieu à des violences certaines du climat. Les marées sont de faible amplitude et l'évaporation ( $3130 \text{ km}^3/\text{an}$ ) y est plus importante qu'en Océan Atlantique, provoquant ainsi des taux de salinité et des températures d'eaux plus élevés que dans ce dernier. Les pertes par évaporation ne peuvent être comblées ni par les apports en eaux douces (apports fluviaux de  $430 \text{ km}^3/\text{an}$ ) ni par les précipitations (pluviométrie de  $1\,000 \text{ km}^3/\text{an}$ ) (Bryden *et al.*, 1994). Cependant, une partie de ce déficit est compensée par les apports hydrologiques de la mer Noire ( $180 \text{ km}^3/\text{an}$ ) mais plus particulièrement par ceux provenant de l'océan Atlantique ( $1520 \text{ km}^3/\text{an}$ ).

La mer Méditerranée est comparée à un petit océan (Bethoux *et al.*, 1999). En effet, elle fait intervenir dans un espace relativement réduit des processus hydrodynamiques similaires à ceux associés au fonctionnement d'un océan, incluant des échelles spatio-temporelles variables. La circulation générale à une échelle supérieure est stationnaire, quant à la formation d'eaux denses dans le bassin Nord Occidental impliquant des processus à l'échelle du kilomètre pendant quelques heures.

## **II.2. Golfe du Lion**

Le Golfe du Lion (GoL), localisé au nord-ouest de la Méditerranée, est formé d'un plateau semi-circulaire qui s'étale sur une superficie de  $6700 \text{ km}^2$  de 90m de profondeur moyenne (170m en profondeur maximale), entre Marseille et le Cap de Creus (à côté de la frontière

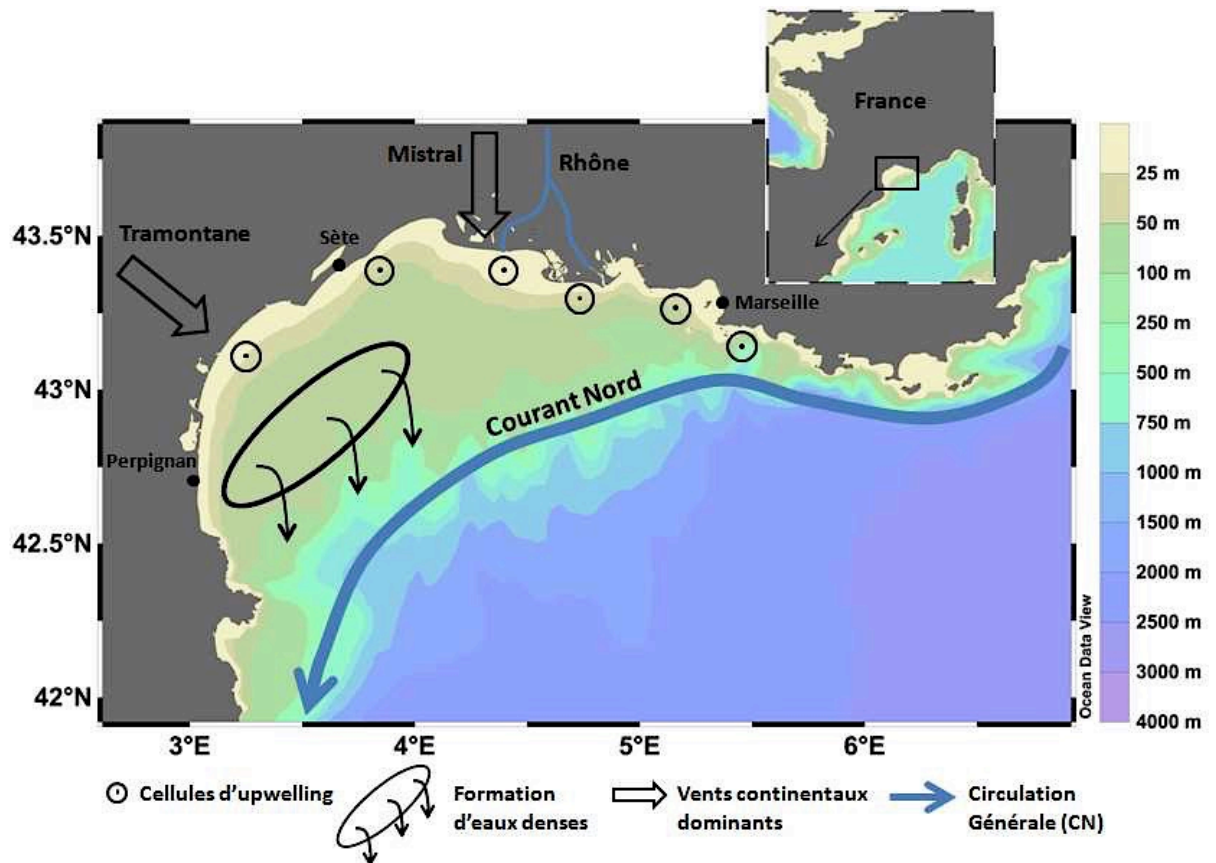
espagnole) Le GoL s'étend à sa bordure sud par un talus entaillé de nombreux canyons marins abrupts, constituant ainsi une frontière entre la région côtière et la plaine abyssale (2500-3000 de profondeur). Le talus et les canyons forment une zone de transition entre la mer côtière et la mer hauturière et jouent un rôle important dans l'hydrodynamique du GoL, aussi bien en termes d'échange d'eaux que de particules entre la côte et le large (*Estournel et al., 2001 ; Uses, 2006 ; Langlais, 2007*).



**Figure 21** : Carte représentative de la Méditerranée et de la géolocalisation du Golfe du Lion.

### II.2.1. Climat

Le GoL présente localement des particularités climatiques dues spécifiquement au régime de vent amplement contraint par la topographie. Ce climat venteux est prédominé par les vents continentaux du secteur nord et nord-ouest canalisés par les massifs montagneux (les Pyrénées, le Massif Central et les Alpes). Les couloirs maritimes atlantiques sont localisés entre le Nord des Pyrénées et le Sud du Massif central donnant naissance à la Tramontane et au Mistral entre l'Est du Massif central et l'Ouest des Alpes (vallée du Rhône) (**Fig. 22**). Ces vents régionaux sont très froids en hiver et souvent violents. Ils peuvent souffler à 100 km/h en plaine et à 10-15m/s en moyenne en été (*Millot, 1990*). De l'automne au printemps, des vents d'origine maritime peuvent également souffler du secteur sud-est et s'avèrent plus violents que ceux continentaux (aux alentours de 25m/s) (*Maillot, 1990 ; Petrenko, 2003 ; André et al., 2005*).



**Figure 22 :** Bathymétrie, circulation des masses d'eau et principaux forçages sur le plateau du GoL (d'après Millot, 1990). Les isobathes 160, 500, 1000 et 2000 sont représentées.

## II.2.2. Caractéristiques hydrodynamiques

Dans une approche générale, la circulation des masses d'eaux dans cette mer microtidale est impactée par la bathymétrie, les conditions atmosphériques (vent et flux de chaleur) ainsi que par les apports fluviaux.

En Méditerranée Nord Occidentale comme dans le GoL, la circulation est forcée par trois masses d'eau principales. On distingue : les eaux de surface d'origine Atlantique MAW (Modified Atlantic Water), les eaux levantines intermédiaires LIW (Levantine Intermediate Water) et les eaux profondes ouest méditerranéennes WMDW (Western Mediterranean Deep Water) (Millot, 1999 ; Send et al., 1999). On peut citer également les eaux denses tyrrhéniennes TDW (Tyrrhenian dense water).

➔ MAW proviennent de l'Atlantique par le détroit de Gibraltar s'écoulant en surface (0-300m), Elles sont soumises directement aux forçages atmosphériques. Elles séjournent en

premier temps dans le bassin Algérien et passe à l'Est de la Corse, puis longe le talus continental en mer Ligure, au large du GoL du plateau Catalan. Elles sont caractérisées par une faible salinité en Méditerranée et par une température sous la couche de mélange de surface de 14-15°C. Elles subissent des mélanges de courant durant son transit et deviennent de plus en plus salées.

➔ LIW (eau levantine intermédiaire) sont des masses d'eau formées en hiver dans le bassin Est, le bassin Levantin, pénétrant le canal de Sicile pour rejoindre la Méditerranée occidentale et se retrouver sous la MAW (300-700m) (*Ovchinnicov and Plakhin, 1984; Lascaratos, 1993; Nittis and Lascaratos, 1998*). Elles sont plus salées (38.45-38.75) et relativement chaudes (13-14°C) (*Wu and Haines, 1996*). Tel le cas des MAW, les LIW subissent des changements au cours de leur déplacement. Il est à noter que Millot (2013) avait proposé de renommer la LIW du bassin occidental en EIW (Eastern Intermediate Water).

➔ WMDW (eau profonde de Méditerranée occidentale) : par l'effet de l'évaporation due aux vents continentaux secs et froids « Mistral et Tramontane », les masses d'eau subissent en hiver, une diminution de la température de la couche de surface et une augmentation de la salinité suffisamment pour générer une masse d'eau dense profonde (jusqu'à 800m). Ceci se produit au large du GoL et en mer Ligure. Les WMDW sont relativement froides (12,7-12,9°C) et salées (38.4-38.48 ‰) (*Rixen et al., 2005*).

➔ TDW (eau dense tyrrhénienne) se produit en mer Tyrrhénienne par mélange entre la WMDW et la LIW pénétrant par le détroit de Sicile (*Millot, 1999 ; Funda et al., 2002*) et s'écoule sous la LIW une fois qu'elle franchit la Sardaigne.

La circulation dans le GoL est contrainte en premier ordre par le courant Nord (CN). Ce dernier est d'une largeur de 40 km et longe le talus continental. Il constitue la branche nord de la circulation cyclonique entreprise par les MAW et les LIW. Sa dynamique est soumise aux vents géostrophiques (les vents résultant de l'équilibre géostrophique entre la force de Coriolis et la force du gradient de pression atmosphérique). Le CN présente une importante variabilité saisonnière. Son flux double en hiver (de janvier à Mars), il est plus profond (environ 450) et sa vitesse peut arriver à  $1\text{m}\cdot\text{s}^{-1}$  (*Sammari et al., 1995*). De juin à décembre, le CN s'élargit (40-50 km), devient peu profond (environ 250m) et ralentit sa vitesse en

conséquence ( $0.5\text{m}\cdot\text{s}^{-1}$ ). Durant l'hiver, quand le flux est à son maximum, le CN présente une activité méso-échelle très marquée (*Flexas et al., 2002 ; Albérola et Millot, 2003 ; André, 2007*), pouvant provoquer la formation des méandres susceptibles d'affecter la circulation sur le plateau (*Petrenko, 2003*) et de favoriser les échanges d'eau entre le plateau continental et le large (*Durrieu de Madron et al., 1999*).

Comme cité précédemment, la circulation est également influencée par les vents continentaux du secteur Nord (Mistral) et Nord-Ouest (Tramontane) (**Fig. 22**). Ces vents induisent la formation d'eaux denses et des upwellings en période hivernale (*Durrieu et Madron et al., 1999 ; Millot, 1990*). Ces cellules d'upwelling côtières possèdent une durée de vie de quelques jours (*Millot et al. 1990*) et ont également été mises en évidence par images satellitaires de la couleur de l'eau (*Forget et André, 2007*). En ce qui concerne la circulation sur le plateau induite par les vents, le mistral engendre une circulation anticyclonique sur l'ensemble du plateau et également un jet orienté vers le Nord-est et parcourt le long des côtes (*Estournel et al., 2003*). Par contre, la Tramontane provoque un tourbillon cyclonique à l'ouest du plateau. La combinaison des deux engendre la formation d'un tourbillon anticyclonique au centre du plateau et cyclonique à l'ouest.

Ces vents génèrent également des houles qui se propagent vers le large. Elles sont plus fréquentes en été (*The Medatlas Group et al., 2004*). Elles sont généralement de faible amplitude, toutefois des tempêtes peuvent se produire et sont susceptibles de générer des houles de forte amplitude. En été, l'absence relative de vent favorise la stratification de la colonne d'eau et la mise en place d'une thermocline. Cette stratification est limitée dans le temps, la raison étant que les périodes sans vent sont de courte durée (*Rabouille et al., 2008*). Les plongées d'eaux denses et les houles induites par les vents marins contribuent à la remise en suspension des sédiments (*Millot, 1990 ; Bourrin et al., 2008 ; Ulses et al., 2008*).

### **II.2.3. Apport fluviaux dans le GoL**

Dans le GoL, le Rhône (**Fig. 21**) est le fleuve le plus important en termes d'apports liquides et solide à la Méditerranée. Près de 90% de l'eau d'origine terrestre apportée dans le GoL provient du Rhône. Les autres fleuves possèdent des débits nettement plus faibles (autour de la centaine de  $\text{m}^3 \text{s}^{-1}$ ), hors période de crue, durant lesquelles les débits liquides de l'Hérault

et le Vidourle peuvent atteindre  $1000 \text{ m}^3 \text{ s}^{-1}$ . Huit affluents contribuent aux apports fluviaux dans le GoL, classés ci-après dans l'ordre de flux annuel décroissant : Rhône, Hérault, Aude, Orb, Têt, Tech, Vidourle et Agly (Bourrin et Durrieu de Madron, 2006). L'apport principal vient du Rhône dont le débit liquide peut être sujet à des crues considérables et atteindre  $10000 \text{ m}^3 \text{ s}^{-1}$ . Son flux annuel est 17.5 fois plus élevé que celui de l'Aude, de l'Hérault et de l'Orb combinés et 135 fois celui du Têt (Ludwig *et al.*, 2003).

Sur le plateau continental du GoL, le Rhône représente la source principale des apports de carbone organique et particulaire (Cauwet *et al.*, 1990) et également des apports terrigènes (80%) (Monaco *et al.*, 1990 ; Durrieu de Madron *et al.*, 2000). Les écosystèmes côtiers présentent une grande productivité expliquée en majeure partie par les apports fluviaux en matière organique terrigène et en éléments nutritifs. La marée de faible amplitude provoque un mélange vertical faible et une dilution lente de l'eau douce ramenée par les fleuves (Lamy *et al.*, 1981).

#### **II.2.4 Caractéristiques biologiques (impact des conditions environnementales et composition)**

La Méditerranée est reconnue comme un écosystème oligotrophe et dont sa composante biologique est soumise à différents facteurs abiotiques (vents, upwelling, limitation en éléments nutritifs) et biotiques (par exemple, la concurrence vis-à-vis des éléments nutritifs et les relations inter-taxa). Par exemple, Ludwig *et al.*, 2009 ont souligné une production primaire faible liée à la limitation en phosphate et en silice. En revanche, Marty *et al.* (2002) ont avancé que la dominance de la biomasse phytoplanctonique due aux phytoplanctons de petites tailles (nanoflagellés et cyanobactéries) a été favorisée par une période de stratification estivale allongée. Ce résultat rejoignait les observations de Goffart *et al.* (2002) dans la mer Ligure (NMW), où ils ont démontré un impact significatif des conditions environnementales, et en particulier de la température sur le phytoplancton. Suite à un suivi à long terme (de 1994 à 2006), 288 espèces de dinoflagellés et de diatomées ont été recensées dans la baie de Marseille (M. Girault, non publié, <http://www.mio.univ-amu.fr/~queguiner/SPEciMed 2011 dossier actualise.pdf>), certaines espèces montraient un développement particulier qui se manifestait par leur apparition ou leur disparition. Par exemple, le dinoflagellé *Prorocentrum minimum* diminuait après 1999, au moment où le genre *Gymnodinium* semblait trouver des conditions plus favorables à son

développement. De manière similaire chez les diatomées, *Chaetoceros curvisetum* devenait très commun alors que le *Thalassionema frauenfeldii* disparaissait presque des enregistrements. Ces espèces peuvent être considérées comme des espèces sentinelles dont la présence ou l'absence est le reflet de changements dans les conditions environnementales. En ce qui concerne les communautés mésozooplanctoniques, Molinero et al. (2005) ont détecté une évolution temporelle opposée des copépodes et des méduses, suggérant que le forçage climatique pourrait agir par des changements de température de l'eau. Au cours du programme SESAME (Southern European Seas : Assessing and Modelling Ecosystem changes), des études sur la distribution du zooplancton à travers le GoL et sa connexion avec le Courant Nord (CN) ont permis de mettre en évidence les différents comportements des espèces de mésozooplancton qui semblent directement influencées par la température. Certaines espèces telles que *Calanus helgolandicus*, *Temora stylifera* et *Centropages typicus* ont réduit leur présence saisonnière dans l'écosystème en réponse aux changements de paramètres environnementaux, y compris la température. Ces travaux ont également souligné l'importance des structures physiques ou topographiques spécifiques dans la survie de *C. typicus* en hiver, tandis que les zones profondes des canyons de la pente continentale permettent à *C. helgolandicus* de rester en diapause pendant l'été, ce qui lui permet d'élargir son aire de répartition en hiver après le refroidissement du plateau. D'autres travaux basés sur de longues séries temporelles ont également souligné l'impact des forçages environnementaux sur la composition et la dynamique des communautés planctoniques (par exemple : *Gómez & Gorsky, 2003; Pereira Quiroga, 2005; Tunin Ley et al., 2007*).

Plusieurs études soulignent l'importance des analyses au niveau communautaire et / ou au fonctionnel du plancton afin de mieux comprendre les différents processus des changements écosystémiques (*Karl et al., 2001; Mackas et al., 2001; Batchelder & Powell, 2002; Beaugrand et al., 2002; Hooff et Peterson, 2006; Chiba et al., 2008*). Néanmoins, jusqu'à présent, les études qui ont été menées dans le GoL concernaient principalement la circulation de l'eau, l'environnement trophique et les structures communautaires de quelques genres ou groupes précis, essentiellement celles indicatrices des changements environnementaux (*Tunin-Ley et al., 2007, 2009*) et celles potentiellement toxiques (Harmful Algal Blooms) (ex, *Ostreopsis sp. (Cohu et al., 2011; Blaufune et al., 2012)*), ainsi qu'*Alexandrium sp. (Bolch et al., 2007)*. Toutefois, les études visant à obtenir une vision globale de la diversité microbienne restent jusqu'ici rares. En effet, la majorité des

travaux réalisés était généralement basée sur l'usage des méthodes classiques (par exemple : la microscopie).

En revanche, l'assemblage des communautés bactériennes a été examiné, dans la région d'étude, via les techniques de séquençage de nouvelle génération (par exemple : *Salter et al., 2014*). Toutefois, à notre connaissance, cette approche n'a que très rarement été utilisée pour étudier la diversité des communautés eucaryotes. Ainsi, une seule étude récente a eu recours à l'approche de séquençage à haut débit pour étudier les assemblages phytoplanctoniques dans trois écosystèmes dont la Méditerranée Occidentale (et plus précisément dans la baie de Marseille) (*Penna et al., 2017*). Cette étude a souligné l'hétérogénéité des assemblages phytoplanctoniques dans la Méditerranée occidentale comparée à sa partie orientale. Toutefois, aucune étude n'a pris en compte la diversité eucaryote microbienne dans sa globalité. L'intérêt de ce projet est donc d'obtenir, dans un premier temps, une vue d'ensemble des communautés protistes eucaryotes microbiens et, dans un second temps, d'évaluer leur dynamique spatio-temporelle.



## **PARTIE 3 : OBJECTIFS**

Dans ce projet de thèse, la grande partie d'étude concernait la Manche Orientale mais nous avons dédié comme précédemment mentionné un chapitre à la Méditerranée. Nous avons visé en premier ordre d'explorer la diversité et la structure des protistes eucaryotes dans deux conditions environnementales contrastes, en combinant à la fois les approches moléculaires (Illumina et pyroséquençage) et classiques (microscopie) et d'évaluer l'impact des variables environnementales sur leur structuration. L'utilisation des deux approches combinées concernait l'écosystème méditerranéen, pour but de maximiser la détection des organismes déjà connus et d'identifier également de nouveaux groupes taxonomiques.

Les objectifs principaux de cette thèse étaient (i) de renforcer nos connaissances sur la diversité et l'activité des protistes dans un écosystème méso-eutrophe (cas de la Manche Orientale) et mettre également en évidence les successions saisonnières communautaires des protistes eucaryotes pendant le bloom de *Phaeocystis globosa* (Rachik et al., en presse) (ii) d'étudier la diversité et la variabilité spatiale des protistes dans un écosystème oligotrophe contrasté à la Manche Orientale (cas du Golfe du Lion) (Rachik et al., in prep). (iii) et enfin, de mettre évidence les relations existantes entre les parasites eucaryotes et les autres taxons ainsi que les paramètres environnementaux (Christaki et al., 2017).

Ce manuscrit de thèse se divisent en deux parties principales: La Manche Orientale (**Chapitre 3 et 4**) et la Méditerranée (**Chapitre 5**).

### **Manche Orientale (Chapitre 3 et 4)**

#### ***1. Etude de la diversité et du potentiel d'activité des protistes eucaryotes en Manche Orientale***

Nous avons exploré en premier temps la diversité et la structure des communautés protistes eucaryotes sur une période de 4 ans (de Mars 2011 à Juillet 2015, 77 échantillons) à la station SOMLIT par séquençage de l'ADN18S par les techniques Illumina (appliquée sur les échantillons prélevés entre 2013-2015) et pyroséquençage (appliqué sur ceux prélevés en 2011-2013). L'intérêt de combiner les deux méthodes avait pour but d'obtenir un bilan général de la diversité sur la période précédemment mentionnée. Nous nous sommes

concentrés par la suite sur la diversité évaluée entre 2013-2015. Afin de la mettre dans son contexte environnemental, nous avons effectué des analyses statistiques afin d'hierarchiser les variables environnementales les plus contributrices aux spécificités saisonnières et ainsi à la structure des communautés microbiennes. Et pour apporter plus d'éléments de réponse, nous avons examiné le potentiel d'activité des communautés eucaryotes présentes par séquençage de l'ARN18S (via la technique Illumina) et calcul du ratio ARN18S:ADN18S. Ceci a été appliqué sur les échantillons prélevés entre 2013-2015.

Dans cette partie, quelques questions majeures ont été abordées :

- Quelles sont la composition et la succession des communautés planctoniques eucaryotes en relation avec les variables environnementales en Manche Orientale ?
- Quelles sont les principales différences et ressemblances entre les deux ensembles de données ADNr et ARNr ?
- Et enfin, est-ce-que l'approche ARNr/ADNr est écologiquement pertinente?

## ***2. Etude des parasites eucaryotes et leurs interactions avec les autres taxons et les conditions environnementales***

L'étude des parasites a pris plus d'importance dans ces dernières années en raison de leur rôle important dans la structuration des communautés planctoniques. Les parasites sont connus par leur faculté d'infecter les organismes à divers niveaux trophiques et d'augmenter la complexité des réseaux trophiques planctoniques en élargissant les chaînes alimentaires (*Grami et al., 2011*). Ceci, à son tour, peut entraîner une augmentation à la fois de la connectivité et de l'efficacité des flux de carbone, bien que de telles conséquences soient encore l'objet de débat. Il est donc essentiel d'étudier comment le parasitisme peut entraîner des successions microbiennes. Pour ceci, une étude de "métadonnée" a été menée pour les données obtenues pendant la période 2011-2013. Cette étude consistant à mettre en évidence des relations existantes entre les parasites eucaryotes et les autres taxons ainsi que les paramètres environnementaux (32 échantillons traités).

Les objectifs principaux de cette étude ont été (i) d'explorer la capacité des réseaux de corrélation pour mettre en évidence les relations en les taxons SymbDec (symbiontes-décomposeurs) et la succession saisonnière du microplancton, et (ii) d'identifier les

paramètres environnementaux affectant des groupes taxonomiques et trophiques spécifiques de la communauté du microplancton dans les eaux côtières tempérées de la Manche Orientale.

### **Méditerranée Occidentale (Golfe du Lion, Chapitre 5)**

#### ***Etude de la diversité des protistes dans un écosystème côtier oligotrophe (Méditerranée Occidentale, Golfe du Lion)***

Comme pour la Manche Orientale, une étude des protistes eucaryotes a été réalisée en Méditerranée Occidentale. Cette étude a été réalisée comme précédemment mentionné dans le cadre du projet SPECiMed (Structures of Planktonic Ecosystems in the North-western Mediterranean) (PI Bernard Queguiner MIO). Notre contribution dans ce projet concernait l'exploration de la variabilité de la structure et dynamique des communautés eucaryotes microbiennes au regard des paramètres environnementaux et de la zone géographique étudiée. L'étude de diversité a été réalisée par séquençage de la région variable du gène de l'ARNr18S par la technique Illumina.

Les principaux objectifs étaient :

- Quelle est la diversité globale des protistes telle que révélée par séquençage à haut débit ?
- Quelles sont les principales différences et similitudes entre les stations côtières et celles du large et entre le sud-est et sud-ouest du Golfe du Lion et leur lien avec les variables environnementales ?

## **CHAPITRE II – MATERIELS ET METHODES**



## CHAPITRE II – MATÉRIELS ET MÉTHODES

### I. APPROCHE MOLECULAIRE

L'avènement des techniques de séquençage à haut débit (NGS) (e.g. Pyroséquençage, Illumina) a rendu possible le séquençage d'un grand nombre de nucléotides (jusqu'à  $10^{12}$  bases séquencées par expérience-avec la technologie Illumina) à un coût nettement moindre qu'avec la méthode de Sanger. De plus, les approches de séquençage à haut débit permettent d'accéder à la diversité taxonomique et fonctionnelle en s'affranchissant les étapes de clonages, longues et coûteuses.

Dans ce projet de thèse, l'intérêt s'est principalement porté sur l'exploration de la variabilité de la structure des communautés microbiennes eucaryotes en Manche Orientale ainsi qu'à la (sub)mesoscale dans le Golfe du Lion (les particularités de chaque milieu sont reportées dans le **Chapitre 1**). Le séquençage de l'ADNr 18S a été réalisé via la méthode Illumina Miseq appliquée sur les échantillons provenant de la Méditerranée (2012-2014) et sur ceux provenant de la Manche Orientale (2013-2015). Cette méthode permet d'évaluer la diversité et la structure des communautés microbiennes planctoniques, à une échelle très fine. Son principe repose sur l'amplification par PCR de la région du gène codant pour l'ARNr 18S en ciblant plus particulièrement les régions variables du gène (différentes entre les espèces). Après cette étape d'amplification, les échantillons sont préparés (ce que l'on appelle communément construction des bibliothèques) afin d'ajouter sur chaque amplicon un adaptateur et un code barre. L'insertion du code barre, identique pour un même échantillon mais différents entre échantillons, permet de réaliser le multiplexage qui consiste à mélanger dans un même tube les produits d'amplification issus de différents échantillons, tout en ayant la capacité d'identifier à *posteriori* la provenance des séquences (Pour plus de détails sur la technique : voir **Chapitre 2**).

En raison du développement rapide et continu des techniques de séquençage à haut débit, nous avons assisté à l'obsolescence de la méthode 454-pyroséquençage populaire à l'époque et à la montée en puissance du séquençage Illumina. En effet, le 454-pyroséquençage utilisé lors du premier suivi (2011-2013) a été progressivement abandonné au profit d'Illumina utilisée lors du second suivi (2013-2015). Ainsi, afin d'établir une vue d'ensemble de la diversité microbienne sur quatre ans (de Mars 2011 à Juillet 2015), nous

devions combiner les données de séquences Illumina issues de l'étude (réalisée entre 2013 et 2015) avec celles précédemment obtenues par pyroséquençage entre Mars 2011 à Juillet 2013 (GenBlanc-SRA accession SRX768577, *Genitsaris et al., 2015*). Les deux plateformes ont fourni une estimation comparable de la structure de la communauté microbienne (*Ferrera et al., 2016*). Cependant, après avoir combiné et analysé *in silico* les données de 454-pyroséquençage et d'Illumina, il est apparu que les deux jeux de données ne peuvent pas être directement regroupés et analysés ensemble. En effet, bien que techniquement possible, la mise en commun des données issues de deux approches de séquençage à haut débit a conduit à des reconstructions communautaires en fonction des méthodes de séquençage et non en fonction de la structuration des communautés par date. Ce résultat n'était pas surprenant car les deux méthodes sont différentes en termes de préparations de bibliothèques, de longueur et de quantités de séquences générées. Etant donné que la comparaison entre les deux techniques ne faisait pas l'objet de cette étude, nous avons décidé d'analyser indépendamment les données, puis de comparer les résultats obtenus. Pour cela, après un traitement bioinformatique indépendant des données (voir section I.1.3 ci dessous), nous avons affilié les OTU obtenues de la même manière (par utilisation de la dernière version des bases de données de références : NCBI, SILVA et PR2), puis de comparer les résultats obtenus pour la diversité et la succession des espèces.

### *I.1. Extraction de l'ADN et de l'ARN totaux, amplification et séquençage par Illumina*

Les différentes étapes concernant le pyroséquençage (suivi du 2011-2013) sont expliquées dans *Genitsaris et al., 2015* (Voir également **Annexe 1**). Pour l'ADN ribosomique des organismes provenant du Golfe du Lion a été extrait, purifié et amplifié de la même manière que ceux provenant de la Manche Orientale (voir **section I.1.1** ci-dessous). Les échantillons traités contenaient entre 2.59 et 53.1 ng.µL<sup>-1</sup> d'ADN. Le déroulement des différentes étapes (extraction, purification, amplification et séquençage) est détaillé davantage dans le "**Chapitre 5**". Dans cette partie, juste les étapes de traitement des échantillons (2013-2015) provenant de la Manche Orientale seront détaillées.

Il est à signaler que la préparation des bibliothèques, ainsi que le séquençage des échantillons du Golfe du Lion (projet SPECIMED) a été réalisé par le laboratoire *Genoscreen* (Institut Pasteur, Lille), tandis que ce travail a été effectué par le laboratoire *GenesDiffusion* (Institut

Pasteur). J'ai notamment contribué à la préparation des bibliothèques nécessaires au séquençage des échantillons de la Manche Orientale.

### **I.1.1 Extraction de l'ADN et de l'ARN totaux**

Après filtration des échantillons d'eau de mer, l'ADN et l'ARN totaux ont été extraits et purifiés simultanément sur le même filtre à l'aide du kit d'isolement d'ADN et d'ARN (Qiagen Allprep DNA/RNA Mini Kit 50) en suivant le protocole du fabricant. Afin d'éliminer l'ADN contaminant l'ARN et inversement, tous les échantillons d'ARN ont été traités avec de la DNase et ceux de l'ADN avec de la RNase. Les échantillons traités contenaient entre 0,26 et 28 ng  $\mu\text{L}^{-1}$  d'ADN et entre 1,14 et 14,8 ng  $\mu\text{L}^{-1}$  d'ARN. Les dosages des quantités d'ADN et d'ARN ont été réalisés avec le Qubit 2.0 Fluorometer (Thermo Fisher Scientific Inc., Massachusetts, USA).

### **I.1.2. Amplification et séquençage**

Pour commencer, l'ARN extrait a été dénaturé à 70°C pendant 5 min et 25°C pendant 10 min en présence des deux amorces universelles d'eucaryotes 18S-82F (5'-GAAACTGCGAATGGCTC-3', *Lopez-Garcia et al., 2003*) et Euk-516r (5'-ACCAGACTTGCCCTCC-3', *Amann et al., 1990*), conçues pour amplifier les régions variables des eucaryotes V2 et V3 du gène de l'ARNr18S (environ 470-480 pb). Ces amorces ont été utilisées avec succès dans des études similaires (*Monchy et al., 2012* ; *Christaki et al., 2014* ; *Georges et al., 2014* ; *Rédou et al., 2014* ; *Jing et al., 2015* ; *Genitsaris et al., 2015, 2016*). L'ARN total a ensuite été reverse transcrit en ADNc en utilisant la transcriptase inverse M-MLV ARNase H Minus ADN Polymérase (Euromedex, Souffelweyersheim, France) à 45°C pendant 60 min puis maintenu à 10°C. Puis, les échantillons d'ADNc ont été traités avec de la RNase A (Euromedex, Souffelweyersheim, France) à 37°C pendant 10 min, afin d'éliminer l'ARNm résiduel, et purifiés en utilisant le kit de nettoyage PCR NucleoFast 96 (Macherey Nagel, Düren, Allemagne). Pour les constructions de banques d'amplicons, les échantillons d'ADN et d'ADNc ont été amplifiés par PCR avec les amorces Euk-82F et Euk-516R. Un second cycle de PCR a été réalisé afin de lier les adaptateurs de séquençage et le code-barres aux amplicons précédemment obtenus. Lors des deux amplifications par PCR, l'ADN Polymérase à haute fidélité « Platinum Taq » (Thermo Fisher Scientific, Waltham, MA, USA) a été utilisé avec les



cycles suivants: 94°C pendant 1 min, [15s à 94°C, 15s à 51°C, 45s à 68°C] (26 cycles pour la 1<sup>ère</sup> PCR et 9 cycles pour la 2<sup>ème</sup> PCR) et 1 minute à 68°C. Les produits de la seconde PCR ont été purifiés avec le système Agencourt AMPure XP (Beckman Coulter, Brea, CA, USA), quantifiés avec le test PicoGreen (Thermo Fisher Scientific, Waltham, MA, USA) et leur qualité évaluée avec le kit d'analyse Bioanalyzer High Sensitivity DNA (Agilent, Santa Clara, CA, États-Unis). Les banques d'amplicons ont été séquencées en mode « paired-end » (2x300bp) grâce à la plate-forme Illumina MiSeq (Plateforme de Génomique LIGAN-PM, Université Lille, CNRS-UMR8199, Lille, France).

### **I.1.3. Traitement de données**

De nos jours, il existe plusieurs logiciels permettant de filtrer les séquences en fonction de leur qualité et de les regrouper en unités taxonomiques opérationnelles (OTU : l'équivalent moléculaire d'une espèce), tels que QIIME (*Caporaso et al., 2010*), ESPRIT (*Sun et al., 2009*) ou MOTHUR (*Schloss et al., 2009*), en se basant sur les similarités de séquences. Les OTU obtenues sont ensuite affiliées via une base de données de référence telle que SILVA (*Quast et al., 2013*) ou PR2 (*Guillou et al., 2013*) grâce au logiciel BLAST. Pour ce travail de thèse, les données ont été traitées avec le Software Mothur v.1.35.0 et l'affiliation a été faite sur les base de données PR2 et SILVA et manuellement sur la plateforme NCBI.

#### *○ Analyse bioinformatique*

L'analyse des données pyroséquencées ou issues d'Illumina ont été réalisées en suivant les procédures bioinformatiques adaptées des «Standard Operating Procedure» (SOP, [www.mothur.org/wiki/454\\_SOP](http://www.mothur.org/wiki/454_SOP) et [http://www.mothur.org/wiki/MiSeq\\_SOP](http://www.mothur.org/wiki/MiSeq_SOP), pour le pyroséquençage et Illumina, respectivement) suggérées par les créateurs du Mothur (*Schloss et al., 2009*). Dans la suite, seule la procédure Illumina (échantillons de Méditerranée (2012-2014) et de Manche Orientale (2013-2015)) sera détaillée. Ainsi, les différentes étapes sont expliquées globalement pour l'ensemble des échantillons peu importe la provenance. Toutefois, à la fin de cette partie, un tableau présentant le nombre de séquences obtenues durant les étapes les plus importantes sera présenté (**Tableau 1**). Sous Mothur, la première étape consistait à combiner les deux ensembles de séquences (paired-end) au sein de chaque échantillon, et pour l'ensemble des échantillons. Ceci a été

effectué par l'utilisation de la commande `make.contigs`, qui nécessite l'utilisation de `stability.files` en entrée. Ce fichier contient un tableau formé de trois colonnes, la première comporte les groupes, la seconde le *forward fastq* et la troisième la *reverse fastq*. Cette commande permet d'extraire les données de séquences et de qualité des fichiers *.fastq* et de créer le complément inverse de la séquence, puis de joindre les séquences en contigs à partir des régions chevauchantes. Elle permet de générer quatre fichiers. Les fichiers *.trim.contigs.fasta* et *.groups* contiennent les données de séquences et l'échantillon d'appartenance pour chaque séquence. Le fichier *.report* nous renseigne sur l'assemblage en contig pour chaque séquence.

```
mothur > make.contigs (file=StabilityEuk.txt, processors=3), Euk pour Eucaryotes.
```

```
Fichiers de sortie :
```

```
StabilityEuk.trim.contigs.fasta
```

```
StabilityEuk.contigs.report
```

```
StabilityEuk.scrap.contigs.fasta
```

```
StabilityEuk.contigs.groups
```

La commande `summary.seqs`, permet d'avoir les statistiques relatives aux séquences présentes dans le fichier *.fasta*. Elle nous renseigne sur le nombre total des séquences, la longueur moyenne (min, max), le nombre de base ambiguë (min, max) ainsi que le nombre d'homopolymères (min, max) pour chaque quartile.

```
mothur > summary.seqs (fasta=Fichier.trim.contigs.fasta, processors=1)
```

```
Fichier de sortie:
```

```
StabilityEuk.trim.contigs.summary
```

La commande `count.groups` nous permet de vérifier le nombre total de séquences correspondant à celles présentes dans le fichier *.fasta* et de nous renseigner également sur le nombre de séquences présentes dans chaque échantillon.

```
mothur > count.groups (group=StabilityEuk.contigs.groups)
```

```
Fichier de sortie :
```

```
StabilityEuk.contigs.count.summary
```

Par la suite, nous avons utilisé la commande `screen.seqs` afin de conserver que les séquences répondant à des critères de qualité bien définis (e.g. la longueur de séquences maximale et le nombre de base ambiguë maximal) et d'éliminer ainsi toutes celles qui ne

répondent pas aux critères exigés. La longueur maximale des séquences correspond à la longueur de séquences à 97.5% de séquences (maxlength= 484 et 455 pour les échantillons provenant de la Manche Orientale et de la Méditerranée, respectivement). Aucune base ambiguë n'était tolérée dans nos analyses.

```
mothur > screen.seqs (fasta=Fichier.trim.contigs.fasta, group=Fichier.contigs.groups,  
maxambig=0, maxlength=484/455, processors=3)  
Fichiers de sortie :  
StabilityEuk.trim.contigs.good.fasta  
StabilityEuk.trim.contigs.bad.accnos  
StabilityEuk.contigs.good.groups
```

Afin de réduire le jeu de données aux seules séquences informatives, la commande *unique.seqs* a été utilisée afin d'éliminer les doublons et pour ne seulement conserver que les séquences uniques. Cette sélection successive des séquences informatives, suivie durant tout le processus d'analyse, permet de réduire la quantité de données, de calcul, et donc le temps d'analyse, tout en conservant dans des fichiers séparés l'intégralité de l'information.

```
mothur > unique.seqs (fasta=StabilityEuk.trim.contigs.good.fasta)  
Fichiers de sortie :  
StabilityEuk.trim.contigs.good.names  
StabilityEuk.trim.contigs.good.unique.fasta
```

Et afin de simplifier l'analyse, la commande *count.seqs* a été utilisée afin de générer une table contenant l'ensemble des données de séquence avec en lignes les noms des séquences et en colonnes les noms des échantillons où ces séquences ont été obtenues. La table est ensuite remplie avec le nombre de fois que chaque séquence unique apparaît dans chaque groupe.

```
mothur > count.seqs (name=StabilityEuk.trim.contigs.good.names,  
group=StabilityEuk.contigs.good.groups)  
Fichier de sortie :  
StabilityEuk.trim.contigs.good.count_table
```

De la même façon que précédemment, un *summary.seqs* a été appliqué afin de vérifier les résultats (le nombre total des séquences et le nombre total des séquences uniques).

```
mothur > summary.seqs (count=StabilityEuk.trim.contigs.good.count_table, processors=1)  
Using StabilityEuk.trim.contigs.good.unique.fasta as input file for the fasta parameter  
Fichier de sortie :  
StabilityEuk.trim.contigs.good.unique.summary
```

Par la suite, les séquences ont été alignées sur un alignement de référence (fichier SilvaEukV2V3.filter.fasta, contenant 26866 séquences 18S eucaryotes) récupéré via la base de données Silva ([www.arb-silva.de](http://www.arb-silva.de))

```
mothur > align.seqs (fasta=StabilityEuk.trim.contigs.good.unique.fasta,  
reference=SilvaEukV2V3.filter.fasta, processors=3)
```

Fichier de sortie :

```
StabilityEuk.trim.contigs.good.unique.align  
StabilityEuk.trim.contigs.good.unique.align.report  
StabilityEuk.trim.contigs.good.unique.flip.accnos
```

Après l'alignement, nous avons lancé encore une fois le *summary.seqs* afin de vérifier l'état de l'alignement général des séquences et d'identifier un éventuel problème.

```
mothur > summary.seqs (fasta=StabilityEuk.trim.contigs.good.unique.align,  
count=StabilityEuk.trim.contigs.good.count_table, processors=1)
```

Fichier de sortie :

```
StabilityEuk.trim.contigs.good.unique.summary
```

Afin de borner l'alignement, la commande *screen.seqs* a été utilisée avec en entrée les coordonnées de l'alignement déterminées via les résultats de la commande *summary.seqs* précédente. A ce niveau du traitement, seules les séquences dont l'alignement commencent aux positions 0 (pour les échantillons de la Manche Orientale) et 19 (pour les échantillons de la Méditerranée) et se terminent aux positions 2578 et 2558 (pour la Manche Orientale et la Méditerranée, respectivement) sont retenues. Les bornes d'alignement sont choisies de façon à ce que 97.5% des séquences soient conservés.

```
mothur > screen.seqs (fasta=StabilityEuk.trim.contigs.good.unique.align,  
count=StabilityEuk.trim.contigs.good.count_table,  
summary=StabilityEuk.trim.contigs.good.unique.summary, start=1, end=2578,  
maxhomop=8, processors=3)
```

Fichiers de sortie :

```
StabilityEuk.trim.contigs.good.unique.good.summary  
StabilityEuk.trim.contigs.good.unique.good.align  
StabilityEuk.trim.contigs.good.unique.bad.accnos  
StabilityEuk.trim.contigs.good.good.count_table
```

Le *summary.seqs* a été relancé à nouveau afin de vérifier que le bornage a bien été effectué et que toutes les séquences commencent bien à la position 1 et 19 (pour les séquences de la

Manche Orientale et la Méditerranée, respectivement) et se terminent bien à la position 2578 et 2558 (pour les séquences de la Manche Orientale et la Méditerranée, respectivement).

```
mothur > summary.seqs (fasta=current, count=current, processors=1)
Using StabilityEuk.trim.contigs.good.good.count_table as input file for the count
parameter.
Using StabilityEuk.trim.contigs.good.unique.good.align as input file for the fasta
parameter.
Fichier de sortie :
StabilityEuk.trim.contigs.good.unique.good.summary
```

Suite à l'alignement de nos séquences, de nombreuses positions de l'alignement commençaient ou se terminaient par des caractères (« . ») et contenaient uniquement des gaps (caractères « - »). Bien qu'en principe cela ne pose pas de problème pour l'analyse, ces caractères sont non informatifs, et prennent donc inutilement de la place en mémoire. Ainsi, afin de simplifier, sans perte d'information de séquences, et accélérer le temps de calcul, la commande *filter.seqs* a été utilisée pour supprimer toutes les colonnes dans l'alignement contenant uniquement le caractère (« - ») via l'option « vertical=T », et éliminer les caractères (« . ») par l'option « trump= ».

```
mothur > filter.seqs (fasta=StabilityEuk.trim.contigs.good.unique.good.align, vertical=T,
trump=.)
Fichiers de sortie :
StabilityEuk.filter
StabilityEuk.trim.contigs.good.unique.good.filter.fasta
```

En éliminant les extrémités des séquences, il se peut que certaines séquences soient de nouveau identiques. Pour les supprimer, nous avons réexécuté la commande *unique.seqs* afin de ne conserver que les séquences uniques.

```
mothur > unique.seqs (fasta=StabilityEuk.trim.contigs.good.unique.good.filter.fasta,
count=StabilityEuk.trim.contigs.good.good.count_table)
Fichiers de sortie :
StabilityEuk.trim.contigs.good.unique.good.filter.count_table
StabilityEuk.trim.contigs.good.unique.good.filter.unique.fasta
```

En partant du principe que plus une séquence est abondante, plus elle est en état de produire des versions erronées d'elle-même, nous avons utilisé la commande *pre.cluster* afin de réduire le nombre de séquences pouvant comporter des erreurs de séquençage. Ainsi, cette commande divisera les séquences par groupe, les triera par abondance, puis analysera chacune d'entre-elle individuellement en commençant par les plus abondantes. Toute séquence ayant jusqu'à deux paires de base de différence avec une séquence plus abondante (diffs=2) sera groupée avec cette dernière. Ainsi, les séquences les plus rares, très similaires à une séquence davantage abondante, seront éliminées, ce qui aura pour conséquence de simplifier une fois de plus le jeu de données et d'accélérer le traitement *in silico* de l'information. Néanmoins, une trace de toutes les séquences éliminées est conservée dans les fichiers afin de les prendre en considération lors du résultat final.

```
mothur > pre.cluster (fasta=StabilityEuk.trim.contigs.good.unique.good.filter.unique.fasta,  
count=StabilityEuk.trim.contigs.good.unique.good.filter.count_table, diffs=2)  
Fichiers de sortie :  
StabilityEuk.trim.contigs.good.unique.good.filter.unique.precluster.fasta  
StabilityEuk.trim.contigs.good.unique.good.filter.unique.precluster.count_table  
stability.trim.contigs.good.unique.good.filter.unique.precluster.Euk1.map  
...
```

Juste après *le pre.cluster*, nous avons utilisé la commande *split.cluster* afin de faire un premier tri parmi les séquences (séparation en fonction de leur abondance).

```
mothur >  
split.abund (fasta=stability.trim.contigs.good.unique.good.filter.unique.precluster.fasta,  
count=stability.trim.contigs.good.unique.good.filter.unique.precluster.count_table,  
cutoff=1, accnos=true)  
  
Fichiers de sortie :  
stability.trim.contigs.good.unique.good.filter.unique.precluster.rare.count_table  
stability.trim.contigs.good.unique.good.filter.unique.precluster.abund.count_table  
rare.accnos  
abund.accnos  
stability.trim.contigs.good.unique.good.filter.unique.precluster.rare.fasta  
stability.trim.contigs.good.unique.good.filter.unique.precluster.abund.fasta
```

L'étape d'amplification par PCR d'un mélange d'acide nucléotidique de séquence proche peut générer des chimères (une séquence issue de l'amplification d'acide nucléique provenant de deux espèces différentes). Ces séquences chimériques ne sont pas

représentatives d'espèce existant réellement. Ainsi, il convient de les éliminer par utilisation de la commande *chimera.uchime*. Le principe de fonctionnement d'*Uchime* consiste à diviser les données par échantillon et à détecter la présence des chimères en prenant comme référence les séquences abondantes. Les séquences abondantes étant moins susceptibles d'être des artefacts de PCR et donc d'être des chimères. Pour cela, chaque séquence du groupe est alignée aux séquences de référence. Si celle-ci s'aligne avec deux séquences de référence, elle sera considérée ainsi comme chimérique. Si une séquence est signalée comme chimérique dans un échantillon, elle sera enlevée de tous les échantillons (*dereplicate=F*). Dans nos études, nous avons utilisé l'option (*dereplicate=T*) afin d'éliminer les chimères par échantillon et non par rapport à l'ensemble des échantillons.

```
mothur > chimera.uchime
(fasta=Stability.trim.contigs.good.unique.good.filter.unique.precluster.abund.fasta,
count=Stability.trim.contigs.good.unique.good.filter.unique.precluster.abund.count_table,
dereplicate=t)
Fichiers de sortie :
stability.trim.contigs.good.unique.good.filter.unique.precluster.abund.uchime.pick.count_
table
stability.trim.contigs.good.unique.good.filter.unique.precluster.abund.uchime.chimeras
stability.trim.contigs.good.unique.good.filter.unique.precluster.abund.uchime.accnos
```

L'exécution de *chimera.uchime* avec le fichier *.count* permet de supprimer les séquences chimériques du fichier *.count* mais pas du fichier *.fasta*. Pour ce faire, nous avons utilisé la commande *remove.seqs* pour supprimer ces séquences listées dans le fichier *.accnos* du jeu de données.

```
mothur > remove.seqs
(fasta=Stability.trim.contigs.good.unique.good.filter.unique.precluster.abund.fasta, accnos=
Stability.trim.contigs.good.unique.good.filter.unique.precluster.abund.uchime.accnos)
Fichier de sortie :
Stability.trim.contigs.good.unique.good.filter.unique.precluster.abund.pick.fasta
```

Après toutes ces étapes, le jeu de données est de nouveau contrôlé avec la commande *summary.seqs*.

```
mothur > summary.seqs (fasta=current, count=current)
Using Stability.trim.contigs.good.unique.good.filter.unique.precluster.abund.pick.fasta as
input file for the fasta parameter.
```

Fichier de sortie:

```
Stability.trim.contigs.good.unique.good.filter.unique.precluster.abund.pick.summary
```

Par la suite, les séquences ont été pré-affiliées par rapport à une base de données de références (PR2 dans les deux cas d'études), par utilisation de la commande *classify.seqs*.

```
mothur > classify.seqs
```

```
(fasta=Stability.trim.contigs.good.unique.good.filter.unique.precluster.abund.pick.fasta,  
count=Stability.trim.contigs.good.unique.good.filter.unique.precluster.abund.uchime.pick.c  
ount_table, reference=mothur_qiime_gb203.fasta, taxonomy=gb203_pr2.tax.txt, cutoff=80)
```

Fichiers de sortie:

```
Stability.trim.contigs.good.unique.good.filter.unique.precluster.abund.pick.tax.wang.  
taxonomy
```

```
Stability.trim.contigs.good.unique.good.filter.unique.precluster.abund.pick.tax.wang.tax.  
summary
```

Après que toutes les séquences soient classées, nous avons utilisé la commande *remove.lineage* afin de supprimer toutes les séquences parasites qui ne doivent être dans le jeu de données (comme les séquences de bactéries, de chloroplastes ou autres). Dans notre cas, aucune séquence n'a été supprimée, ce qui suggère la qualité du choix de nos primers.

```
mothur >
```

```
remove.lineage
```

```
(fasta=Stability.trim.contigs.good.unique.good.filter.unique.precluster.abund.pick.fasta,  
count=Stability.trim.contigs.good.unique.good.filter.unique.precluster.abund.uchime.pick.  
count_table,  
taxonomy=Stability.trim.contigs.good.unique.good.filter.unique.precluster.abund.pick.tax.  
wang.taxonomy, taxon=Bacteria-Archaea)
```

Fichiers de sortie

```
Stability.trim.contigs.good.unique.good.filter.unique.precluster.abund.pick.tax.wang.pick.ta  
xonomy
```

```
Stability.trim.contigs.good.unique.good.filter.unique.precluster.abund.pick.pick.fasta
```

```
Stability.trim.contigs.good.unique.good.filter.unique.precluster.abund.uchime.pick.pick.cou  
nt_table
```

Afin de regrouper les OTU selon leur séquence, nous avons utilisé la commande *cluster.split*. Cette commande utilise les informations taxonomiques afin de regrouper les séquences en classes. Cette commande est une meilleure alternative à l'approche traditionnelle consistant à utiliser les commandes *dist.seqs* (permettant de calculer les distances entre chaque séquence alignée et de générer une matrice de distances et à la



commande *cluster* permettant de regrouper les séquences par OTU à différents niveaux d'identité (100%, 99%, ..., 90%). En effet, dans notre cas, la taille importante du jeu de données due à un grand nombre de séquences, nous impose l'usage de la commande *cluster.split* car elle est plus rapide, utilise moins de mémoire et est exécutable sur plusieurs processeurs en parallèle. Dans cette commande nous utilisons *taxlevel=4*, ce qui correspond au niveau de l'ordre. Le fichier *.list* généré comporte, pour tout niveau d'identité, les identifiants des séquences regroupées au sein de chaque OTU.

```
mothur >
cluster.split(fasta=StabilityEuk.trim.contigs.good.unique.good.filter.unique.precluster.abund.pick.pick.fasta,
count=StabilityEuk.trim.contigs.good.unique.good.filter.unique.precluster.abund.uchime.pick.pick.count_table,
taxonomy=StabilityEuk.trim.contigs.good.unique.good.filter.unique.precluster.abund.pick.pick.tax.wang.pick.taxonomy, splitmethod=classify, taxlevel=4, cutoff=0.15, processors=2)

Fichier de sortie :
StabilityEuk.trim.contigs.good.unique.good.filter.unique.precluster.abund.pick.pick.an.unique_list.list
```

A cette étape du traitement, nous avons utilisé la commande *make.shared* afin de créer le tableau final *.shared* comportant toutes les OTU, à un seuil d'identité précis (ici *label=0.03*, ceci signifie que le regroupement de deux séquences au sein d'une seule OTU ne s'effectue que si leur similarité dépasse les 97%), ainsi que leur abondance respective (en terme de nombre de séquences) dans chaque échantillon.

```
mothur > make.shared
(list=StabilityEuk.trim.contigs.good.unique.good.filter.unique.precluster.abund.pick.pick.an.unique_list.list,
count=StabilityEuk.trim.contigs.good.unique.good.filter.unique.precluster.abund.uchime.pick.pick.count_table, label=0.03)

Fichier de sortie :
StabilityEuk.trim.contigs.good.unique.good.filter.unique.precluster.abund.pick.pick.an.unique_list.shared
```

A la fin du processus Mothur et avant toute analyse des jeux de données, nous avons éliminé les « single singleton ». Ces derniers sont des OTU représentées que par une seule séquence dans l'ensemble de la base de données, et sont donc suspectées d'être des séquences erronées. Les « single singletons » ont été supprimés dans plusieurs études, afin

d'éviter la surestimation de la richesse spécifique (Monchy et al., 2012 ; Christaki et al., 2014, 2017 ; Georges et al., 2014 ; Genitsaris et al., 2015, 2016).

A la fin du processus, les jeux de données ont été normalisés. Le nombre de séquences a été standardisé par échantillon, de manière à ce que tous les échantillons aient le même nombre de séquences. Pour ce faire, nous avons déterminé le plus petit nombre de séquence contenu dans un échantillon (size=2082 et 5366 pour la Manche Orientale et la Méditerranée, respectivement), servant comme base de normalisation. Afin de générer le fichier de séquences normalisées, nous avons lancé la commande *sub.sample*.

```
mothur > sub.sample (shared=Somlit/SPECIMED_Euk.shared, size=2082 ou 5366)
Fichier de sortie
Somlit_Euk.0.03.subsample.shared ou SPECIMED_Euk.0.03.subsample.shared
```

Après le *sub.sample* nous avons lancé la commande *dist.seqs* afin de calculer les distances entre chaque séquence alignée. Par défaut, *dist.seqs* génère une matrice de distance triangulaire au format *phylip* (option : output=lt).

```
mothur > dist.seqs
(fasta=StabilityEuk.trim.contigs.good.unique.good.filter.unique.precluster.abund.pick.pick.
fasta, output=lt, processors=3)
Fichier de sortie:
StabilityEuk.trim.contigs.good.unique.good.filter.unique.precluster.abund.pick.pick.phylip
.dist
```

Après calcul des distances, la commande *get.oturep* a été exécutée afin de générer un fichier de séquences au format *fasta* contenant uniquement une séquence représentative (déterminée en fonction de la matrice de distance précédemment générée) pour chaque OTU. Un fichier *.rep.fasta* et *.rep.names* est généré pour chaque niveau de similarité entre séquences représentatives d'une OTU (1%, 2%, 3%, ..., 10%).

```
mothur > get.oturep
(phylip=StabilityEuk.trim.contigs.good.unique.good.filter.unique.precluster.abund.pick.pic
k.phylip.dist, fasta=
StabilityEuk.trim.contigs.good.unique.good.filter.unique.precluster.abund.pick.pick.fasta,
list=StabilityEuk.trim.contigs.good.unique.good.filter.unique.precluster.abund.pick.pick.an
unique_list.list)
StabilityEuk.trim.contigs.good.unique.good.filter.unique.precluster.abund.pick.pick.an.uniq
ue_list.unique.rep.names
...
```

```
StabilityEuk.trim.contigs.good.unique.good.filter.unique.precluster.abund.pick.pick.an.uniq  
ue_list.0.01.rep.fasta
```

...

- *Affiliation taxonomique*

Toutes les OTUs (*SOMLIT.fasta* ou *SPECIMED.fasta*) ont été affiliées par comparaison avec les séquences présentes dans les bases de données PR2 (*Guillou et al., 2013*) et Silva (*Quast C. et al., 2013*). Pour choisir une affiliation, nous nous sommes basés sur la/les séquence(s) ayant un maximum de similarité avec une séquence présente dans les bases de données et correspondant à un organisme cultivé et clairement identifiable. Nous avons également vérifié manuellement cette annotation par comparaison des OTU (essentiellement celles plus abondantes ou exprimant un intérêt) avec celles se trouvant dans la base de données GenBank ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)).

L'affiliation a été assurée à l'aide du programme BLAST (*Altschul et al., 1990*) et se présentait sous forme d'un tableau avec pour chaque OTU la séquence la plus proche dans la base de donnée, ainsi que les statistiques d'alignement ( $E_{value}$ , Score, pourcentage de similarité...). La commande pour l'affiliation des séquences par rapport à la base de données PR2 était la suivante :

```
Blastn-db PR2.fasta -query SOMLIT/SPECIMED.fasta -outblast_SOMLIT/SPECIMED-  
PR2.txt -outfmt 6
```

	MO (ADN & ARN)	SPECIMED (ADN)
Nombre de séquences avant normalisation	5087016	4539054
screen.seqs	4062876	4428697
unique.seqs	Séq. Uniques : 557783	Séq. Uniques : 1899752
screen.seqs (après alignement)	Séq. Uniques : 336115 Séq. Totales : 1548565 start=1, end=2578	Séq. Uniques : 1858620 Séq. Totales : 4355800 start=19, end=2558
pre.cluster (diffs=2) + split.abund + chimera + remove.seqs	Séq. Uniques : 11691 Séq. Totales : 1420717 start=1, end= 1104	Séq. Uniques : 61245 Séq. Totales : 3084392 start=1, end= 1257
classify.seqs + cluster.split + shared (0.03) + dist.seqs	199872 séq. 2039 OTU	1067834 séq. 7791 OTU
blast + normalisation (sans single singleton)	199440 séq. 952 OTU	1067237 séq. 7605 OTU
	ADN (59251 séq., 696 OTU) ARN (60557 séq., 700 OTU) Séq. Totales : 119808 OTU totales : 868	Séq. Totales : 248788 OTU totales: 3358 SOLA (59003 séq., 1648 OTU) MOLA (53497 séq., 1324 OTU) SOLEMIO (78690 séq., 1895 OTU) JULIO (57598 séq., 1560 OTU)
Echantillons utilisés dans l'étude		

**Tableau 2 :** Nombre de séquences après chaque étape de traitement des jeux de données sous Mothur. MO : concerne les échantillons prélevés en Manche Orientale entre 2013-2015 et dont l'ADN et l'ARN ont été extraits. SPECIMED : concerne les échantillons prélevés en Golfe du Lion entre 2010-2014. Pour ces derniers, juste les échantillons prélevés entre 2012-2014 ont été intégrés dans l'étude.  
Séq.: Séquence, OTU : Unité taxonomique Opérationnelle

## II. APPROCHE MICROSCOPIQUE

L'identification des protistes est très complexe. Cette complexité est due majoritairement à leur grande diversité phylogénétique mais également à leurs tailles variées (allants de 0.2µm à quelques mm), leur énorme abondance pouvant atteindre des milliers de cellules par cm<sup>3</sup> ainsi qu'à leur morphologie. De par leur grande diversité, leur variété de taille et leur morphologie dans certains cas semblable, leur identification précise est difficile. Néanmoins, la microscopie (inversée ou autres), depuis sa mise en place, reste la seule technique pouvant fournir des données quantitatives (e.g. dimensions) relatives aux espèces présentes dans l'échantillon traité, permettant d'évaluer leur biovolume et leur abondance qu'on peut convertir en biomasse selon des facteurs propres à la taille et/ou l'espèce. Malgré le fait que la microscopie possède de nombreux avantages, principalement quantitatifs, elle a également plusieurs inconvénients. La microscopie est une technique chronophage, l'analyse d'un échantillon peut prendre de 2 à 6h (tout dépend de l'écosystème étudié, la saison, etc.). Elle nécessite également une grande expertise de la taxinomie des organismes : l'identification n'est pas toujours possible en raison des morphologies hyper-semblables des espèces appartenant à la même famille (e.g. la famille des Gymnodiniaceae : cas de *Gymnodinium sp.*, *Katodinium sp.*, *Nematodinium sp.*). Par exemple, la disposition des plaques chez les dinoflagellés ou encore l'insertion du flagelle, sont des critères d'identification importants qui ne sont pas souvent visibles. Un autre point majeur pouvant influencer la bonne identification des organismes est celui lié aux artefacts causés par la fixation au Lugol, pouvant induire par exemple le rétrécissement des cellules (Leakey et al., 1994 ; Stoecker et al., 1994). De plus, quelques études ont pu montrer une surestimation des organismes de grande taille (20 à 40 µm) en dépit du nanoplancton, engendrée par la formation d'agrégats suite à la fixation au Lugol (Zarauz et al., 2008). Et finalement, l'observateur doit avoir une grande maîtrise des techniques de comptage. Tous ces inconvénients peuvent avoir des conséquences sur des résultats selon le savoir-faire de l'observateur (Tillmann et Rick, 2003).

Dans ce projet de thèse, le comptage a concerné les échantillons provenant de la Méditerranée Occidentale (SOLA, MOLA, SOLEMIO, JULIO ; **Fig.1-Chapitre 5**). Les échantillons ont été prélevés dans 4 stations situées dans le Golfe du Lion entre 2012 et 2014. Les échantillons ont été fixés au Lugol (2%) et les comptages du microplancton ont été effectués avec un Nikon EclipseTE2000-s au grossissement x400 (voir partie site

d'échantillonnage et prélèvement, **Chapitre 5**). L'identification des organismes a été basée sur plusieurs références bibliographiques (*Petz, 2005; McMinn & Scott, 2005; Hoppenrath et al., 2009*). Un trombinoscope des organismes observés dans l'ensemble des stations a été également réalisé faisant partie des informations supplémentaires reporté dans le **Chapitre 5**.

### III. MESURES DES PARAMETRES PHYSICO-CHIMIQUES

#### III.1. Les paramètres physico-chimiques de la Manche Orientale

##### III.1.1. Mesures *in situ*

Des profils verticaux de la température (T ; °C) et la salinité (S ; p.s.u.) ont été réalisés *in situ* à l'aide d'une sonde CTD Seabird 19 SeaCat. Le pH a été mesuré *in situ* à l'aide d'un pH1970i -mètre (WTW). La luminosité (PAR ;  $\mu\text{mol.m}^2.\text{s}^{-1}$ ) a été mesurée par un capteur PAR (QSP 2300, Biospherical Instruments) couplé à la CTD. Les données correspondants aux prélèvements de surface (e.g. température, salinité) ont été moyennées sur le premier mètre de la colonne d'eau.

Le coefficient d'atténuation  $k$  a été déterminé à partir de la relation entre la luminosité ( $E_z$ ), la luminosité de surface ( $E_0$ ) et la profondeur de prélèvement  $z$  (m) en utilisant les données de PAR :

$$E_z = E_0 e^{(-kz)}$$

Dont  $k$  représente la pente de régression linéaire entre  $\ln E_z$  et la profondeur  $z$ .

##### III.1.2. Sels nutritifs

Les prélèvements d'eau de mer ont été effectués à l'aide d'une bouteille Niskin (5L) en sub-surface (1m). Pour tout prélèvement, des échantillons d'eau ont été conservés en triplicat dans des tubes à hémolyse en polyéthylène (10ml), stockés à -18°C jusqu'au moment de l'analyse des ions nitrates ( $\text{NO}_3^-$ ), nitrites ( $\text{NO}_2^-$ ) et phosphates ( $\text{PO}_4^{3-}$ ) et à 4°C pour les ions silicates ( $(\text{SiOH}_4)^{2-}$ ) (*Aminot et Chaussepied, 1983*).

Les nitrates, nitrites, silicates et phosphates ont été dosés à l'aide d'une chaîne de dosage semi-automatique Alliance® Integral Futura, selon les protocoles de Tréguer et le Corre (1974). Le dosage est basé sur des réactions colorimétriques avec lecture par spectrophotométrie d'absorption (*Aminot et Chaussepied, 1983*).

Les méthodes analytiques de références sont respectivement :

- Pour les ions ( $\text{NO}_2^-$ ), la méthode de dosage appliquée est celle proposée par Bendshneider et Robinson (1952) fondée sur la réaction de Griess. Ces ions forment un composé diazoïque avec le sulfanilamide en milieu acide ( $\text{pH} < 2$ ). Le composé ainsi produit réagit avec le N-naphtyl-éthylènediamine pour former un complexe coloré (rose). La concentration en nitrites est mesurée à 543 nm à l'aide d'un spectrophotomètre intégré à l'appareil Alliance.
- Pour les ions nitrates ( $\text{NO}_3^-$ ), le dosage a été effectué selon la méthode de Woods et al. (1967). Les ions nitrates sont réduits en nitrites après passage dans une colonne réductrice (cadmium et cuivre). Après dosage des ions nitrites ainsi formés, la concentration en nitrates est déterminée par soustraction.
- Pour les ions silicates ( $(\text{SiOH}_4)^{2-}$ ), le dosage suit la méthode développée par Mullin et Riley (1962). Les ions silicates réagissent avec le molybdate d'ammonium en milieu acide pour former un complexe silicomolybdique, réduit par la suite par l'acide ascorbique pour former un précipité bleu. Ce dernier est mesuré à une longueur d'onde de 810 nm.
- Pour les ions phosphates ( $\text{PO}_4^{3-}$ ), sont dosés suivant la méthode proposée par Murphy et Riley (1962). Les ions phosphates réagissent en milieu acide avec le molybdate d'ammonium formant l'acide molybdophosphorique. En présence d'antimoine, ce dernier est réduit par l'acide ascorbique et génère un complexe phosphomolybdique bleu. Le précipité bleu ainsi obtenu est mesuré à 885 nm.

### **III.1.3. Carbone et azote organiques particuliers, matières en suspension et oxygène**

Le carbone organique particulaire (POC,  $\text{mg L}^{-1}$ ) et l'azote organique particulaire (PON,  $\text{mg L}^{-1}$ ) sont déterminés par combustion du matériel particulaire récupéré par combustion. Le contenu organique de chacun des composés a été évalué sur un analyseur CHN NA2100 Frisons (NA2100 Frisons Analyzer) après séchage des filtres à 60°C pendant 24h et

exposition aux vapeurs de HCl 1N pendant 5h. La matière en suspension (SPM, mg L<sup>-1</sup>) a été déterminée par pesée avant et après filtration au travers de filtres GF/F (porosité environ 0.7 µm) en filtrant un volume d'eau suffisant pour avoir 0.5 à 1 mg de matériel déposé sur le filtre. Les concentrations en oxygène (ml/L) ont été mesurées en triple par microtitration de Winkler (1988) (*Aminot et Kerouel, 2004*). Cette méthode exploite la capacité d'un sel de manganèse à réagir en milieu fortement basique avec l'oxygène dissous de l'échantillon. En abaissant le pH de cette réaction, les ions iodure préalablement ajoutés dans le réactif produisent de l'iode. Ce dernier composé est dosé par une solution étalonnée de thiosulfate.

#### **III.1.4. Chlorophylle *a***

Les concentrations en Chl *a* (µg. L<sup>-1</sup>) sont mesurées selon les équations de Lorenzen (1967). Des échantillons de 0,5 à 1 L d'eau de mer sont filtrés sur des filtres en fibre de verre (0,7 µm; Whatman GF/F). En attendant l'analyse, les filtres sont placés à l'obscurité, au congélateur à -20°C. L'extraction des pigments photosynthétiques a lieu dans l'acétone à 90% (8 mL, 4°C pendant 12h) après laceration des filtres (*Aminot et Kerouel, 2004*). Les broyats de filtre sont ensuite centrifugés (3000 tours/min, à 4°C pendant 15 min) et le surnageant est récupéré pour lecture de la fluorescence au fluorimètre (Turner Designs model 10-AU). La lecture des extraits chlorophylliens s'effectue avant et après acidification à l'acide chlorhydrique (1%) pour évaluer la fraction des phéopigments.

Pour plus de détails concernant l'ensemble des mesures, voir <http://somlit.epoc.u-bordeaux1.fr/fr/>.

Il est à noter que les mesures à bord ont été réalisées par Eric Lecuyer (technicien CNRS-LOG ; <http://log.cnrs.fr/Lecuyer>) ainsi que les différents dosages avec l'aide de Muriel Crouvoisier (assistante ingénieure CNRS-LOG ; <http://log.cnrs.fr/Crouvoisier>).

### *III.2. Les paramètres physico-chimiques de la Méditerranée Occidentale*

#### **III.2.1. Température, salinité et éléments nutritifs**

La température (°C) et la salinité de l'eau de mer ont été mesurées par une CTD Seabird 19 SeaCat. Les échantillons d'éléments nutritifs (NO<sub>3</sub>, NO<sub>2</sub>, PO<sub>4</sub> et SiOH<sub>4</sub>, µM) ont été



recueillis dans des flacons en polyéthylène propres de 20 ml et empoisonnés avec du HgCl<sub>2</sub>, avant stockage à 4°C. Les concentrations ont été déterminées à l'aide de l'auto-analyseur Brann Luebbe selon Aminot et Kerouel (2007). Ces mesures ont été assurées par Olivier Grosso (OCEANOMED).

### **III.2.2. Chlorophylle *a***

Entre 0.5 et 2 L d'eau de mer ont été recueillis à partir de bouteilles Niskin dans des bouteilles en polycarbonate en surface (1-5m) pour la chlorophylle *a* (Chl *a*, µg L<sup>-1</sup>) et stockés à -20°C jusqu'à l'analyse. Les pigments chlorophylliens ont été filtrés sur un filtre GFF de 25 mm (non brûlé) puis extraits pendant 24h dans 90% d'acétone suivant le protocole fluorimétrique d'Aminot et Kerouel (2004). Les mesures de la Chl *a* ont été assurées par Sandra Helias-Nunige (OCEANOMED).

### **III.2.3. Carbone et azote organiques particuliers**

Entre 1 à 2 L d'eau de mer ont été recueillis à partir de bouteilles Niskin dans des bouteilles en polycarbonate puis stockées au froid jusqu'à la filtration. Le carbone organique particulaire (POC, µg L<sup>-1</sup>) et l'azote organique particulaire (PON, µg L<sup>-1</sup>) ont été filtrés sur des filtres GFF de 25 mm puis traités à l'aide d'un analyseur élémentaire CHN Perkin Elmer 2400 Series CHNS/O. Tous les filtres ont été séchés au four à 60°C. L'analyse des composés organiques particulaire a été assurée par Karine Leblanc (OCEANOMED).

### **III.2.4. Mesures supplémentaires**

Aux stations SOLA et MOLA, des mesures supplémentaires ont été également réalisées. Les concentrations en oxygène dissous (ml/L) ont été mesurées par microtitration, telle proposée par Winkler (1888) (*Aminot et Chaussepied, 1983*). Les concentrations en matières en suspension (MES ou SPM, mg L<sup>-1</sup>) ont été déterminées par pesée avant et après filtration au travers de filtres GF/F (porosité d'environ 0.7 µm) (pour plus de détails: <http://somlit.epoc.u-bordeaux1.fr/fr/>).

#### IV. ANALYSES STATISTIQUES

La diversité a été caractérisée par le calcul des indices de diversité ( $S_{\text{chao1}}$ , Simpson et équitabilité) avec le logiciel Past 3.05 (Hammer et al., 2001) pour tous les échantillons provenant de la Manche Orientale. L'approche de  $S_{\text{chao1}}$  consiste à utiliser le nombre de singletons et de doubletons afin d'estimer le nombre théorique d'OTU (Unité taxonomique opérationnelle, proxy d'espèce en biologie moléculaire) attendu sur la base des singletons et des doubletons dans chaque échantillon. L'indice de Simpson (D) mesure la probabilité que deux individus soient de la même OTU et varie entre 0 et 1. Plus il se rapproche de 0, plus les chances d'obtenir, par tirage aléatoire, des individus d'OTU différentes sont élevés. L'indice d'équitabilité mesure la régularité avec laquelle les individus sont répartis entre les taxons selon  $H'/H'_{\text{max}}$  ( $H'$  : indice de Shannon). Ces indices de diversité ont été comparés pour chaque date avec une «*procédure de randomisation par bootstrapping*» et la  $p_{\text{value}}$  a été calculée sur la base de 1000 paires de permutations aléatoires.

Les assemblages microbiens ont été classés à travers les dates d'échantillonnage par analyse de classification hiérarchique en utilisant PRIMER 6.0 (Clarke et Gorley, 2006). Les coefficients de dissimilarité de Bray-Curtis ont été calculés pour construire une matrice basée sur l'abondance des OTU. Un test de profil de similarité (SIMPROF) a été réalisé pour définir des groupes saisonniers significatifs.

Le reste des analyses a été effectué sous R (v 3.2.3). **(1)** L'analyse en Composante Principale (ACP) a été effectuée pour explorer les inter-corrélations entre les variables environnementales, en utilisant le package «*ade4*» (Dray and Dufour, 2007). Pour se rapprocher d'une distribution normale des données, la transformation BoxCox (Box and Cox, 1964), a été appliquée en utilisant le package «*caret*» (Kuhn, 2016). La clusterisation des dates/sites a été effectuée avec la méthode hiérarchique de Ward basée sur la variance minimale (Ward, 1963) (**Annexe 2**). **(2)** Pour déterminer le degré de couplage entre les variations environnementales et celles des abondances des protistes une analyse de Co-Inertie (COIA, Dolédec & Chessel 1994, Dray et al. 2003) a été effectuée avec le package «*ade4*» (**Annexe 3**). Elle consiste à trouver des axes de Co-inertie qui maximisent les covariances entre les coordonnées de lignes (dates) des matrices de variables environnementales et celles d'abondances des groupes taxonomiques. Elle a été calculée à partir d'une ACP normalisée sur les variables environnementales et une ACP centrée sur

les abondances. Le degré de couplage a été évalué avec le coefficient de corrélation multidimensionnelle (RV) et en effectuant un test de Monte-Carlo. Les données environnementales et d'abondance ont été transformées en  $\log(x+1)$  afin de minimiser l'influence des valeurs élevées et pour assurer une distribution normale (*Legendre & Legendre, 1998*). Les données environnementales ont de plus été normalisées pour les rendre sans dimension (z-scores, *Kenkel, 2006*). La significativité statistique a été fixée à  $p < 0.05$ . Cette analyse a été appliquée sur les données issues du séquençage des échantillons provenant de la Méditerranée Occidentale.

**ANNEXE 1 :** Les différentes étapes de traitement des 47 échantillons Pyro-séquencés telles que renseignées dans *Genitsaris et al., 2015*. Les échantillons ont été prélevés bimensuellement (sauf en cas de conditions environnementales difficiles) à la station SOMLIT en Manche Orientale entre Mars 2011 et Juillet 2013.

*1<sup>ère</sup> étape : Extraction d'ADN ribosomique*

L'ADN ribosomique des organismes protistes a été extrait et purifié avec le Kit d'isolement d'ADN PowerWaterDNA (MoBio Laboratories Inc., CA, USA), en suivant le protocole du fabricant. Les échantillons traités contenaient entre 0.5 et 4.5 ng.µL<sup>-1</sup> d'ADN. Les dosages des quantités d'ADN ont été réalisés avec le Qubit 2.0 Fluorometer (Thermo Fisher Scientific Inc., Massachusetts, USA).

*2<sup>ème</sup> étape : Amplification et Pyro-séquençage*

Les échantillons d'ADN ont été amplifiés en utilisant les mêmes amorces que dans l'amplification des échantillons séquencés par la technique Illumina : 18S-82F (5'-GAAACTGCGAATGGCTC-3', *Lopez-Garcia et al., 2003*) et Euk-516r (5'-ACCAGACTTGCCCTCC-3', *Amann et al., 1990*). Ces amorces ont été utilisées afin d'amplifier les régions variables des eucaryotes V2 et V3 du gène de l'ARNr18S (environ 470-480 pb). Il a été également ajouté à ces amorces, un code barre de 10pb spécifique à chaque échantillon, une clé TCAG 160 de 4pb et un adaptateur GsFLX de 26pb.

L'amplification en chaîne par polymérase (ACP ou PCR) a été réalisée selon les conditions standards pour l'ADN Polymérase à haute fidélité « Platinum Taq » (Invitrogen, Carlsbad, CA, USA), avec 5 ng d'ADN environnemental comme matrice, en utilisant l'appareil GeneAmp PCR System (Applied Biosystems, Foster City, Californie, États-Unis). Après l'étape de dénaturation à 95°C pendant 5 min, 30 cycles d'amplification ont été effectués à 95°C et 50°C pendant 30 secondes, respectivement, puis à 72°C pendant 1 minute. Une dernière étape d'extension de 7 min à 72°C a été incluse.

La banque d'amplicons a été préparée selon les procédures décrites par Roche (Bâle, Suisse) et utilisée dans une seule plaque sur un séquenceur de 454 GsFLX Titanium. Les Pyro-séquences ont été soumises à GenBank-SRA sous le numéro d'accès SRX768577. Il

est à noter que l'étape du Pyro-séquençage a été réalisée par **GenoScreen** (Institut Pasteur, Lille, France)

**ANNEXE 2 :** Script élaboré pour la réalisation des différentes Analyses en Composante Principale (ACP) appliquées sur les variables environnementales des différentes stations étudiées.

```
library (ade4)
library (caret)
setwd ("D:/")
Env <- read.table ("Fichier Env.txt", header=TRUE, sep="\t", na.strings="NA", dec=".",
strip.white=TRUE, row.names=1, fill=TRUE)
Env <- as.matrix (Env)
is.numeric (Env)
Env = as.data.frame (Env)
transenv <- preprocess(Env, c("BoxCox", "center", "scale"))
envTrans = data.frame(transenv = predict(transenv, Fichier Env))
acp <- dudi.pca(envTrans, scale=TRUE, scan = F, nf=3)
print (acp$eig)
summary (acp)
inertia.dudi (acp,col.inertia = T)$col.abs
print (acp$co)
print (acp$li)
Cluster <- hclust(dist(acp$li), method = "ward.D")
plot(Cluster)
date.group <- as.factor(cutree(Cluster, k = 2))
print(date.group)
s.class(acp$li, date.group, col=c("cornflowerblue", "darkorchid2", "violetred", "
aquamarine2"))
s.arrow(4*acp$co, clab = 0.7, add.plot = TRUE)
```

**ANNEXE 3 :** Script des analyses de co-inertie appliquées sur les grands groupes taxonomiques et les variables environnementales des échantillons prélevés au niveau des quatre stations du Golfe du Lion (SOLA, MOLA, SOLEMIO et JULIO).

RV : coefficient de corrélation multidimensionnelle

Env. : tableau des paramètres environnementaux

Taxonomie : tableau d'abondances des différents groupes taxonomiques

```
library (ade4)
```

```
library (caret)
```

```
taxonomie <- read.table ("D:/Fichier taxons.txt", header=TRUE, sep="\t", na.strings="NA",  
dec=".", row.names=1, strip.white=TRUE)
```

```
taxonomie <- as.matrix (taxonomie)
```

```
is.numeric (taxonomie)
```

```
summary (taxonomie)
```

```
print (taxonomie)
```

```
pca_taxonomie <- dudi.pca (taxonomie, scale=TRUE, scan=FALSE, nf=3)
```

```
scatter (pca_taxonomie)
```

```
summary (pca_taxonomie)
```

```
inertia.dudi (pca_taxonomie, col.inertia = T)$col.abs
```

```
print (pca_taxonomie$co)
```

```
print (pca_taxonomie$li)
```

```
Env <- read.table ("D:/Fichier Env.txt", header=TRUE, sep="\t", na.strings="NA", dec=".",  
row.names=1, strip.white=TRUE)
```

```
Env <- as.matrix (Env)
```

```
is.numeric (Env)
```

```
Env = as.data.frame (Env)
```

```
transenv <- preProcess (Env, c ("BoxCox", "center", "scale"))
```

```
envtrans <- data.frame (transenv = predict (transenv, Fichier Env))
```

```
pca_env <- dudi.pca (envtrans, scale=TRUE, scan=FALSE, nf=3)
```

```
scatter (pca_env)
```

```
summary (pca_env)
```

```
inertia.dudi (pca_env, col.inertia = T)$col.abs
```

```
print (pca_env $co)
```

```
print (pca_env $li)

coin_taxo_env <- coinertia (pca_taxonomie, pca_env, scan = FALSE, nf = 3)
print (coin_taxo)
summary (coin_taxo_env)
print (coin_taxo_env $co)
print (coin_taxo_env $li)
inertia.dudi (coin_taxo_env, col.inertia = T)$col.abs
inertia.dudi (coin_taxo_env, row.inertia = T)$row.abs
plot (coin_taxo_env)
rv <- RV.rtest(pca_env_$stab, pca_taxonomie$stab, 999)
plot (rv)
print (rv)
```





# **CHAPITRE III**

**DIVERSITE ET ACTIVITE DES EUCARYOTES MICROBIENS DANS UN ECOSYSTEME  
COTIER MESO-EUTROPHE (MANCHE ORIENTALE)**



**Diversity and potential activity patterns of planktonic eukaryotic microbes in a mesoeutrophic coastal area (eastern English Channel) (En presse)**

**S. RACHIK, U. CHRISTAKI, L.L. LI, S. GENITSARIS, E. BRETON, S. MONCH**

**RESUME DU CONTEXTE ET OBJECTIFS**

L'étude de la diversité et de l'activité métabolique des eucaryotes microbiens est fondamentale dans la compréhension du fonctionnement des écosystèmes (Caron, 2013), en raison de leur large diversité ainsi que leur rôle important dans les écosystèmes marins (Caron et al., 2012). Avec l'avènement des techniques de séquençage à haut débit (e.g. Illumina, Pyro-séquençage), une vaste diversité a pu être détectée.

Dans cette première étude, nous nous sommes intéressés à l'étude de ces eucaryotes microbiens en Manche Orientale. Cet écosystème est sujet à des proliférations récurrentes de l'Haptophyte *Phaeocystis globosa* au printemps (Monchy et al., 2012). Les études conduites dans cet écosystème ont bien élucidé son impact sur la structuration des communautés planctoniques (par exemple : bactéries (Lamy et al., 2009), microplancton (Grattepanche et al. 2011a, b)). La Manche Orientale a été également le siège des études de diversité et de succession saisonnière des communautés eucaryotes basées sur l'ADNr (par exemple : Genitsaris et al., 2015, 2016). Toutefois, aucune étude de leur dynamique saisonnière et de leur activité cellulaire relative n'a été auparavant effectuée dans le siège d'étude.

Ce travail a été réalisé sur des échantillons prélevés entre Mars 2011 et Juillet 2015 à la station côtière (SOMLIT : Service d'Observation du Milieu LITtoral) de la Manche Orientale. Nous avons pour objectifs d'établir en premier temps une vue d'ensemble de la diversité des communautés eucaryotes sur une période de 4 ans et d'étudier également les relations pouvant exister entre la diversité et l'activité potentielle relative dans les communautés eucaryotes par calcul des rapports ARNr : ADNr.

## Abstract

The diversity of planktonic eukaryotic microbes was studied at a coastal station of the eastern English Channel (EEC) from March 2011 to July 2015 (77 samples) using high throughput sequencing (454-pyrosequencing and Illumina) of the V2-V3 hypervariable region of the 18S SSU rDNA gene. Similar estimations of OTU relative abundance and taxonomic distribution for the dominant higher taxonomic groups (contributing >1% of the total number of OTUs) were observed with the two methods (Kolmogorov-Smirnov  $p$ -value=0.22). Eight super-groups were identified throughout all samples: Alveolata, Stramenopiles, Opisthokonta, Hacrobia, Archeplastida, Apusozoa, Rhizaria, and Amoebozoa (ordered by decreasing OTU richness). To gain further insight into microbial activity in the EEC, ribosomal RNA was extracted for samples from 2013-2015 (30 samples). Analysis of 18S rDNA and rRNA sequences led to the detection of 696 and 700 OTUs, respectively. Cluster analysis based on OTUs' abundance indicated three major seasonal groups that were associated to spring, winter/autumn, and summer conditions. The clusters inferred from rRNA data showed a clearer seasonal representation of the community succession than the one based on rDNA. The rRNA:rDNA ratio was used as a proxy for relative cell activity. When all OTUs were considered, the average rRNA:rDNA ratio showed a linear trend around the 1:1 line, suggesting a linear relation between OTU abundance (rDNA) and activity (rRNA). However, this ratio was highly variable over time when considering individual OTUs. Interestingly, the OTU affiliated with *P. globosa* displayed rRNA:rDNA ratio that allowed to delimit high vs low abundance and high vs low activity periods. It unveiled quite well the *Phaeocystis* bloom dynamic regarding cell proliferation and activity, and could even be used as early indicator of an upcoming bloom.

**Keywords:** planktonic eukaryotic microbes, taxonomical succession, high-throughput sequencing, biodiversity, marine ecosystem, activity, rRNA: rDNA ratio, *Phaeocystis globosa*

## INTRODUCTION

Planktonic eukaryote microbes are abundant, ubiquitous in aquatic environments and extremely diverse in terms of taxonomy and metabolism [1]. They play a crucial role in the functioning of marine ecosystems through primary production (photosynthetic phytoplankton), predation (heterotrophs and mixotrophs) and/or symbiosis (mutualists and parasites) [1,2]. Studying their diversity and metabolic activity is, therefore, fundamental for understanding the functioning of marine ecosystems [3,4,5,6]. High-throughput sequencing targeting the 16S/18S rDNA genes coupled with analysis using efficient bioinformatics tools, made it possible to reveal the vast diversity of marine microbes. In marine ecosystems, this approach has been used to explore spatial and temporal prokaryote and eukaryote diversity in different regions worldwide [7,8,9,10,11,12] and revealed the diversity overlooked by conventional methods (e.g microscopy) [11,13,14]. Sequencing of rDNA allows detecting a large number of microorganisms despite their viability and activity. However, estimating microbial activity is fundamental to understanding the functioning of ecosystems. Generally, the quantity of rRNA is proportional to both the number of ribosomes and total RNA concentration [15,16], and can therefore be used as an indicator of cell activity which cannot be assessed targeting only rDNA. Several studies have applied this strategy to marine microbial communities of bacteria [17,18,19,20], archaea [21], and eukaryotes [10,12,22,23,24,25,26]. The approach is based on the calculation of rRNA:rDNA ratio in order to normalize the rRNA concentration according to the number of cells (DNA concentration), as rDNA is more constant per cell [15,27,28]. However, due to variations depending on microbial life strategies, life histories and non-growth activities, limitations exist for using this ratio to infer cells activity [29]. Rather, such ratio may potentially infer species-specific life-cycle events [29]. While favorable environmental conditions are known triggers for increased cell activities, several studies showed that gene regulatory networks governing microbial life strategy is more complex than initially thought. Indeed, microorganisms experiencing repeated patterns of changing environmental conditions develop “anticipatory life strategies” based on adaptive regulatory gene networks [30,31,32]. Cells use past conditions as predictive signals to anticipate upcoming conditions. For example, they accumulate rRNA during unfavorable conditions, in period of low activity, shortly before favorable conditions return [33,34]. To

our knowledge, such “anticipatory life strategy” has not been reported for eukaryote microbes in marine environments.

The present study was carried out in the eastern English Channel (EEC) in the framework of the SOMLIT network (French Network of Coastal Observatories). EEC is a meso-eutrophic coastal ecosystem presenting great seasonal fluctuations both in its biotic and abiotic environment. This ecosystem is characterized by recurrent massive blooms of *P. globosa* that develop under nitrogen-replete conditions and a silicate limitation at the end of winter [35,36,37], associated with sufficient underwater light intensity in order to achieve high growth rates [35,38]. *Phaeocystis* spp. exhibits alternate life cycles between solitary cells and gelatinous colonies [39], its bloom is preceded and followed by communities of colonial diatoms and dinoflagellate grazers (e.g [40,41,42,43] and references therein). In the EEC, the structure and seasonal succession of microbial eukaryotes community was thoroughly investigated through DNA-based high throughput sequencing [44,45,46]. However, information on the seasonal dynamic of relative cell activity is still lacking.

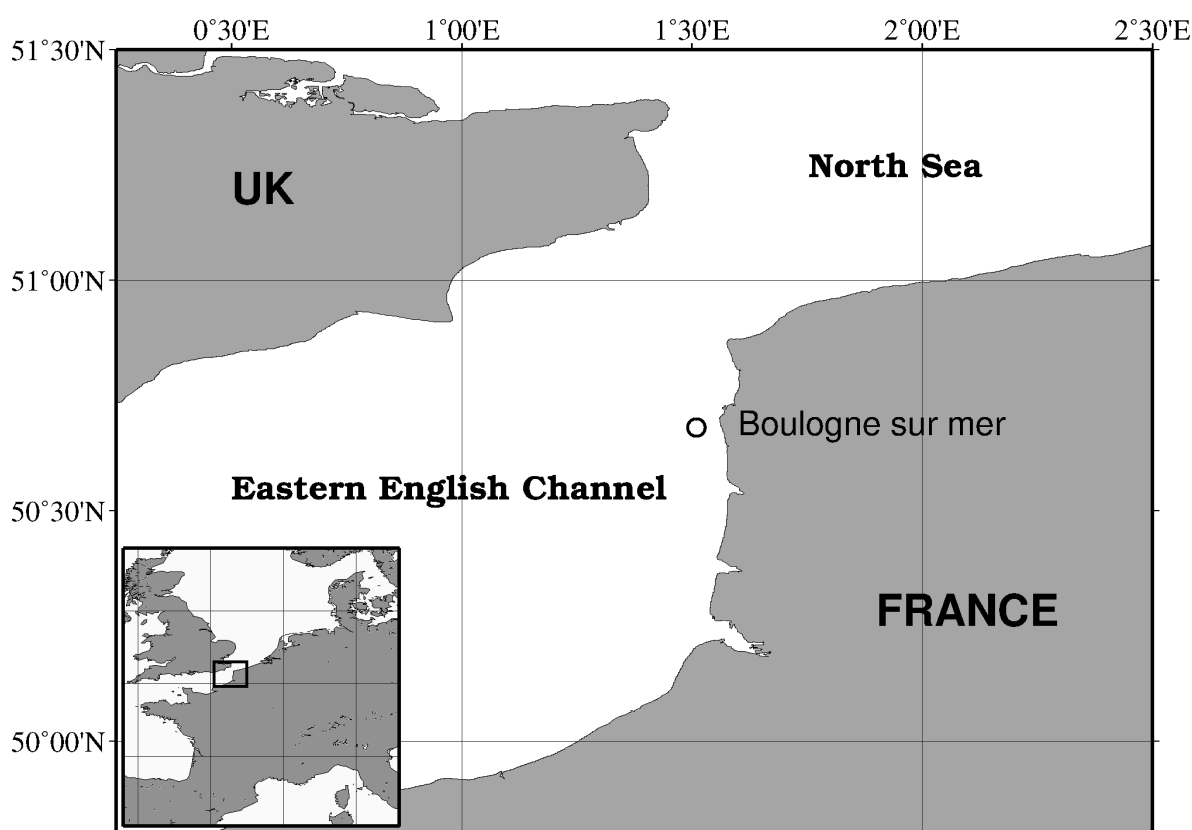
Our main objectives were to establish an overview of diversity in the mesoeutrophic coastal area of EEC and to study relationships between diversity and potential activity in microbial eukaryote communities by establishing rRNA:rDNA ratios. The study was conducted over 4 years at a coastal station of EEC from March 2011 to July 2015, using high-throughput sequencing. Specifically, the main questions were as follows: (i) what are the composition and succession of planktonic eukaryote communities in relation to environmental parameters in the EEC? (ii) What are the main differences/similarities between the rDNA and rRNA datasets? (iii) What is the ecological relevance of the rDNA:rRNA ratio values?

## **MATERIALS AND METHODS**

### **Sampling area**

The sampling was carried out at the SOMLIT station (French Network of Coastal Observatories; <http://somlit.epoc.u-bordeaux1.fr/fr/>) in the EEC, one mile from the coast (50°40' 75" N, 1° 31' 17" E; 20–25 maximum depth, Fig 1). For diversity assessment, subsurface samples (2-3 m water depth) were collected in 5 L polyethylene bottles, stored in the dark at *in situ* surface temperature, and filtered within 2h. Before filtration, the

seawater sampled was prefiltered using a 150- $\mu\text{m}$  mesh, in order to remove metazoa and large particles. Filtration, through 0.2  $\mu\text{m}$  nucleopore filters (47 mm diameter), was performed using a very low filtration pressure peristaltic pump (15 rpm) in order to avoid filter clumping and minimize organism disruption. The filters were immediately stored at -80°C until DNA and RNA extraction. A total of 77 samples were collected between March 2011 and July 2015 (12, 20, 25, 12 and 8 samples for each year, respectively).



**Fig 1. Study area in the eastern English Channel.** The dot (●) indicates the location of the sampling station.

### **Environmental variables**

Sampling was performed using Niskin bottles coupled to a CTD Seabird (SBE 19) for measuring Sea Surface Temperature (SST, °C), salinity (S), and the diffuse attenuation coefficient for down-welling irradiance ( $K_d$ ,  $\text{m}^{-1}$ ), using a biospherical PAR light sensor (QSP 2300, Biospherical Instruments) connected to the CTD. The average sub-surface daily light intensity ( $I$ ,  $\text{E m}^{-2} \text{d}^{-1}$ ) reaching phytoplankton was estimated using the formula (1) based on [47], where,  $Z$  was the depth at which samples were collected, and  $I_0$  was the



daily incident light estimated from global solar radiation (GSR,  $W m^{-2}$ ).  $I_0$  was measured continuously every 5 min with a solar radiation sensor (Vantage Pro, Davis).

$$I = \frac{I_0(1-e^{-k_d Z})}{k_d Z} \quad (1)$$

Oxygen concentrations ( $ml L^{-1}$ ) were measured in triplicate by Winkler microtitration [48]. The pH was determined with a pH1970i (WTW) pH meter. Nutrient samples (Nitrate  $NO_3^-$ , Nitrite  $NO_2^-$ , Phosphate  $PO_4^{3-}$  and Silicate  $SiOH_4$ ,  $\mu M$ ) were analyzed on Alliance Integral Futura Autoanalyzer II according to Aminot and Kerouel (2004). Suspended Particulate Matter (SPM,  $mg L^{-1}$ ) was determined by weighing before and after filtration through GF/F filters ( $0.7\mu m$ ) (for more details: <http://somlit.epoc.u-bordeaux1.fr/fr/>). Particulate Organic Carbon and Nitrogen (POC and PON,  $\mu g L^{-1}$ ) were estimated using a NA2100 Frisons Analyzer. Chlorophyll a (Chl. a,  $\mu g L^{-1}$ ) samples were extracted for 24h in 90% acetone following the fluorimetric protocol of Aminot and Kerouel (2004). The concentrations were measured using a 10-AU Turner Designs® fluorometer and the Lorenzen equations (1967).

### **DNA and RNA extraction**

Total DNA and RNA of planktonic microorganisms were extracted and purified simultaneously from the same filter with the Qiagen AllPrep DNA/RNA Mini kit (Qiagen, Hilden, Germany) following the manufacturer's protocol. To remove contaminating DNA from RNA and inversely, all RNA samples were treated with DNase and all DNA samples were treated with RNase. Each resulting extraction ( $100\mu L$ ) had DNA concentration between 0.26 and 28  $ng \mu L^{-1}$  and RNA concentration between 1.14 and 14.8  $ng \mu L^{-1}$ , as measured with the Qubit® 2.0 fluorometer (Thermo Fisher Scientific, Waltham, MA, USA).

### **PCR and tag Illumina sequencing**

Extracted RNA was first denatured in presence of two primers, 18S-82F (5'-GAAACTGCGAATGGCTC-3', [49]) and Euk-516r (5'-ACCAGACTTGCCCTCC-3', [50]), that were designed to amplify the variable V2-V3 eukaryotic 18S rRNA gene regions (around 470-480 bp), at 70°C for 5 min and 25°C for 10 min. Then, it was reverse transcribed into cDNA using M-MLV reverse transcriptase RNase H Minus DNA

Polymerase (Euromedex, Souffelweyersheim, France) at 45°C for 60 min then held at 10°C. Afterwards, cDNA samples were treated with RNase A (Euromedex, Souffelweyersheim, France) at 37°C for 10 min and were purified by using NucleoFast 96 PCR Clean-up Kit (Macherey Nagel, Düren, Germany). For amplicon library constructions, DNA and cDNA samples were first PCR amplified with primers Euk-82F and Euk-516R, and resulting products were subsequently attached to sequencing adapters-index through the second PCR step (Pegase platform, GenesDiffusion, Lille, France). The Platinum *Taq* High-Fidelity DNA Polymerase (Thermo Fisher Scientific, Waltham, MA, USA) was used for PCR amplifications with program set as: 94°C for 1 min, [15 s at 94°C, 15s at 51°C, 45 s at 68°C](26 cycles for 1st PCR, 9 cycles for 2nd PCR), and 1 min at 68°C. Products of second PCR were purified with the Agencourt AMPure XP system (Beckman Coulter, Brea, CA, USA), quantified with the PicoGreen Assay (Thermo Fisher Scientific, Waltham, MA, USA), and examined for quality with Bioanalyzer High sensitivity DNA Analysis Kits (Agilent, Santa Clara, CA, USA). Amplicon libraries were sequenced with the Illumina MiSeq platform 2x300bp paired-end run at the Plateforme de Génomique LIGAN-PM (Université de Lille 2, CNRS-UMR8199, Lille, France).

### **Sequences processing**

The rDNA and rRNA sequences obtained from Illumina were all processed together using the MOTHUR v1.34.0 software [51] following the standard operating procedure ([http://www.mothur.org/wiki/MiSeq\\_SOP](http://www.mothur.org/wiki/MiSeq_SOP)) [52]. Sequences were extracted and separated according to their index tag. The dataset was dereplicated to unique sequences and aligned against the SILVA 108 database (<http://www.arb-silva.de/>). Suspected chimeras were removed by using the UCHIME software [53]. After quality filtering, an average of 19,878 rDNA reads and 10,091 rRNA reads per sample were clustered into operational taxonomical units (OTUs) at 97% similarity threshold [5], using the average neighbor method in Mothur. Single singleton, referring to OTU that has a single representative sequence in the whole data set, were removed as these are most likely erroneous sequencing products [5,54]. After normalization of the entire dataset, all remaining OTUs (933 OTUs) sequences were searched against the PR2 curated database [55] and SILVA 114 database [56] by using BLASTN [57]. Careful examination and manual curation of BLASTN results was done to assign putative taxonomic affiliations for each OTU. Those identified as metazoa were removed from analyses to only target microbial eukaryotes.

## Data analyses

In order to establish an overview of microbial diversity over four years, we combined data from February 2013 to July 2015 (this study, GenBank-SRA accession SRP136006) with data previously obtained by our group from March 2011 to July 2013 (GenBank-SRA accession SRX768577, [45]). To note, the same set of primers (i.e. Euk-82F and Euk-516R) were used in both surveys for the amplification of 18S ribosomal gene. The two datasets were analyzed independently, and obtained results were compared for OTUs richness and succession.

Alpha diversity estimators (the richness estimator Chao-1, Simpson and Equitability indices) were calculated using the Past 3.05 software [58] for all samples. The Chao-1 estimator uses the number of singletons and doubletons to estimate the theoretically number of expected OTUs based on the singletons and doubletons in each sample. Simpson's index (D) measures the probability that two individuals belong to the same OTU, it ranges from 0 (all taxa are equally present) to 1 (one taxon dominates in the community). The Shannon equitability index measuring the evenness of the community was calculated by dividing the Shannon diversity index (H') by the maximum diversity (H'max). These diversity estimators were compared for each date with a 'bootstrapping randomization procedure' and the *p*-value was computed based on 1,000 random permutation pairs.

Principal Component Analysis (PCA) was performed on environmental variables with R (v 3.2.3) [59] using the 'ade4' package [60]. Before PCA analysis, the BoxCox transformation [61] was applied using the 'caret' package [62] to approach normal distribution in the data. Sampling dates were grouped using the hierarchical Ward's method [63].

Microbial assemblages (based on OTUs) were grouped across sampling dates by hierarchical cluster analysis using the PRIMER version 6.0 [64]. The Bray-Curtis dissimilarity coefficients were calculated to build a matrix based on OTUs abundance. Similarity profile test (SIMPROF) was performed to define significant seasonal clusters [64].

The relative read abundance of each OTU was calculated at each date by dividing its number of reads (rRNA or rDNA) by the total reads number (rRNA or rDNA) for that

OTU. For each OTU, the relative abundance of rRNA (y-axis) and rDNA (x-axis) reads per date were plotted on logarithmic scales and linear regression was calculated. This regression, used to estimate the relation between OTU abundance (rDNA) and relative activity (rRNA), was calculated using the Past 3.05 software [58], with RMA (Reduced Major Axis) to minimize both the x and y errors. For each OTU and date, rRNA:rDNA ratio was calculated by dividing the relative abundance of rRNA reads for an OTU by the relative abundance of rDNA reads of the same OTU (2). OTUs occurring only in the DNA or only in the RNA dataset were excluded for calculation of rRNA:rDNA ratio and for the linear regression analysis.

$$rRNA:rDNA\ ratio_{(OTU_i)} = \frac{(NbRNAreads_{OTU_i} / \sum_{i=1}^R NbRNAreads_{OTU_i})}{(NbDNAreads_{OTU_i} / \sum_{i=1}^R NbDNAreads_{OTU_i})} \quad (2)$$

## RESULTS

### Environmental variables and biological seasonality

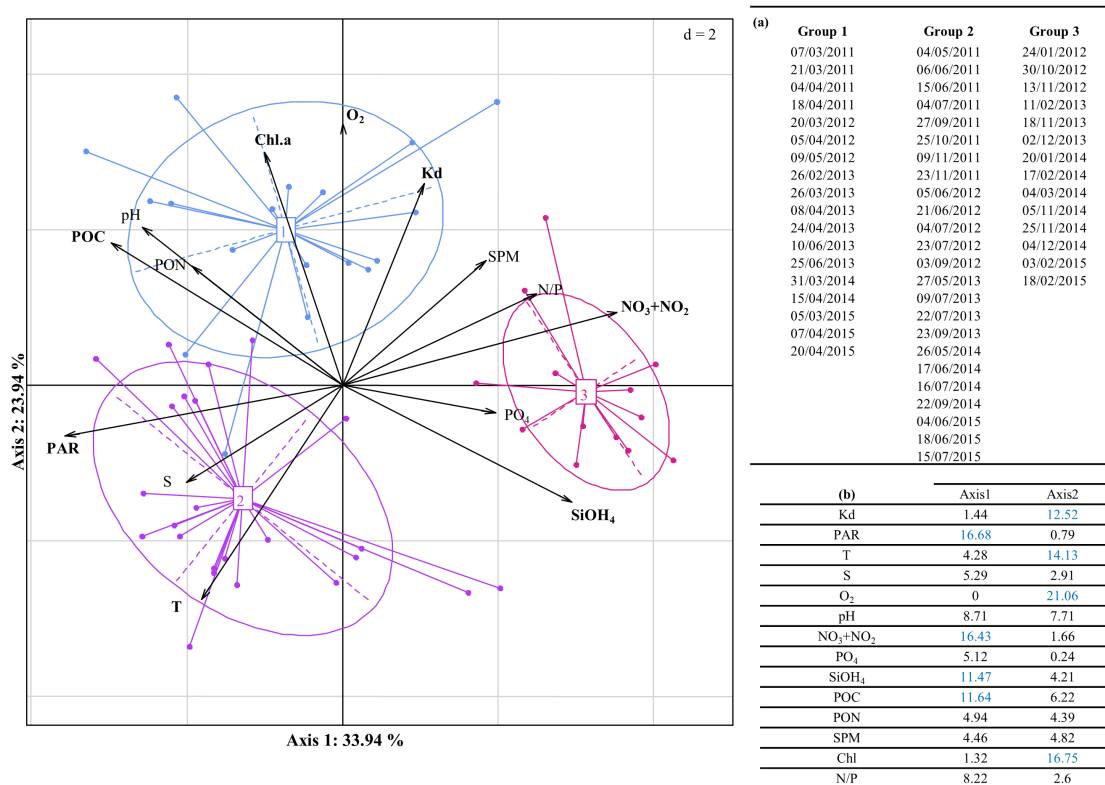
During the study period, seawater temperature (SST) and salinity (S) ranged from 5.4 to 18.5°C and from 33.11 to 34.07, respectively (Table 1). Each year, the largest amounts of nutrients (26.7 µM for NO<sub>3</sub><sup>-</sup>+NO<sub>2</sub><sup>-</sup>, 1.2 µM for PO<sub>4</sub><sup>3-</sup>, and 13.6 µM for SiOH<sub>4</sub>) were observed between November and March (Table 1). The N/P ratio dropped during *P. globosa* blooms (as low as 0.19 mol:mol in April 2012, S1 Table). Highest O<sub>2</sub> concentrations were recorded every year during the phytoplankton spring bloom (up to 8.2 ml.L<sup>-1</sup> in April of most years) and then decreased from May to September (down to 4.9 ml.L<sup>-1</sup> in September 2014) (S1 Table). Chl. a amounts varied from 0.4 to 14.2 µg.L<sup>-1</sup> with recurrent maximum values between March and June. The highest values of SPM were recorded in winter every year, except for June 2015 where it reached its maximum (32.9 mg.L<sup>-1</sup>) (S1 Table).

**Table 1. List of environmental variables, means ( $\pm$ SD) of all measurements, minimal and maximal value measured per variable at the SOMLIT station (mean $\pm$ sd, min-max).**

	mean $\pm$ sd	min-max
Kd ( $\text{m}^{-1}$ )	0.37 $\pm$ 0.1	0.17-0.73
PAR ( $\text{E m}^{-2} \text{d}^{-1}$ )	22.35 $\pm$ 13.5	2.28-51.95
T ( $^{\circ}\text{C}$ )	12.07 $\pm$ 3.9	5.44-18.5
S	34.11 $\pm$ 0.3	33.11-34.7
O <sub>2</sub> ( $\text{mL.L}^{-1}$ )	6.42 $\pm$ 0.9	4.9-8.2
pH	8.09 $\pm$ 1.1	8.01-8.49
NO <sub>3</sub> <sup>-</sup> +NO <sub>2</sub> <sup>-</sup> ( $\mu\text{M}$ )	5.89 $\pm$ 7.3	0.06-26.71
PO <sub>4</sub> <sup>3-</sup> ( $\mu\text{M}$ )	0.3 $\pm$ 0.2	0.01-1.2
SiOH <sub>4</sub> ( $\mu\text{M}$ )	2.57 $\pm$ 3.4	0.01-13.6
POC ( $\mu\text{gC.L}^{-1}$ )	367.03 $\pm$ 283.5	64.78-1242.8
PON ( $\mu\text{gN.L}^{-1}$ )	55.11 $\pm$ 29.68	5.33-180.53
SPM ( $\text{mg.L}^{-1}$ )	5.18 $\pm$ 6.1	0.1-32.93
Chl. a ( $\mu\text{g.L}^{-1}$ )	4.03 $\pm$ 3.3	0.4-14.19
N/P	29.9 $\pm$ 57.8	0.12-373.5

Kd: diffuse attenuation coefficient, PAR: photosynthetically active radiation, T: temperature, S: salinity, O<sub>2</sub>: Oxygen, pH: potential hydrogen, NO<sub>3</sub><sup>-</sup>: nitrate, NO<sub>2</sub><sup>-</sup>: nitrite, PO<sub>4</sub><sup>3-</sup>: phosphate, SiOH<sub>4</sub>: silicate, POC: particulate organic carbon, PON: particulate organic nitrogen, SPM: suspended particular matter, Chl. a: chlorophyll a, and N/P: nitrate/phosphate ratio.

Principal Component Analysis (PCA) applied to environmental variables showed that the first two components (PCA1 and PCA2) represented 58% of the total variability in the data. All environmental variables except for salinity, pH and SPM, contributed >10% to the formation of the two first axis of the PCA (Figs 2a and 2b). PCA1 strongly correlated to, by decreasing order, PAR, NO<sub>3</sub><sup>-</sup>+NO<sub>2</sub><sup>-</sup>, SiOH<sub>4</sub> and POC, while PCA2 significantly correlated to O<sub>2</sub>, Chl. a, SST, and Kd. The environmental variables PON and PO<sub>4</sub><sup>3-</sup> (14% and 32% of total contribution, respectively) contributed mainly to the third PCA (not shown). The Hierarchical Cluster Analysis (HCA) with Ward's method, allowed to distinguish three seasonal periods (Fig 2). The first period corresponded to *P. globosa* spring blooms (from March to June), and was associated to high values of Kd, O<sub>2</sub>, Chl. a, and POC, that are characteristics of spring conditions. The second period (from early May to early November) was associated to warm and nutrient-limited waters. Finally, the third period (from mid-November to early March) was associated to low SST and high nutrient levels, typical of winter conditions (Figs 2 and 2b).

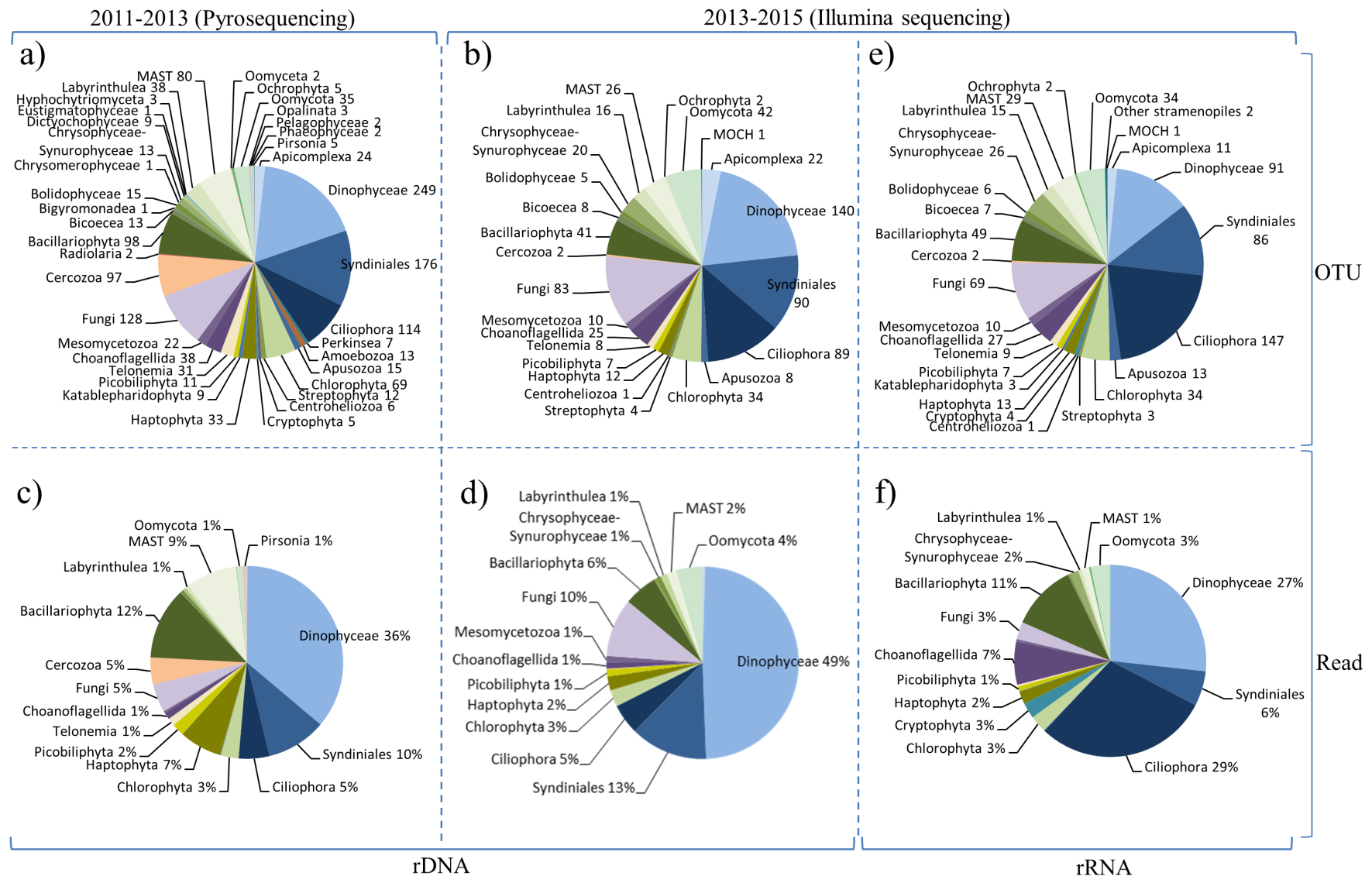


**Fig 2. Principal Component Analysis (PCA) applied to the environmental variables recorded at SOMLIT station.** Projection of the environmental variables (arrows) and the sampling dates (colored points) on the first factorial plane explaining 58% of the total data inertia. The blue, red and purple ellipses (numbered 1, 2, 3, respectively) correspond to the three groups of sampling dates revealed by the HCA with Ward' method of the PCA dates coordinates on the factorial plane. Most contributing environmental variables to the variability on the first two axes are marked in bold. **(a)** Sampling dates grouping on each cluster **(b)** Contributions of each environmental variable on the first two axes of the PCA in percentage. In blue: strong contributions of environmental variables (>10%).

### Eukaryotic diversity (rDNA-based)

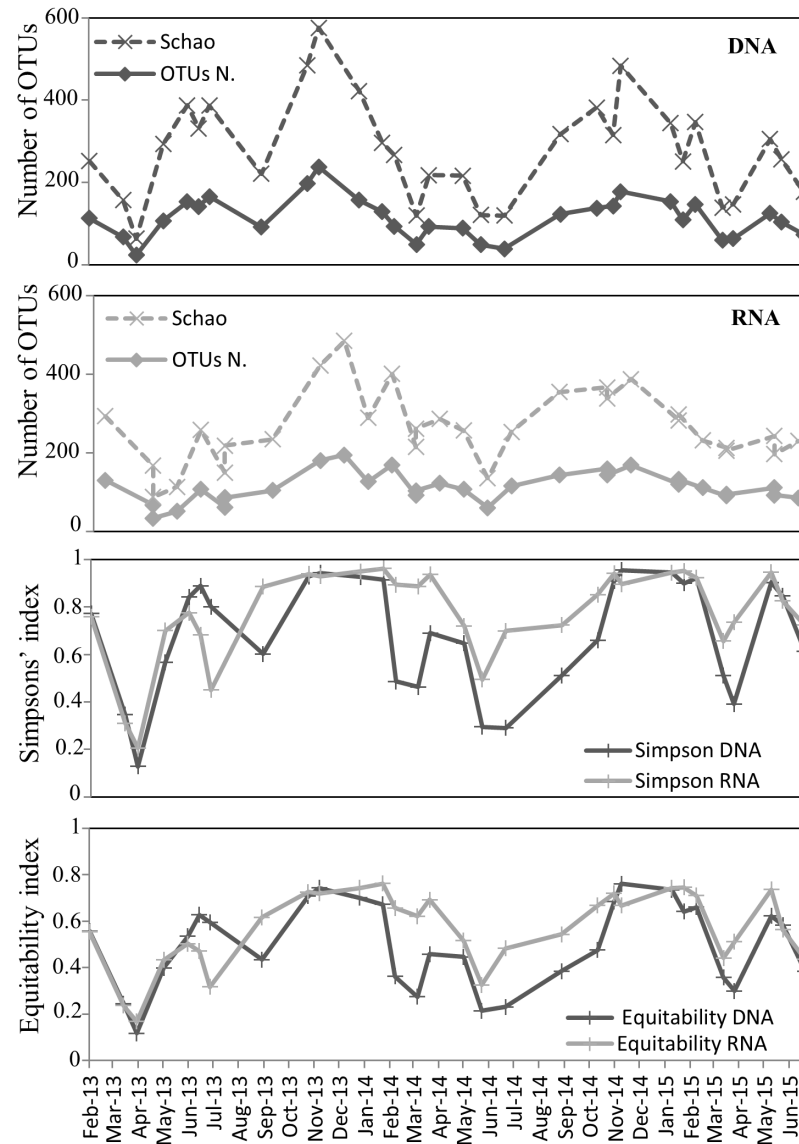
The DNA-based diversity was conducted during a four-year survey (from 2011 to 2015). During this period, the 454-pyrosequencing method (used during the first survey from 2011 to 2013) was gradually discontinued at the benefit of Illumina sequencing (used during the second survey from 2013 to 2015) due to fast and continuous development of high throughput sequencing technologies. The rDNA-based analysis identified 1388 OTUs with the pyrosequencing approach (47 samples, 2011-2013) and 696 OTUs with Illumina

sequencing (30 samples, 2013-2015). Overall, the OTUs identified were affiliated to eight super-groups (Alveolata, Amoebozoa Apusozoa, Archeplastida, Hacrobia, Opisthokonta, Rhizaria and Stramenopiles) with Amoebozoa only identified during the 2011-2013 sampling period (13 OTUs) (Fig 3a, Table 2). The two datasets were analyzed separately. For both periods and sequencing methods, no significant difference was observed on the OTU abundance and taxonomic distribution of dominant higher taxonomic groups (representing at least 1% of the total OTUs) based on the Kolmogorov-Smirnov test ( $p$ -value=0.22). However, a significant difference (Kolmogorov-Smirnov  $p$ -value < 0.005) was observed between the two sequencing methods and sampling periods when including rare groups. For example, the Radiolaria were only observed in the pyrosequencing dataset (Figs 3a and 3b, Table 2). Overall, among the 40 groups identified (Table 2), 24 were found in both surveys, 15 were found only in the pyrosequencing dataset, and one group (MOCH) was identified only in the Illumina dataset (in May, September, November 2013 and in February 2014) (Fig 3b, Table 2). The highest number of OTUs was consistently found in November-December (with 373 OTUs for 2011-2013 and 238 OTUs for 2013-2015), while the lowest diversity was repeatedly found in April (with 46 OTUs in April 2012 and 24 in April 2013, for the pyrosequence and Illumina datasets, respectively) (Fig 4). Alveolata was always the most diverse group, comprising 570 OTUs (in 2011-2013) and 341 OTUs (in 2013-2015). This group was mostly composed of Dinophyceae (~20% of all the OTUs), followed by Syndiniales (~13%) and Ciliophora (~11%). Stramenopiles (327 OTUs in 2011-2013 and 161 OTUs in 2013-2015) was also a highly diverse group composed of Bacillariophyta (~6% of all the OTUs), MAST (~4%) and Oomycota (~6%). Finally, Fungi represented around 10% of all OTUs (Figs 3a and 3b). Although the number of reads can be only considered in terms of relative and not absolute abundance, it can be noted that Dinophyceae showed the highest reads number (36% and 49% of the total number of reads, in the 2011-2013 and 2013-2015 surveys, respectively), followed by Syndiniales (10% and 13%), Bacillariophyta (12% and 6%), Fungi (5% and 10%) and Ciliophora (5% and 5%) (Figs 3c and 3d). Overall, significant differences were found in the distribution of taxonomic groups read's abundances (Kolmogorov-Smirnov  $p$ -value < 0.001) between the two surveys. For example, pyrosequencing generated a higher proportion of reads for MAST (9%) and Haptophyta (7%) compared to Illumina (Figs 3c and 3d).



**Fig 3. Overview of microbial community structure in the EEC, relative to the studied periods and sequencing methods.** Number of OTUs unveiled by pyrosequencing (a) and Illumina (b) methods, and relative number of reads obtained by pyro-sequencing (c) and Illumina (d) methods, based on rDNA sequencing. Number of OTUs (e) and relative number of reads (f) based on rRNA Illumina sequencing. To facilitate reading, the percentage of reads number assigned to a specific group is given when  $\geq 1\%$ . Taxonomic affiliation was based on BLASTN searches against the PR2 and Silva databases.





**Fig 4. Diversity indices, based on rDNA and rRNA Illumina sequencing, for samples taken in the EEC (2013-2015).** Curves represent the number of OTUs (solid lines), the richness estimator ( $S_{\text{chao}1}$ , dash lines) and indices for diversity heterogeneity (Simpson, Equitability). The rDNA-based dataset is depicted with dark grey lines and the rRNA-based dataset with light grey lines.

**Table 2: Taxonomic groups unveiled in the EEC from 77 samples either by pyrosequencing (2011-2013, 47 samples) or by Illumina sequencing (2013-2015, 30 samples).**

Groups	2011-2013	2013-2015
	Pyrosequencing	Illumina
<b>Alveolata</b>	✓	✓
Apicomplexa	✓	✓
Dinophyceae	✓	✓
Syndiniales (MALV)	✓	✓
Ciliophora	✓	✓
<u>Perkinsea</u>	✓	✓
<b>Amoebozoa</b>	✓	ND
<b>Apusozoa</b>	✓	✓
<b>Archaeplastida</b>	✓	✓
Chlorophyta	✓	✓
Streptophyta	✓	✓
<b>Hacrobia</b>	✓	✓
Centrohelioza	✓	✓
<u>Cryptophyta</u>	✓	ND
Haptophyta	✓	✓
<u>Katablepharidophyta</u>	✓	ND
Picobiliphyta	✓	✓
Telonemia	✓	✓
<b>Opisthokonta</b>	✓	✓
Choanoflagellida	✓	✓
Mesomycetozoa	✓	✓
Fungi	✓	✓
<b>Rhizaria</b>	✓	✓
Cercozoa	✓	✓
<u>Radiolaria</u>	✓	ND
<b>Stramenopiles</b>	✓	✓
Bacillariophyta	✓	✓
Bicoecia	✓	✓
<u>Bigyromonadea</u>	✓	ND
Bolidophyceae-and-relatives	✓	✓
<u>Chrysomerophyceae</u>	✓	ND
Chrysophyceae-Synurophyceae	✓	✓
<u>Dictyochophyceae</u>	✓	ND
<u>Eustigmatophyceae</u>	✓	ND
<u>Hyphochytriomyceta</u>	✓	ND
Labyrinthulea	✓	✓
MAST	✓	✓
<i>MOCH</i>	ND	✓
<u>Oomyceta</u>	✓	ND
Ochrophyta	✓	✓
Oomycota	✓	✓
<u>Opalinata</u>	✓	ND
<u>Pelagophyceae</u>	✓	ND
<u>Phaeophyceae</u>	✓	ND
<u>Pirsonia</u>	✓	ND
<u>Other Stramenopiles</u>	✓	ND

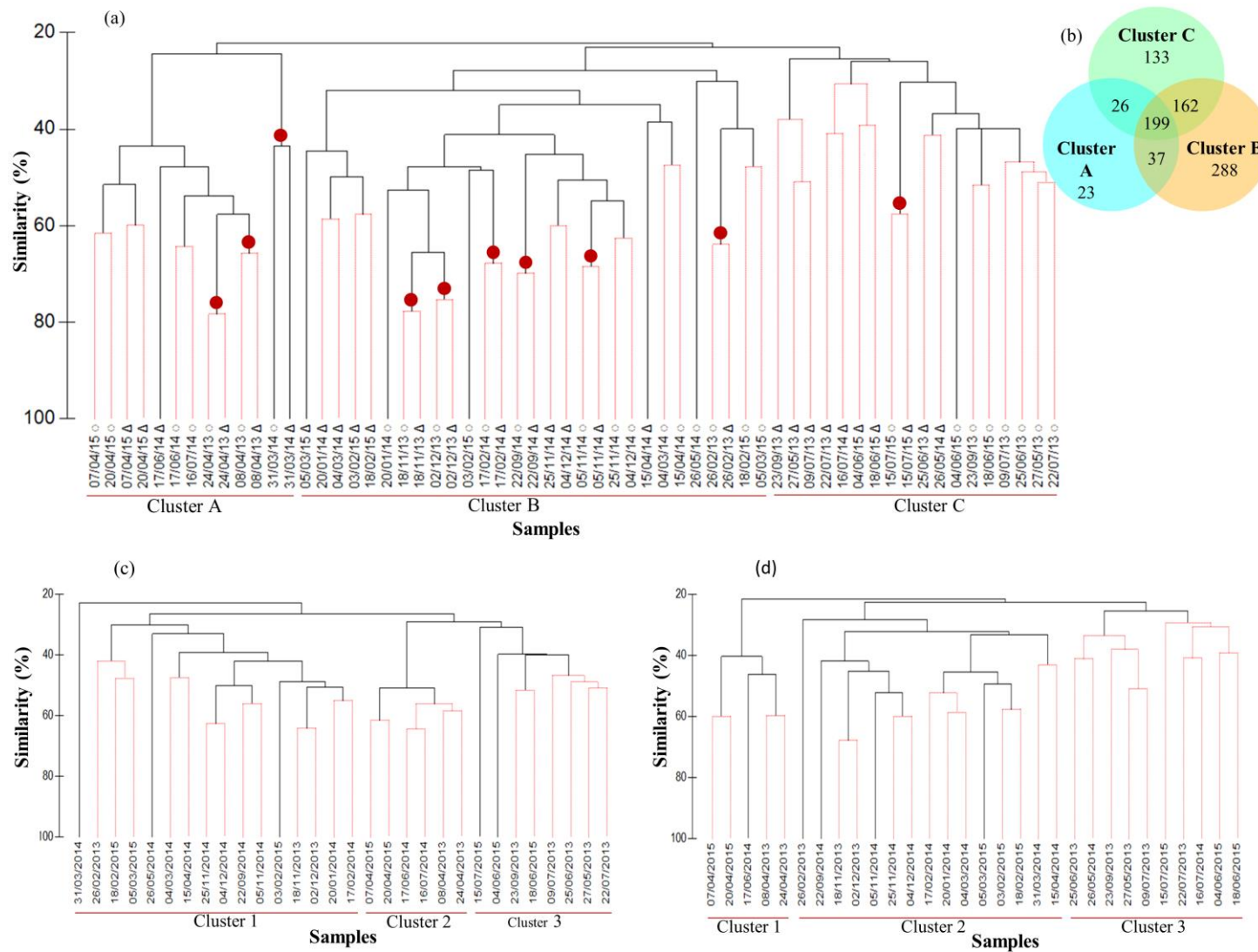
Groups identified only between 2011 and 2013 (pyro-samples) are underlined. The MOCH group (in italic) was only identified between 2013 and 2015 by Illumina sequencing. ND: Not Detected.

## Microbial community composition and seasonal succession revealed by rDNA and rRNA sequencing

Between 2013 and 2015, a total of 696 OTUs and 700 OTUs were detected by rDNA- and rRNA-based Illumina sequencing, respectively (Figs 3b and 3e). The rDNA- and rRNA-based diversity followed a similar pattern of OTU abundance (Fig 4). The mean ratio of observed to expected OTUs (Chao-1) was  $69\pm 8\%$  for DNA samples and  $78\pm 7\%$  for RNA samples. The Chao-1, Simpson and Equitability indices reflected the same pattern, with relatively higher values from October to February, and lower values during *P. globosa* bloom (from mid-February to May) (Fig 4). Over 60 % of total OTUs were shared between the rDNA and rRNA datasets, whereas 168 OTUs were only identified in the rDNA dataset and 172 OTUs were only identified in the rRNA dataset. A total of 868 OTUs corresponding to 26 higher taxonomic groups were found when considering both rDNA and rRNA datasets (Figs 3b and 3e). The proportion of OTUs affiliated to these major super groups was similar whether considering the rDNA or rRNA dataset (~49% of Alveolata, ~23% of Stramenopiles, ~15% of Opisthokonta, 4~5% of Archaeplastida and Hacrobia, and 1~2% of Apusozoa) (Figs 3b and 3e). In the rRNA dataset, the most diverse group was the Ciliophora (147 OTUs gathering together 29% of the total reads), with the majority being affiliated to the Spirotrichea (60 OTUs) (Figs 3e and 3f). The second most diverse group was the Dinophyceae (91 OTUs, 27% of all reads) followed by Syndiniales (86 OTUs, 6% of all reads), Fungi (69 OTUs, 3% of all reads) and Bacillariophyta (49 OTUs, 11% of all reads) (Figs 3e and 3f). No significant difference was observed between the rDNA and rRNA datasets regarding OTU abundance and taxonomic distribution for higher taxonomic groups (Kolmogorov-Smirnov p-value=0.98).

The cluster analysis applied together on the total read abundance of OTUs from rDNA and rRNA samples showed three main clusters at a similarity level of 25% (Fig 5b). The first cluster (Cluster A) grouped mainly samples from March to July (including the bloom period of *P. globosa*). The second cluster (cluster B) grouped autumn and winter samples (September to March). Finally, the third cluster (cluster C) included summer and early autumn samples (late May to early September). Protist community structure based on rDNA and rRNA sequencing exhibited minor differences. In total, 10 of the 30 DNA/RNA pairs (i.e. rDNA and rRNA from the same date) appeared side by side in the dendrogram (red dots in Fig 5a), and 17 of the 30 pairs belonged to the same cluster at 40 % similarity.

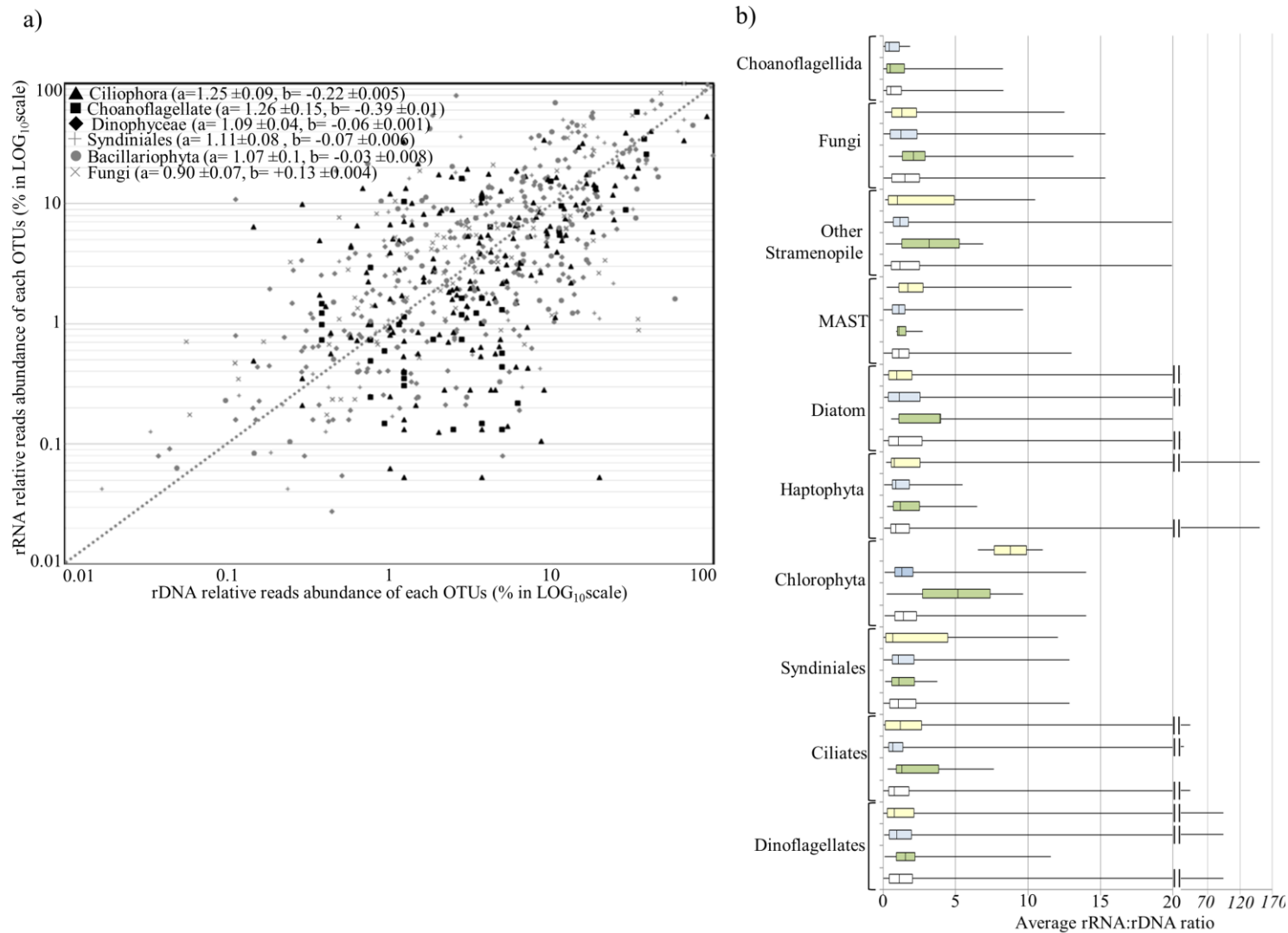
The remaining pairs of samples showed lower similarity, but they still belonged to the same seasonal cluster, except for two samples (July 16<sup>th</sup>, 2014 and May 26<sup>th</sup>, 2014) (Fig 5a). The Venn diagram showed that only 23% of OTUs were shared among the three clusters (Fig 5b). While fall-winter (cluster B) and summer (cluster C) periods had relatively distinct diversity (~43% of shared OTUs), the blooming period of *P. globosa* (cluster A) shared most of its diversity (>90 % of OTUs) with the other two clusters (Fig 5b). When the same cluster analysis was applied to total OTU read abundance from rDNA (Fig 5c) and rRNA (Fig 5d) separately, the clusters inferred from rRNA data showed a clearer seasonal representation of the community succession than the one based on rDNA. In fact, 79% of the rRNA samples (compared to 72% for the rDNA samples) belonged to a cluster in agreement with those that were identified in the PCA of environmental variables (Figs 2 and 5c-d).



**Fig 5. Cluster-based microbial community structure in the EEC.** (a) Cluster diagram based on Bray–Curtis dissimilarities calculated based on the non-transformed number of reads for all OTUs found during the study. Red lines in the dendrogram indicate significant differences ( $p > 0.05$ ) between bifurcations, based on the SIMPROF significance test.  $\circ$  DNA samples,  $\Delta$  RNA samples. Red dots indicate grouping of samples from the same sampling date (b) Venn diagram of shared diversity between three clusters. The total richness for all groups was 868 OTUs. The number of species shared between the three clusters was 199, corresponding to 23% of the total richness. Similar Bray-Curtis cluster diagrams were constructed from rDNA-based (c) and rRNA-based (d) number of reads.

### **Temporal variation of abundance and activity of most abundant OTUs**

The rRNA:rDNA ratio was used as a proxy of relative cellular activity for the 100 most abundant OTUs (>0.1% of the total reads abundance) which represented over 90% of all reads. Despite of important variability, the relative abundance of each OTU (rDNA-based dataset) generally increased with cellular relative activity (rRNA-based dataset) (Fig 6a).



**Fig 6. Relation between abundance (rDNA reads) and relative cell activity (rRNA reads) for the 100 most abundant OTUs.** (a) Each OTUs was plotted according to their relative number of rDNA (x-axis) and rRNA (y-axis) reads. Each point represents paired percentages (in logarithmic scales) of each taxonomic group from each sample. (b) Boxplot representing rRNA:rDNA ratios for the major taxonomic groups of the 100 most abundant OTUs. The boxplots represent the variation of the rRNA:rDNA ratio around the median for each seasonal clusters (white: all years, green: cluster A representative of *P. globosa* bloom period, blue : cluster B corresponding to winter conditions, yellow: cluster C corresponding to spring-summer conditions). Whiskers indicate the maximum and minimum values.

**Dinoflagellates (mostly heterotrophic micrograzers and mixotrophs):** this group represented almost half (48%) of the 100 most abundant OTUs' sequences. The large number of rDNA reads for this group is related to high rDNA gene copy number present in the genome of Dinoflagellates [65,66]. The most abundant OTU during the entire survey was affiliated to *Gyrodinium spirale* (OTU1: 35% of all sequences) and was found in all samples. Three OTUs (OTU1, OTU2 and OTU10), having *G. spirale* as their closest relative, all reached recurrent highest abundance in April-June (Table 3). *G. fusiforme* (OTU21) followed the same seasonal succession as *G. spirale*. The second most abundant dinoflagellate was *Torodinium robustum* (OTU61), mostly present from September to February (Table 3). The rRNA:rDNA ratio (as a proxy of cells activity) for dinoflagellates was generally low, but showed some increase in spring compared to the rest of the year (Fig 6b). In general, the number of rRNA reads appeared to be lower than rDNA reads, suggesting that dinoflagellates were abundant in the EEC but with a relatively low cell activity (Fig 6b). *Warnowia sp.* (OTU4) and *Katodinium glaucum* (OTU8) were present all year long. *Warnowia sp.* displayed its highest abundance and higher rRNA:rDNA ratio from September to February, while peak abundance of *Katodinium glaucum* was detected in June/July. Other unclassified Dinophyceae (8 OTUs) were detected throughout the year with seasonal preferences depending on the OTU. The most abundant unclassified Dinophyceae sp. (OTU3) was observed from March to July with a peak of activity in March or June. Another Dinophyceae sp. (OTU7) was found all year with a peak of activity in September, and an increased abundance in November. Several other Dinophyceae sp. (OTU24, 126, 131 and 245) were also mostly detected between November and February (Table 3).



**Table 3. Heat-map showing abundance (rDNA-based survey) and relative activity (rRNA/rDNA ratio) of the 100 most abundant OTUs (representing more than 92% of all the reads) detected between 2013 and 2015. Positive versus negative ratios can be used as an indication of relative cell activity. Negative rRNA:rDNA ratio indicates low cell activity, in contrast to positive ratio which corresponds to higher cells activity (for example, a ratio of -40 corresponds to an OTU having 40-time more rDNA reads than rRNA reads).**

OTU	Affiliation	2013										2014										2015										
		26 feb.	8 Apr.	24 Apr.	27 May	25 Jun.	9 Jul.	22 Jul.	23 Sep.	18 Nov.	2 Dec.	20 Jan.	17 Feb.	4 Mar.	31 Mar.	15 Apr.	26 May	17 Jun.	16 Jul.	22 Sep.	5 Nov.	25 Nov.	4 Dec.	3 Feb.	18 Feb.	5 Mar.	7 Apr.	20 Apr.	4 Jun.	18 Jun.	15 Jul.	
<b>Dinoflagellates</b>																																
1	<i>Gyrodinium spirale</i>	1.4	1.9	1.7	-33	1.4	-8.9	-2.5	10.8	1.4	2.0	-9.4	-2.8	-3.5	1.2	-4.3	2.4	1.5	-3.1	1.3	1.2	-3.3	-2.4	-3.6	-16	-3.0	1.5	1.2	-4.2	-27	2.0	
2	<i>Gyrodinium spirale</i>		1.9	2.2	-	2.3	-	-3.1		3.4	1.7	-		-		1.1		2.0	-	1.8	1.7	-	-3.6	-		-	-2.1	1.2	-1.4	-		
3	Dinophyceae sp.	-			-	1.1	-	-1.3	-								10.5	2.2	-		-			-	-	-	-	1.8	-2.8	-4.6	3.4	
4	<i>Warnowia</i> sp.	-1.1	-2.1	-	-	-	-3.6	-1.5	6.2	1.9	3.0	-8.0	1.2	-1.1		-4.6	3.3	-1.8	1.4	2.4	2.5	-1.6	-2.3	-1.8	-2.9	1.9		1.4	-	-	7.0	
7	Dinophyceae sp.	1.8	1.8		1.4	-	-	-2.2	3.8	1.3	1.9	-14	-1.1	-5.9	11	-2.0	-2.7	1.4		2.6	1.3	-1.8	-2.3	-1.3		-	1.4	2.4	-	-1.8	2.0	
8	<i>Katodinium glaucum</i>	2.3	2.1	2.1	-63	1.2	-10	-6.2	-8.0	2.0	1.7	-		-	-	94	-1.0	-2.3	-1.1	1.4	-3.8	1.1						-	-6.2	-	2.7	
10	<i>Gyrodinium spirale</i>		1.7	3.3	-	-2.9	-	-			-						1.8	-6.6									-1.2	-1.1	-2.2			
21	<i>Gyrodinium fusiforme</i>		-1.6	3.1	-	-1.1	-	-1.5			-					-	-1.1	-								-	-3.3	-1.8	-			
24	Dinophyceae sp.	-1.6	3.7			-	-		1.2	1.5	2.4	-5.3	2.6	-5.7		-4.9	-6.5			2.9	2.0	-2.0	-2.5	-1.6	-2.5	-	2.0	1.4	-	-		
32	<i>Pelagodinium beii</i>	1.7	3.8			-		-	-			-	1.3			-	-4.0	-2.0	-			-	-	1.3	-				-	-	6.3	
51	<i>Gyrodinium</i> sp.	-2.2	-10	-	-	-	-	-5.4	30	-2.5	-1.3	-	-1.8	-		-3.1	1.2	-		2.7	-2.1	-1.8	-1.8	-14	-			-	-	-1.8		
60	<i>Heterocapsa pygmaea</i>	3.0	-		-		-			1.8	5.6	-4.2	3.5	1.1		1.4	-		-	3.1	-1.6	2.8		-1.4	-		2.8			1.1		
61	<i>Torodinium robustum</i>					-1.1	-	-	3.8	-1.2	2.9	-	-2.7	-1.1		-				2.8	2.5	-11	-2.9	-					-			
70	Dinophyceae sp.		6.2		-	1.3	-	-2.3		-							28.7			-										-		
76	<i>Gyrodinium</i> sp.	-			-1.5		-		39	1.1	-1.4	-12	-1.2	-1.7	2.6	1.1	-1.3	4.6			-		1.3	-		-	4.9	1.2				
113	<i>Chrysophyceae</i> sp.	-								-1.0	-1.6		-1.0							1.7			-11	-		20						
121	<i>Azadinium cuneatum</i>																															
126	Dinophyceae sp.		-		-		-	1.3	-3.8	2.0	2.4	-2.4	1.4	-3.4	-	-	4.9	5.8		5.8	2.4	-1.1	-2.2	2.0							-1.0	



14	<i>Chrysochromulina leadbeateri</i>	-2.8			-4.3	-		-3.1	-1.6	1.8	-1.0	-1.1	-1.1	3.2	6.5	-1.2	2.0			1.5	1.5	-1.4	-1.4	-1.7	3.8	-2.0	1.6	-	-1.9	-1.2		
16	<i>Leptocylindrus aporus</i>	-2.4		1.3	8.2	1.0	2.0	-1.6	-36								29.0											-1.8	2.3	1.6	-	
17	<i>Thalassiosira hispida</i>							-		-4.1	-11	-4.4	-4.1	2.3						-6.2	-12		-4.9	6.5	2.3							
27	<i>Phaeocystis globosa</i>	-2.9	-2.3	-3.5	-1.4	-		-	150	-1.3	1.3	-10	1.2	-2.3	5.2	-1.9	-			1.6	2.0	-4.2	-10	-6.4	-1.2	-1.2	1.4	-1.3	3.5	-	-	
36	<i>Haptolina fragaria</i>		-	-	-	-	-	-	-4.0	3.6	1.8	-1.5	-	-			5.5			-	1.8	-1.5	-					2.7	-			
41	Chlorellales sp.	-10	-		-					1.8	-1.2	7.3		7.3					11	-	-	5.5	11	2.0	-1.2	-1.3	-					
47	<i>Ditylum brightwellii</i>							-		2.2	2.6	-3.0	3.3	1.9	1.1						3.3	-		3.7	2.3	-1.7						
54	<i>Bathycoccus prasinus</i>	-1.2	-		-	-	-	-	-	-1.0	1.3	1.1	-1.4	1.9		4.8	-			-1.2	1.9	-1.2	-1.2	-1.2	14	12					-	
66	<i>Micromonas pusilla</i>	-			-	-	-	-	-	1.6	-1.9	1.7	1.1	5.5		1.1	-			1.1	1.4	-1.3	-1.8	-1.5	5.8						-	
86	<i>Polar-centric-Mediophyceae</i> sp.	3.9			-	-	-	-	-	2.5	1.3		-2.0								4.3	-2.0	-2.7							-	-	
88	<i>Guinardia delicatula</i>							-1.0	-3.6	2.0	2.8			-	-	-	4.0			-	8.0	-1.5	1.7			-	4.0	4.0	-	-1.9		
103	<i>Picomonas judraskeda</i>		-		-					2.0	-1.0	-	1.5	-		-1.8				3.3	1.1	-3.7	-4.3	-1.8	-	3.3	-		-			
104	Picobiliphyta sp.	-2.6	-1.3		-	-	-	-2.1	2.8	-1.1	1.8	-7.8	-2.1	-		-2.4	50			2.1		1.0	-2.3	-	-	-		3.6	-1.3	-		
110	<i>Pycnococcus provasolii</i>	1.4	1.6			-	-	-	-	1.6	2.2	-	2.5	-	-	2.5	-		6.5	-1.1	1.6	-	6.5	1.6	-4.0	-		-	-	-		
132	<i>Guinardia flaccida</i>			3.0	-3.7	1.1	-5.5	-1.7								-										1.3	-	1.0	-2.8	-7.4	-3.7	
168	<i>Bolidophyceae</i> sp.				-					-1.2	-3.3	-1.2	-1.3								-1.8	1.7	-	4.3	1.1	1.5	1.4					
213	<i>Micromonas pusilla</i>				-			-		1.2	-1.7	2.5	1.7	1.4						-2.8	-3.0	-1.8	1.3	1.3								
221	<i>Eucampia</i> sp.				-			1.0	-	-7.3	-14				6.5						-1.6	-1.8	-4.9					3.1	-1.7			
266	<i>Guinardia delicatula</i>	-1.8		-1.8	-	-	-2.6	6.7							3.9	-1.1												-				
370	Picomonas sp.	-			-					2.9	-2.5									1.8	3.7	1.2	-5.3									
581	<i>Guinardia flaccida</i>	-				-	1.9	-				-		23	11		-										-1.4	-	-1.1			
1085	Picobiliphyta sp.				-	-	-2.4	-	4.4	1.5							22.0	-		4.4												-

**MALV, Fungi and other potential parasites**

9	<i>Olpidiopsis porphyrae</i>					-	-	-	-1.0	3.4	-			2.7	1.3					-1.1	-1.1	-3.9	-5.2	-								1.0
11	Dino-Group-I-Clade-1 sp.	2.6	1.2	2.5	-5.7	-3.2	5.2	-	6.7	2.3	2.8	-23	-1.2	-2.7	3.7	-1.0	-18	1.0	-6.3	2.1	1.8	-1.4	-1.4	-3.0	20.3	-6.0	1.1	-2.2	-8.4	-	-2.2	
15	<i>Haliphthoros</i> sp.	3.7	-5.2	6.9	-1.7	-1.7	-	-		2.8	2.3	1.1	-1.5	-		-			6.9				-	-7.6	-14	-	3.4	-	1.3	-	-	
19	<i>Xenomeris raetica</i>	4.8	-		-	2.4	-	-	-	1.2	4.8	9.6				-	-			-			2.4	-1.7	-	-7.5	-1.5		-	2.4	3.2	
20	<i>Penicillium</i> sp.	-	2.9	-	-2.9			-							-1.0				-	2.3	-		-	2.9	-	-2.4	1.2		-	-		

23	<i>Ichthyospora</i> sp.	-	-	-	1.3	-	-2.2	-	1.9	3.7	1.2	7.4	-	6.2	-	-10		1.9	-	-	-	-	3.7	-	3.4	-	-1.7	3.2					
31	<i>Cryptomycotina</i> sp.	1.5							-	-	-										-	-	-	-		-							
35	<i>Engyodontium album</i>	1.8	2.7	4.2	-2.2	1.1	-	-1.6	-	2.8	-1.0	-1.9	-1.9	-2.6	2.9	-	9.7	4.2	12.5	4.2		1.3	-1.1	-2.1	-8.6	-3.7	2.5	5.0	-2.0	-	-1.2		
42	<i>Xenobotrytis acaducospora</i>	2.1	1.6		-					-	-				1.7	-			-2.7			-	-	-1.9	-	-6.4	-1.3	1.4	-10	-	-1.3		
45	<i>Thraustochytriaceae</i> sp.				-1.0	5.0	-	-5.4	-	-2.4	-1.9						9.7	-1.0	4.9					-	-	-	-		-3.2	-1.7	1.8		
48	<i>Gregarines</i> sp.						-																										
53	<i>Meira miltonrushii</i>	15	1.5	13	-	2.2	-	-		1.6	1.9	-32			-1.1		-	1.9		2.2	-1.4	-	-	-	-2.1	2.2	2.7	-1.4	-	1.5			
58	Oomycota sp.	1.2										1.1									1.3	-	-	8.3	-1.1	-							
62	<i>Abeoformidae</i> MAIP 2 sp.		-1.3					-	-1.1	-	-4.3	2.5	-				-1.1	-	-2.3	-1.1	7.5	2.4	1.1	-3.7					-1.2	1.2			
72	<i>Thraustochytriaceae</i> sp.				-	7.4	-							1.9		-	-									-2.4			-6.5				
74	<i>Achlya bisexualis</i>	-	2.3																							1.1							
87	<i>Gregarines</i> sp.				-	-				-	3.7		-					-											-		5.8		
98	<i>Pseudoperkinsus tapetis</i>			3.0	-	-	-	-	2.0							-		5.9						-	-	1.5	3.5	8.9	-	1.5	-		
150	<i>Abeoformidae</i> MAIP 2 sp.													-1.4	-	1.1			-1.8		-								-				
155	Dino-Group-II-Clade-13																			1.6	1.6	1.5	-1.5	-			-	-3.6	-				
173	Dino-Group-II-Clade-22									7.0	7.0	-																					
181	Dino-Group-III sp. strain1					-				-	-1.5	1.5		1.2				-15			-	5.4	-	-8.8	-				-	3.7			
184	<i>Olpidiopsis</i> sp.	-	1.4																								2.7	1.2					
186	Dino-Group-I-Clade-1 sp.	1.1	-		-	-				1.4	1.1	-																					
215	Dino-Group-I-Clade-1 sp.				-	-	-	1.6		5.1	2.6							3.4		-	13								-				
220	Dino-Group-I-Clade-1 sp.	2.7	1.1		-	-	8.6						-5.1	-	-					4.2	1.1	-2.8	-3.9	-2.1	-	-			-	-	-1.1		
229	Dino-Group-II-Clade-8 sp.				-	-	1.5	-	-	2.7	-1.5						12			3.6	1.5	-2.2	-3.6										
265	<i>Chytridiomycetes</i> sp.	2.8	-									1.7	-																				
281	Dino-Group-II-Clade-10-and-11 sp.																																
294	<i>Exobasidiomycetes</i> sp.	-	2.0		-	2.0																											
352	Dino-Group-II-Clade-10-and-11 sp.	-																															
361	<i>Haliphthorales</i> sp.					-	-	-	7.1											6.4													

519	<i>Cryptomycota</i> sp.																					-	-	-															
638	Dino-Group-I-Clade-5 sp.				-	-	2.0	-1.0									1.5	2.2	-1.1	-1.8	-																		-

<b>Legend:</b>	0 %	<0% to 1%	<1% to 5%	<5% to 10%	<10% to 25%	<25% to 50%	<50% to 75%	<75%
----------------	-----	-----------	-----------	------------	-------------	-------------	-------------	------

OTUs presented here accounted for 92% of total reads and are displayed according to their taxonomic group’s affiliations. Grey shade represents the relative abundance (rDNA-based) of each OTU within each sample, white box indicate absence while darker grey represent higher abundance (see ribbon). Values in each box represent the rRNA:rDNA ratio (for calculation see material and method section).

**MALV, Fungi and other potential parasites and decomposers:** This group constituted about a third of the 100 most abundant OTUs (~27% of the sequences). Diversity was highest in winter and lowest during the spring bloom of *P. globosa*. No clear seasonal pattern was observed for this group (Table 3). The most abundant OTU of the group was Syndiniales *GroupI-Clade1* (9.9% of all sequences), which appeared mainly from September to February. More generally, Syndiniales from groups 1, 2, and 3 displayed their highest activity in September/November, just before their peak of abundance that occurred in November/December depending on the year (Table 3). A similar pattern was observed for *Thraustochytriaceae*, with increased activity happening prior the highest abundance that occurred between May and September. Overall, throughout the year it was possible to reconstitute the following succession pattern: *Chytridiomycota* thrived in winter conditions; *Olpidiopsis* sp. (OTU184) appeared only in February; *Pseudoperkinsus tapetis* was observed from February or April until September depending on the year; *Thraustochytriaceae* sp. appeared from May to July; *Haliphthorales* sp. was almost exclusively found in July; *Olpidiopsis porphyrae* was abundant from July to February; *Cryptomycotina* sp. was identified during winter conditions from November to March depending on the year (Table 3). *Engyodontium album* and *Ichthyospora* sp. were found year round in almost every sample during the study period. Overall, their rRNA and rDNA relative reads abundance plotted close to the 1:1 line (Fig 6a), but this linear relation decreased to 0.9:1 for fungi. Except during their winter peak, fungi showed lower potential activity than the rest of this community (Figs 6a and 6b, Table 3).

**Ciliates:** This group represented around 4% of the 100 most abundant OTUs' sequences. The most abundant organism, observed all year long, was *Strombidium basimorphum* (OTU5). Other *Strombidinopsis* (OTU44, OTU92 and OTU112), along with *Choreotrichia* sp. (OTU78), *Cyclotrichia* sp. (OTU56), and *Pelagostrobilidium* sp. (OTU141) were also abundant between November and March. In contrast, two OTUs, affiliated to *Parastrombidinopsis shimi* and *Pseudocohnilembus persalinus*, were present during a narrower period, from December to February with high abundance and activity (Table 3). All these ciliates are nano-grazers feeding on small sized phytoplankton. Overall, the most abundant ciliate OTUs displayed high rDNA and low rRNA reads number, possibly corresponding to species with high abundance but relatively low cell activities (Fig 6a).

Higher relative activities for ciliates were observed during the spring and summer period (Fig 6b).

**MAST and choanoflagellates:** This group accounted for about 2% of the 100 most abundant OTU sequences and was restricted to Marine Stramenopile (MAST) and Choanoflagellida. The Choanoflagellida, including its most abundant OTU, *Stephanoeca cauliculata* (OTU6), were mostly observed between September and February, with increasing rRNA reads in the middle of this period (Table 3). The most abundant MAST-7C (OTU18) appeared in nearly all samples, with the highest cell activity in September and the highest abundance in December. MAST-1A (OTU52) peak of abundance occurred in February/March, MAST-2C was mostly present from April to early July, while MAST-1B showed irregular abundance peaks from May to June and from November to February (Table 3). Overall, higher MAST cells activity was restricted to the summer period (Fig 6b).

**Phytoplankton (diatoms, nano and picophototrophs):** This group accounted for 9% of the 100 most abundant OTUs. The most abundant OTU of this group was affiliated with the diatom species *Leptocylindrus aporus* (OTU16: 3.4% of all reads) and was detected between April and September. The increase of *L. aporus* rDNA and rRNA reads displayed similar temporal pattern, but with earlier increase of rRNA reads (Table 3). *L. aporus* (OTU16), *Guinardia flaccida* (OTU132), and *G. delicatula* (OTU266) seemed to increase at the same period. The species *G. delicatula* (OTU88) and the genus *Eucampia* sp. (OTU221) increased between September and December. It should be noted that different OTUs affiliated to the same species: *G. delicatula* (OTU88 and OTU266) and *G. flaccida* (OTU132 and OTU581) displayed different succession patterns. Several phytoplankton species, including the diatoms *Ditylum brightwellii* (OTU47) and *Thalassiosira hispida* (OTU17), the Mamiellophyceae *Bathycoccus prasinus* (OTU54) and *Micromonas* sp. (OTU66, OTU213), the Trebouxiophyceae *Chlorellales* sp. (OTU41), the Pyramimonadophyceae *Pycnococcus provasolii* (OTU110) and *Bolidophyceae* sp. (OTU168), increased between September and March. Generally, their RNA:DNA ratios were higher just before their peak of abundance and decreased drastically when their abundance became the highest (e.g. OTU16). Phytoplankton rRNA:rDNA ratio was highest

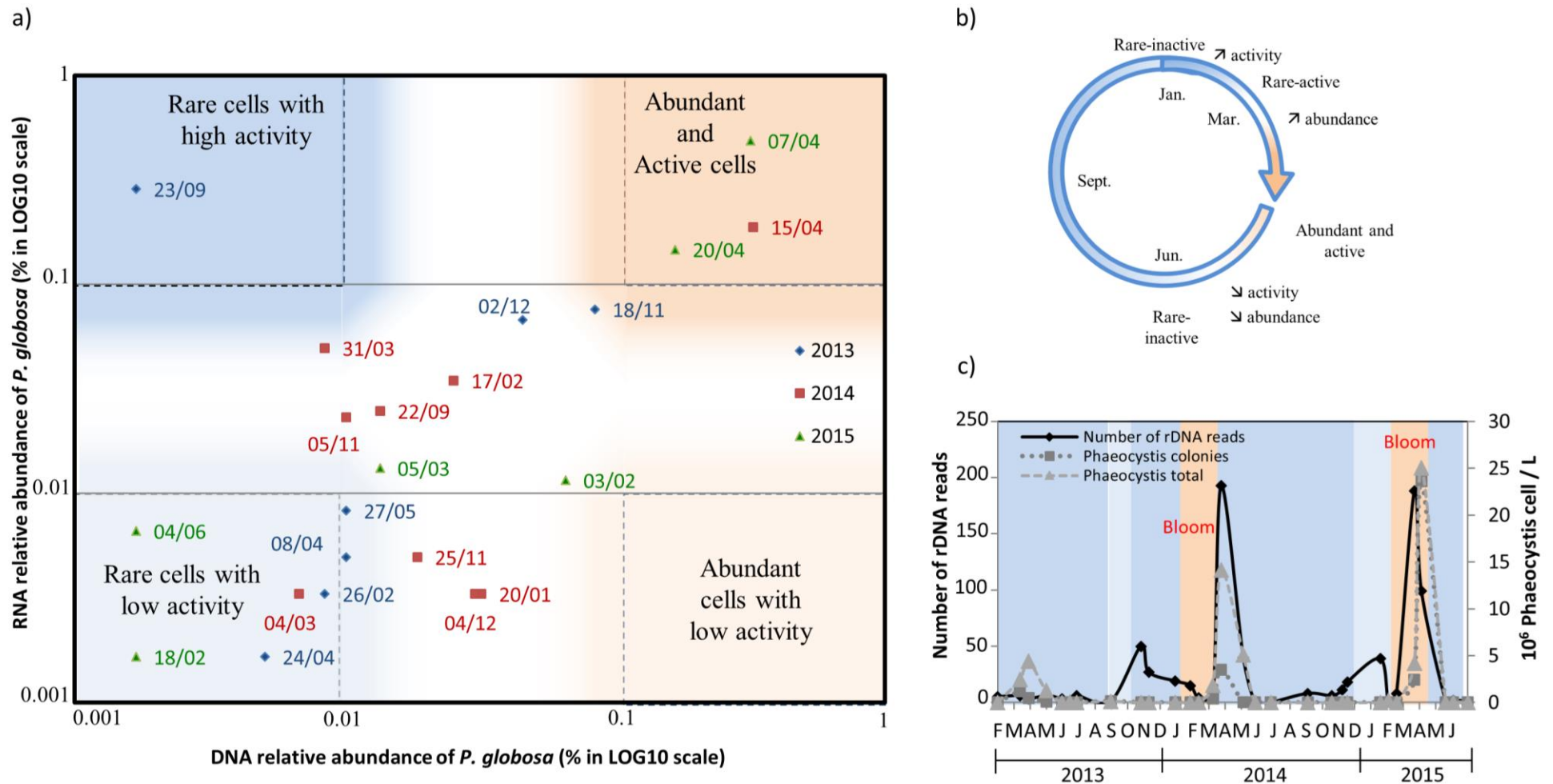
in spring and summer (Fig 6b), with Chlorophytes displaying a median rRNA:rDNA ratio ~8.8 and ~6.5 in spring and summer, respectively, and Diatoms ~3.5 in spring (Fig 6b).

### ***Phaeocystis globosa* life-cycle revealed by rRNA:rDNA ratio**

From February to early-March, *P. globosa* was rare (rDNA reads abundance < 0.01% of all reads) and showed very low rRNA content (rRNA reads abundance < ~0.01% of all reads). In late March, *P. globosa* activity increased according to rRNA data but its abundance remained low according to rDNA data (rDNA < 0.01%, rRNA > 0.01%) (Fig 7a and 7b). Subsequently, it became abundant during its April bloom (rDNA>0.1%, rRNA>0.1%). Abundance (rDNA) began to decrease at the end of the bloom in mid-May, which was confirmed by decreasing number of *P. globosa* cells observed under microscope (Fig 7c). No rRNA reads were detected at the end of the senescence phase. From June to early February, *P. globosa* was again rare and display low relative cell activity (DNA<0.1%, RNA<0.1%, Fig 7a and 7b).

In 2013, the *P. globosa* bloom was observed with lower magnitude under microscope as compared to next two years, and was undetected with the Illumina approach (Fig 7c). Notwithstanding in November 2013 and February 2015, *P. globosa* rDNA read-increase was detected by Illumina sequencing suggesting a micro-bloom (Fig 7a and 7c). Interestingly, before the micro-bloom in November 2013, a significant increase of rRNA read abundance was observed in late September. This rRNA increase suggests higher relative cell activity in response to increase in nutrients (i.e  $\text{NO}_3^- + \text{NO}_2^-$  19.1  $\mu\text{M}$ ) and SST (14.7°C).





**Fig 7. Relation between abundance (rDNA reads) and relative cell activity (rRNA reads) for *Phaeocystis globosa*.** a) Relative rRNA and rDNA reads abundance corresponding to *P. globosa* between 2013 and 2015. Relative reads abundance for each sample were plotted on logarithmic scale. Considering abundance (rDNA-based data) and relative cell activity (rRNA-based data), four quadrants stand out: rare cells with low activity (DNA<0.01%, RNA<0.01%), rare-active cells (DNA<0.01%, RNA>0.1%), abundant-active cells (DNA>0.1%, RNA>0.1) and abundant cells with low activity (DNA>0.1%, RNA<0.01), respectively. Colors represent rDNA reads abundance (blue: rare, red: abundant), and shades represent rRNA reads abundance as proxy for relative cell activity (dark shade: active, light shade: low activity). b) Schematic representation of *P. globosa* annual pattern of abundance and activity in the EEC. c) Microscopic counts of *P. globosa* (dash light grey line: total number of *Phaeocystis* cells; dash dark grey line : number of *Phaeocystis* colony) and number of rDNA Illumina reads (black line).

## DISCUSSION

### Association between environmental variables and community structure

During this four-year study, environmental variables suggested three distinct seasonal periods: spring-early summer (bloom event of *P. globosa*), summer-early autumn, and late autumn-winter. As demonstrated by cluster analysis, microbial community structure (OTU) was generally associated with these three environmental periods. At the end of winter, the abundance of nitrate and high N:P and N:Si ratios triggered blooms of *P. globosa*, following by rises of colonial diatoms and heterotrophic dinoflagellates [41,42,43]. Microbial eukaryote diversity was at the lowest during *P. globosa* bloom period, but gradually increased to reach its highest in fall-winter. Symbionts and degraders accounted for more than one third of the OTU diversity in fall-winter [67]. Microbial eukaryote community structure and seasonal succession were similar between surveys and years for the most abundant taxonomic groups, representing the majority of reads (>90). Differences for rare groups (composed of less than ~10 OTUs) might result from natural diversity fluctuations in an open ecosystem. However, other biases due to random sampling, sequencing methods, PCR and sequencing errors might have also contributed to this disparity [68]. In addition, previous studies in the area have suggested that inter-taxa relations, rather than environmental variables, were the main drivers of microbial community structure and temporal succession [45,67]. These studies showed that environmental variables could explain only around 30 % of microbial eukaryote succession [45], and few correlations were observed, through network analysis, between OTUs and environmental variables while correlations between microbes dominated the network [67].

### rDNA and rRNA-based microbial diversity survey

The diversity of planktonic eukaryotes was similar whether the target was rDNA or rRNA for the most abundant higher taxonomic groups (Fig 3b and 3e). In the cluster analysis, when considering the number of reads, one third of rDNA/rRNA pairs were grouped together, and 28 out of 30 pairs appeared in the same seasonal cluster. This result, although expected, suggests similar abundance estimation for most OTUs whether rRNA or rDNA is

sequenced. However, weaker cluster definition and lower percentage of similarity is observed when considering the diversity of rare OTUs. This suggests that either reads of rare OTUs reached the detection limit of the sequencing method and therefore were not statistically significant, or that biases mentioned earlier hampered accurate detection for rare OTUs. In addition, we observed that RNA-based Bray-Curtis dissimilarity analysis has more consistent sample clustering according to environmental variables compared to DNA-based analysis (Fig 5c and 5d). This suggests that OTUs revealed by rRNA sequencing provides clearer seasonal representation of species succession and level of activity than rDNA-based OTU detection [3,26,69,70]. On the other hand, the DNA molecule is more stable than the RNA molecule and therefore allows thorough investigation for all present microbial species [70]. However, such stability could lead to several biases by including dead or dormant cells, and/or including DNA from the dissolved extracellular pools [71]. The rRNA-based surveys better depict microbial diversity, since they are less affected by ribosomal operon copy number between taxa [23] and reflect more accurately environmental changes due to their shorter lifetime compared to rDNA [25]. Overall, analyzing both rRNA and rDNA provide complementary information and should be considered for comprehensive understanding of ecosystems functioning.

### **Relative activities of major microbial taxonomic and trophic groups**

The comparison between rRNA- and rDNA-based surveys face several limitations including variations in gene copy number [66,72,73], cell size [65,66], and differences in life histories (e.g. dormant cells), and non-growth activities of certain species [29]. Nevertheless, this approach is nowadays generally accepted amongst environmental microbiologists to distinguish and quantify active species among present species, through calculation of the rRNA:rDNA ratio [12,22,23,74]. In our study, taxonomic groups in general displayed an average rRNA:rDNA ratio around 1:1, suggesting a linear relation between cell abundance (rDNA) and activities (rRNA). Two groups (i.e. Ciliophora and Choanoflagellida) appeared to have  $> 1$  mean rRNA:rDNA ratio (slope  $> 1.25$ , Fig 6a). Therefore, these groups might be classified as RNA-prevalent, according to the definition of Massana et al., 2015. In addition, for these two groups, the negative b-value (intersection of the x- and y-axis) (Fig 6a) suggests that the relatively abundant OTUs were active, while the less abundant OTUs had lower than average relative activity. On the opposite, Fungi

displayed  $< 1$  mean rRNA:rDNA ratio (slope=0.9  $\pm$ 0.07, Fig 6) and a positive b-value (b=+0.13  $\pm$ 0.004, Fig 6), that placed fungi in the DNA-prevalent category. According to these results, rare Ciliophora and Choanoflagellida probably have lower relative activity than the abundant ones, while rare Fungi display higher relative activity as compared with abundant fungi. However, care should be taken not to over-interpret the ecological relevance of prevalent (rDNA or rRNA) groups, and rRNA:rDNA regression. In fact, difference in reads' abundance between rDNA- and RNA-based surveys might not necessarily result from variation in relative activities, but can also reflect differences in genome architecture between taxonomic groups, with RNA-prevalent groups having lower rDNA copy number [12]. It is worthy to underline that, while the overall rRNA:rDNA ratio was close to 1 for most taxonomic groups (ranging from 0.9 to 1.25), it was highly variable when considering individual OTUs (Fig 6a). Consequently, we suggest that the rRNA:rDNA ratio is more informative for individual OTUs, rather than whole taxonomic groups. Therefore, for each particular OTUs, significant changes in the rDNA:rRNA ratio may indicate ecological seasonal events, as for instance, it was observed for *P. globosa* (see the next section). Following RNA:rDNA ratio for individual OTUs, we often observed an increase of relative cells activity (rRNA reads) before increase of abundance (rDNA), suggesting the rRNA reads might be used as an early indicator of short-term microbial community dynamics in natural ecosystems. However, increased rRNA can be an indicator of any type of activities and not only those related to cell proliferation, thus it should always be interpreted with caution for "predicting" upcoming cell increase. Still, the rRNA:rDNA ratio provided insights into seasonal changes in activities of eukaryotic microbial taxa. We found that most groups (i.e Dinoflagellates, Ciliates, Chlorophytes, Haptophyta, Diatoms, and Fungi) had OTUs displaying their highest activity in spring and summer (Fig 6b). A similar observation was made at a coastal station in the eastern North Pacific where the majority of protistan groups had significant higher rRNA:rDNA ratios in April [26]. However, the relative activity across OTUs is very different, as shown by long whiskers in the rRNA:rDNA boxplot diagram (Fig 6b). Therefore, thorough investigation of each individual OTU's variation in rRNA:rDNA ratio, and their association with environmental parameters, is necessary to understand species succession and seasonal response in ecosystems. Such analysis performed for *P. globosa* demonstrated its relevance (next section). Finally, while rRNA:rDNA ratio may reveal cells activity, information about which kind of activity is expressed remains unknown. In future studies, *in situ* metatranscriptomic analysis on time-series will help to better understand how species adapt

their physiology and metabolism in response to changing environmental conditions, how they interact with other organisms, and what are their roles and functions in ecosystems.

### **Phases of *Phaeocystis globosa* life-cycle emerge from the rRNA:rDNA ratio**

*Phaeocystis* has a complex life cycle alternating between solitary cells and colonies, and is capable of forming massive, quasi-monospecific blooms in naturally or anthropogenically induced nitrogen rich areas [75]. The present data based on high throughput sequencing of rDNA and rRNA detected the bloom of *P. globosa* but also showed that *P. globosa* is always present in the ecosystem throughout the entire year. Yet, while rRNA:rDNA ratio used to evaluate cells activity was informative, it appeared that simply calculating this ratio only gave partial results for activity as rRNA and rDNA sequences often increase proportionally (Fig 6). In order to determine the general trend for rRNA and rDNA increase or decrease during time-series, graphical relation between relative amount of rRNA and rDNA reads was constructed (Fig 7a). To note that determining the level of relative reads abundance to differentiate rare / abundant and active / inactive organisms is not straightforward, as it is affected by nucleic acids extraction efficiency, primer specificities, and by variation of rDNA gene copy number per cell. Our choice was to select levels of abundance and activity presented on Fig 7a considering that cells were rare when their relative rDNA reads abundance on total rDNA reads was below 0.01 % and abundant when the relative rDNA reads number was above 0.1% [10]. The same thresholds were considered for rRNA in order to differentiate high versus low activity cells. These levels are only indicative and can be refined in the future by additional data. This approach allowed delimiting four regions corresponding to: (1) rare cells with low activity, (2) rare and active cells, (3) abundant cells with low activity, and (4) abundant and active cells (Fig 7a). Samples that did not belong to any of the above categories could be considered in transition state between phases. Overall, the levels used here to delimit abundance and activity unveiled quite well the bloom of *Phaeocystis* in early- / mid-April, but more interestingly, the increase of cells activity before the bloom and activity decrease after the bloom. According to these results, rRNA reads number could be an indicator of an upcoming bloom, provided that the sampling is frequent enough. In 2013, the bloom of *P. globosa* was not detected in the illumina dataset but was detected in the pyrosequencing dataset [45,46]. This could be due to the degradation of ribosomal DNA prohibiting its

detection, combined with the weaker bloom observed that year under microscope. Finally, a fall-winter increase of *P. globosa* was shown by molecular analysis (rDNA) at the end of 2013 and, to a lower extent, in early 2015. In end-September 2013, an increase of rRNA sequences, suggesting higher cell activities, preceded the increase of abundance (rDNA sequences) three weeks later, in mid-November 2013. Data point characteristics for the increase are positioned within transition phases (e.g November 18<sup>th</sup> and December 2<sup>nd</sup> 2013, Fig 7a). This is not surprising considering the lower magnitude of this bloom, as compared with the main mid-April bloom. It also highlights the difficulty to clearly define boundaries between rare/abundant and active/inactive cells.

Concluding, this study showed consistent detection of microbial eukaryote diversity whether targeting rDNA or rRNA, and the possibility to use rRNA:rDNA relative read abundance as a "life-event" indicator. Generalizing this approach on individual OTUs would contribute into improving our understanding of microbial community succession and of the underlying factors structuring the community.

## **ACKNOWLEDGMENTS**

This study was supported by a PhD grant from region Nord Pas de Calais (France), FRB-DEMO (FRB 2013) program from region "Nord Pas de Calais" (France) and the SOMLIT network.

## REFERENCES

1. Caron DA, Countway PD, Jones AC, Kim DY, Schnetzer A (2012) Marine Protistan Diversity. *Annual Review of Marine Science* 4: 467-493.
2. Sherr B, F., Sherr EB, Caron DA, Vaulot D, Worden AZ (2007) Oceanic Protists. *Oceanography* 20: 130-134.
3. Countway PD, Vigil PD, Schnetzer A, Moorthi SD, Caron DA (2010) Seasonal analysis of protistan community structure and diversity at the USC Microbial Observatory (San Pedro Channel, North Pacific Ocean). *Limnology and Oceanography* 55: 2381-2396.
4. Medinger R, Nolte V, Pandey RV, Jost S, OttenwÄlder B, et al. (2010) Diversity in a hidden world: potential and limitation of next-generation sequencing for surveys of molecular diversity of eukaryotic microorganisms. *Molecular Ecology* 19: 32-40.
5. Behnke A, Engel M, Christen R, Nebel M, Klein RR, et al. (2011) Depicting more accurate pictures of protistan community complexity using pyrosequencing of hypervariable SSU rRNA gene regions. *Environmental Microbiology* 13: 340-349.
6. Caron DA (2013) Towards a Molecular Taxonomy for Protists: Benefits, Risks, and Applications in Plankton Ecology. *Journal of Eukaryotic Microbiology* 60: 407-413.
7. Amaral-Zettler LA, McCliment EA, Ducklow HW, Huse SM (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.
8. Cheung MK, Au CH, Chu KH, Kwan HS, Wong CK (2010) Composition and genetic diversity of picoeukaryotes in subtropical coastal waters as revealed by 454 pyrosequencing. *The ISME Journal* 4: 1053-1059.
9. Edgcomb V, Orsi W, Bunge J, Jeon S, Christen R, et al. (2011) Protistan microbial observatory in the Cariaco Basin, Caribbean. I. Pyrosequencing vs Sanger insights into species richness. *The ISME Journal* 5: 1344-1356.
10. Logares R, Audic S, Bass D, Bittner L, Boutte C, et al. (2014) Patterns of Rare and Abundant Marine Microbial Eukaryotes. *Current Biology* 24: 813-821.
11. Massana R, Pedros-Alio C (2008) Unveiling new microbial eukaryotes in the surface ocean. *Curr Opin Microbiol* 11: 213-218.

12. Massana R, Gobet A, Audic S, Bass D, Bittner L, et al. (2015) Marine protist diversity in European coastal waters and sediments as revealed by high-throughput sequencing. *Environmental Microbiology* 17: 4035-4049.
13. Guillou L, Viprey M, Chambouvet A, Welsh RM, Kirkham AR, et al. (2008) Widespread occurrence and genetic diversity of marine parasitoids belonging to Syndiniales (Alveolata). *Environmental Microbiology* 10: 3349-3365.
14. Bråte J, Krabberød AK, Dolven JK, Ose RF, Kristensen T, et al. (2012) Radiolaria Associated with Large Diversity of Marine Alveolates. *Protist* 163: 767-777.
15. Kerkhof L, Ward BB (1993) Comparison of Nucleic Acid Hybridization and Fluorometry for Measurement of the Relationship between RNA/DNA Ratio and Growth Rate in a Marine Bacterium. *Applied and Environmental Microbiology* 59: 1303-1309.
16. Bremer H, Dennis PP (1996) Modulation of Chemical Composition and Other Parameters of the Cell at Different Exponential Growth Rates. In: Neidhardt ea, editor. *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, 2nd ed. 2nd ed. Washington, DC 20036-2904. USA: ASM Press.
17. Mills HJ, Martinez RJ, Story S, Sobecky PA (2005) Characterization of Microbial Community Structure in Gulf of Mexico Gas Hydrates: Comparative Analysis of DNA- and RNA-Derived Clone Libraries. *Applied and Environmental Microbiology* 71: 3235-3247.
18. Moeseneder MM, Arrieta JM, Herndl GJ (2005) A comparison of DNA- and RNA-based clone libraries from the same marine bacterioplankton community. *FEMS Microbiology Ecology* 51: 341-352.
19. Gentile G, Giuliano L, D'Auria G, Smedile F, Azzaro M, et al. (2006) Study of bacterial communities in Antarctic coastal waters by a combination of 16S rRNA and 16S rDNA sequencing. *Environmental Microbiology* 8: 2150-2161.
20. Campbell BJ, Yu L, Heidelberg JF, Kirchman DL (2011) Activity of abundant and rare bacteria in a coastal ocean. *Proceedings of the National Academy of Sciences* 108: 12776-12781.
21. Hugoni M, Taib N, Debroas D, Domaizon I, Jouan Dufournel I, et al. (2013) Structure of the rare archaeal biosphere and seasonal dynamics of active ecotypes in surface coastal waters. *Proceedings of the National Academy of Sciences* 110: 6004-6009.
22. Stoeck T, Zuendorf A, Breiner HW, Behnke A (2007) A molecular approach to identify active microbes in environmental eukaryote clone libraries. *Microb Ecol* 53: 328-339.
23. Not F, del Campo J, Balagué V, de Vargas C, Massana R (2009) New Insights into the Diversity of Marine Picoeukaryotes. *PLoS ONE* 4: e7143.



24. Orsi W, Biddle JF, Edgcomb V (2013) Deep sequencing of subseafloor eukaryotic rRNA reveals active Fungi across marine subsurface provinces. *PLoS ONE* 8: e56335.
25. Charvet S, Vincent WF, Lovejoy C (2014) Effects of light and prey availability on Arctic freshwater protist communities examined by high-throughput DNA and RNA sequencing. *FEMS Microbiology Ecology* 88: 550-564.
26. Hu SK, Campbell V, Connell P, Gellene AG, Liu Z, et al. (2016) Protistan diversity and activity inferred from RNA and DNA at a coastal ocean site in the eastern North Pacific. *FEMS Microbiology Ecology* 92: fiw050.
27. Kemp PF, Lee S, LaRoche J (1993) Estimating the Growth Rate of Slowly Growing Marine Bacteria from RNA Content. *Applied and Environmental Microbiology* 59: 2594-2601.
28. Poulsen LK, Ballard G, Stahl DA (1993) Use of rRNA fluorescence in situ hybridization for measuring the activity of single cells in young and established biofilms. *Applied and Environmental Microbiology* 59: 1354-1360.
29. Blazewicz SJ, Barnard RL, Daly RA, Firestone MK (2013) Evaluating rRNA as an indicator of microbial activity in environmental communities: limitations and uses. *The ISME Journal* 7: 2061-2068.
30. Tagkopoulos I, Liu Y-C, Tavazoie S (2008) Predictive behavior within microbial genetic networks. *Science* 320: 1313-1317.
31. Mitchell A, Romano GH, Groisman B, Yona A, Dekel E, et al. (2009) Adaptive prediction of environmental changes by microorganisms. *Nature* 460: 220-224.
32. Barnard RL, Osborne CA, Firestone MK (2013) Responses of soil bacterial and fungal communities to extreme desiccation and rewetting. *The ISME Journal* 7: 2229-2241.
33. Zinser ER, Lindell D, Johnson ZI, Futschik ME, Steglich C, et al. (2009) Choreography of the Transcriptome, Photophysiology, and Cell Cycle of a Minimal Photoautotroph, *Prochlorococcus*. *PLoS ONE* 4: e5135.
34. Lepp PW, Schmidt TM (1998) Nucleic acid content of *Synechococcus* spp. during growth in continuous light and light/dark cycles. *Archives of Microbiology* 170: 201-207.
35. Peperzak L, Colijn F, Gieskes WWC, Peeters JCH (1998) Development of the diatom-*Phaeocystis* spring bloom in the Dutch coastal zone of the North Sea: the silicon depletion versus the daily irradiance threshold hypothesis. *Journal of Plankton Research* 20: 517-537.
36. Egge JK, Aksnes DL (1992) Silicate as Regulating Nutrient in Phytoplankton Competition. *Marine Ecology Progress Series* 83: 281-289.

37. Breton E, Rousseau V, Parent J-Y, Ozer J, Lancelot C (2006) Hydroclimatic modulation of diatom/Phaeocystis blooms in nutrient-enriched Belgian coastal waters (North Sea). *Limnology and Oceanography* 51: 1401-1409.
38. Jahnke J (1989) The light and temperature dependence of growth rate and elemental composition of *Phaeocystis globosa* scherffel and *P. Pouchetii* (HAR.) Lagerh. in batch cultures. *Netherlands Journal of Sea Research* 23: 15-21.
39. Rousseau V, Chrétiennot-Dinet MJ, Jacobsen A, Verity P, Whipple S (2007) The life cycle of *Phaeocystis*: State of knowledge and presumptive role in ecology. *Biogeochemistry* 83: 29-47.
40. Breton E, Brunet C, Sautour B, Brylinski JM (2000) Annual variations of phytoplankton biomass in the Eastern English Channel: comparison by pigment signatures and microscopic counts. *Journal of Plankton Research* 22: 1423-1440.
41. Schapira M, Vincent D, Gentilhomme V, Seuront L (2008) Temporal patterns of phytoplankton assemblages, size spectra and diversity during the wane of a *Phaeocystis globosa* spring bloom in hydrologically contrasted coastal waters. *Journal of the Marine Biological Association of the United Kingdom* 88: 649-662.
42. Grattepanche JD, Vincent D, Breton E, Christaki U (2011) Microzooplankton herbivory during the diatom–*Phaeocystis* spring succession in the eastern English Channel. *Journal of Experimental Marine Biology and Ecology* 404: 87-97.
43. Grattepanche J-D, Breton E, Brylinski J-M, Lecuyer E, Christaki U (2011) Succession of primary producers and micrograzers in a coastal ecosystem dominated by *Phaeocystis globosa* blooms. *Journal of Plankton Research* 33: 37-50.
44. Christaki U, Kormas KA, Genitsaris S, Georges C, Sime-Ngando T, et al. (2014) Winter–summer succession of unicellular eukaryotes in a meso-eutrophic coastal system. *Microbial Ecology* 67: 13-23.
45. Genitsaris S, Monchy S, Viscogliosi E, Sime-Ngando T, Ferreira S, et al. (2015) Seasonal variations of marine protist community structure based on taxon-specific traits using the eastern English Channel as a model coastal system. *FEMS Microbiology Ecology* 91: fiv034.
46. Genitsaris S, Monchy S, Breton E, Lecuyer E, Christaki U (2016) Small-scale variability of protistan planktonic communities relative to environmental pressures and biotic interactions at two adjacent coastal stations. *Marine Ecology Progress Series* 548: 61-75.
47. Riley GA (1957) Phytoplankton of the North Central Sargasso Sea, 1950–521. *Limnology and Oceanography* 2: 252-270.

48. Aminot A, K erouel R (2004) Dissolved organic carbon, nitrogen and phosphorus in the N-E Atlantic and the N-W Mediterranean with particular reference to non-refractory fractions and degradation. *Deep Sea Research Part I: Oceanographic Research Papers* 51: 1975-1999.
49. L pez-Garc a P, Philippe H, Gail F, Moreira D (2003) Autochthonous eukaryotic diversity in hydrothermal sediment and experimental microcolonizers at the Mid-Atlantic Ridge. *Proceedings of the National Academy of Sciences of the United States of America* 100: 697-702.
50. Amann RJ, Binder BL, Chisholm SW, Devereux R, Stahl DA (1990) Combination of 16S rRNA targeted oligonucleotide probes with flow-cemetry for analysing mixed microbial populations. *Appl Environ Microbiol* 56: 1919-1925.
51. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, et al. (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* 75: 7537-7541.
52. Schloss PD, Gevers D, Westcott SL (2011) Reducing the Effects of PCR Amplification and Sequencing Artifacts on 16S rRNA-Based Studies. *PLoS ONE* 6: e27310.
53. Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26: 2460-2461.
54. Kunin V, Engelbrekton A, Ochman H, Hugenholtz P (2010) Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. *Environmental Microbiology* 12: 118-123.
55. Guillou L, Bachar D, Audic S, Bass D, Berney C, et al. (2013) The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. *Nucleic Acids Research* 41: D597-D604.
56. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, et al. (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research* 41: D590-D596.
57. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215: 403-410.
58. Hammer  , Harper DAT, Ryan PD (2001) PAST: paleontological statistics software package for education and data analysis. *Palaeontologia Electronica* 4: 9.
59. R-Development-Core-Team (2014) R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

60. Dray S, Dufour A-B (2007) The ade4 Package: Implementing the Duality Diagram for Ecologists. *Journal of Statistical Software* 22: 1-20.
61. Box GEP, Cox DR (1964) An analysis of transformations. *Journal of the Royal Statistical Society Series B (Methodological)* 26: 211-252.
62. Kuhn M, Jed Wing, Steve Weston, Andre Williams, Chris Keefer, et al. (2016) caret: Classification and Regression Training R package (version 6.0-68).
63. Ward JHJ (1963) Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association* 58: 236-244.
64. Clarke KR, Gorley RN (2006) Primer v6: User Manual/Tutorial. Plymouth, UK: PRIMER-E.
65. 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 Microbiology Ecology* 52: 79-92.
66. Godhe A, Asplund ME, Härnström K, Saravanan V, Tyagi A, et al. (2008) Quantification of Diatom and Dinoflagellate Biomasses in Coastal Marine Seawater Samples by Real-Time PCR. *Applied and Environmental Microbiology* 74: 7174-7182.
67. Christaki U, Genitsaris S, Monchy S, Li LL, Rachik S, et al. (2017) Parasitic Eukaryotes in a Meso-Eutrophic Coastal System with Marked *Phaeocystis globosa* Blooms. *Frontiers in Marine Science* 4: 416.
68. Ferrera I, Giner CR, Reñé A, Camp J, Massana R, et al. (2016) Evaluation of Alternative High-Throughput Sequencing Methodologies for the Monitoring of Marine Picoplanktonic Biodiversity Based on rRNA Gene Amplicons. *Frontiers in Marine Science* 3: 147.
69. Schnetzer A, D. Moorthi S, Countway P, Gast R, Gilg I, et al. (2011) Depth matters: Microbial eukaryote diversity and community structure in the eastern North Pacific revealed through environmental gene libraries.: Elsevier.
70. Karl DM, Bailiff MD (1989) The measurement and distribution of dissolved nucleic acids in aquatic environments. *Limnology and Oceanography* 34: 543-558.
71. Xu D, Li R, Hu C, Sun P, Jiao N, et al. (2017) Microbial Eukaryote Diversity and Activity in the Water Column of the South China Sea Based on DNA and RNA High Throughput Sequencing. *Frontiers in Microbiology* 8: 1121.
72. Prokopowich CD, Gregory TR, Crease TJ (2003) The correlation between rDNA copy number and genome size in eukaryotes. *Genome* 46: 48-50.

73. Gong J, Dong J, Liu X, Massana R (2013) Extremely High Copy Numbers and Polymorphisms of the rDNA Operon Estimated from Single Cell Analysis of Oligotrich and Peritrich Ciliates. *Protist* 164: 369-379.
74. Jones SE, Lennon JT (2010) Dormancy contributes to the maintenance of microbial diversity. *Proceedings of the National Academy of Sciences* 107: 5881-5886.
75. Verity PG, Brussaard CP, Nejstgaard JC, Van Leeuwe MA, Lancelot C, et al. (2007) Current understanding of Phaeocystis ecology and biogeochemistry, and perspectives for future research. *Biogeochemistry* 83: 311-330.

## SUPPORTING INFORMATION

**S1 Table. Physical-chemical parameters and Chlorophyll a (Chl. a) from March 2011 to July 2015 at the SOMLIT station in the eastern English Channel (SOMLIT network, <http://somlit.epoc.u-bordeaux1.fr/fr/>). Kd: diffuse attenuation coefficient, PAR: photosynthetically active radiation, T: temperature, S: salinity, O<sub>2</sub>: Oxygen, pH: potential hydrogen, NO<sub>3</sub><sup>-</sup>: nitrate, NO<sub>2</sub><sup>-</sup>: nitrite, PO<sub>4</sub><sup>3-</sup>: phosphate, SiOH<sub>4</sub>: silicate, POC: particulate organic carbon, PON: particulate organic nitrogen, SPM: suspended particular matter, Chl. a: chlorophyll a, and N/P: nitrate/phosphate ratio.**

Dates	Kd (m <sup>-1</sup> )	PAR (E m <sup>-2</sup> d <sup>-1</sup> )	T (°C)	Salinity	O <sub>2</sub> (mg L <sup>-1</sup> )	pH	NO <sub>3</sub> +NO <sub>2</sub> (µM)	PO <sub>4</sub> (µM)	SiOH <sub>4</sub> (µM)	POC (µgC L <sup>-1</sup> )	PON (µgN L <sup>-1</sup> )	SPM (mg L <sup>-1</sup> )	Chl a (µg L <sup>-1</sup> )	N/P
07/03/2011	0,31	19,28	6,80	34,20	7,50	8,26	4,32	0,10	0,25	219,72	38,67	4,36	5,80	43,20
21/03/2011	0,17	19,03	7,60	33,90	8,10	8,37	7,10	0,10	0,10	207,06	45,17	2,91	4,50	71,00
04/04/2011	0,43	17,85	9,00	34,40	7,60	8,40	1,57	0,07	0,17	300,77	45,63	3,29	9,90	22,40
18/04/2011	0,36	25,76	10,50	34,20	8,10	8,42	0,20	0,34	2,37	394,48	46,09	7,04	11,70	0,59
04/05/2011	0,30	37,02	11,80	34,50	6,50	8,30	0,10	0,13	0,43	355,94	60,26	6,29	1,90	0,77
06/06/2011	0,21	38,66	15,20	34,50	5,50	8,24	0,31	0,30	1,57	278,82	58,62	3,69	1,30	1,00
15/06/2011	0,28	34,31	15,50	34,70	5,60	8,22	0,25	0,26	1,29	297,75	51,32	1,52	0,90	1,00
04/07/2011	0,22	51,96	17,70	34,50	5,90	8,22	0,34	0,05	1,66	237,84	35,61	2,28	1,30	6,80
27/09/2011	0,21	23,98	17,00	34,50	5,60	8,20	3,12	0,16	2,88	178,97	35,01	3,13	2,20	19,50
25/10/2011	0,23	17,84	14,20	34,20	5,80	8,22	3,65	0,12	0,87	231,95	50,59	3,21	6,30	30,40
09/11/2011	0,21	4,43	14,20	34,60	5,50	8,16	8,67	0,41	1,82	152,37	19,11	4,16	0,70	21,10
23/11/2011	0,18	10,12	12,10	34,50	5,90	8,11	11,20	0,43	0,35	153,49	22,20	1,59	0,40	26,00
24/01/2012	0,52	3,62	8,10	33,90	6,40	8,18	15,10	0,75	6,21	154,60	25,30	10,31	1,00	20,10
20/03/2012	0,55	19,68	7,40	34,10	8,20	8,34	4,53	0,59	0,51	425,70	25,92	1,41	1,90	7,70
05/04/2012	0,67	19,80	8,50	34,50	8,20	8,38	0,20	1,06	3,10	576,48	5,33	2,76	4,50	0,19
09/05/2012	0,23	20,62	10,90	34,70	6,70	8,33	0,14	1,20	1,07	328,81	36,22	1,86	3,40	0,12
05/06/2012	0,20	27,64	14,30	34,60	6,40	8,30	0,11	0,10	0,83	520,67	35,74	0,96	2,10	1,10
21/06/2012	0,19	47,41	15,80	34,40	5,70	8,24	1,02	0,42	0,93	341,90	41,85	0,10	0,50	2,43
04/07/2012	0,32	40,65	17,00	34,00	6,20	8,22	0,61	0,21	0,37	494,52	56,78	1,16	8,00	2,90
23/07/2012	0,36	38,93	17,60	33,50	5,90	8,23	0,18	0,22	0,96	481,78	61,56	2,09	7,90	0,82
03/09/2012	0,28	25,74	18,50	34,20	5,40	8,14	0,10	0,22	1,74	268,62	37,72	1,27	2,80	0,45
30/10/2012	0,50	7,84	12,80	33,30	5,70	8,12	12,10	0,42	7,70	230,00	50,00	7,86	1,00	28,80

13/11/2012	0,61	5,91	11,80	34,20	5,90	8,10	11,10	0,45	5,30	150,00	18,00	3,01	1,80	24,70
11/02/2013	0,45	3,80	6,50	33,80	6,80	8,19	9,23	0,43	2,92	260,00	44,85	3,17	3,60	21,45
26/02/2013	0,49	7,46	5,87	34,30	7,08	8,22	7,35	0,41	1,14	370,43	71,69	3,07	9,95	17,93
26/03/2013	0,27	16,32	5,40	34,40	7,30	8,30	3,15	0,15	0,28	391,24	75,97	2,14	3,20	21,00
08/04/2013	0,49	21,41	5,44	34,41	7,20	8,29	2,89	0,25	0,01	412,04	80,24	6,16	4,42	11,56
24/04/2013	0,30	26,30	7,86	34,41	7,60	8,37	0,60	0,09	0,01	584,20	83,56	2,70	8,33	6,67
27/05/2013	0,23	40,31	11,26	34,10	6,57	8,38	2,10	0,04	0,01	271,12	48,88	1,18	1,94	52,50
10/06/2013	0,64	32,72	11,90	34,30	6,40	8,43	0,29	0,89	1,12	756,00	72,24	2,16	1,20	0,33
25/06/2013	0,66	21,52	14,98	33,64	6,63	8,49	0,21	0,18	0,01	1242,80	95,60	3,14	14,19	1,17
09/07/2013	0,29	42,47	15,71	34,16	5,96	8,39	0,36	0,12	0,10	713,08	68,15	1,90	5,21	3,00
22/07/2013	0,23	46,11	17,31	34,18	5,31	8,28	0,28	0,06	0,79	451,24	50,37	0,65	2,07	4,67
23/09/2013	0,29	21,61	17,01	33,81	5,10	8,22	1,09	0,22	1,88	189,39	32,58	1,77	1,81	4,95
18/11/2013	0,33	4,75	12,39	33,39	5,86	8,17	19,10	0,66	13,60	154,68	39,95	5,42	2,16	28,94
02/12/2013	0,36	4,44	10,20	33,73	6,09	8,17	24,10	0,59	12,90	64,78	28,95	6,39	1,51	40,85
20/01/2014	0,19	4,82	8,69	33,11	6,36	8,21	26,12	0,61	12,32	149,57	24,06	1,67	1,54	42,82
17/02/2014	0,46	9,19	7,89	33,28	6,47	8,01	26,71	0,56	10,66	207,58	44,25	6,15	2,12	47,70
04/03/2014	0,35	13,27	8,42	33,84	6,51	8,05	18,43	0,46	6,44	297,24	38,43	5,27	2,95	40,07
31/03/2014	0,66	19,14	9,62	34,06	7,76	8,30	4,30	0,02	0,26	440,12	52,98	1,72	5,01	215,00
15/04/2014	0,68	28,95	10,93	34,39	7,46	8,33	0,56	0,36	0,51	820,24	140,32	3,56	5,91	1,56
26/05/2014	0,19	33,99	14,16	34,29	5,91	8,19	0,07	0,01	1,80	330,65	84,45	1,20	1,34	7,00
17/06/2014	0,26	38,25	15,81	34,42	5,48	8,15	0,27	0,06	1,69	333,08	89,16	1,47	3,22	4,50
16/07/2014	0,29	38,65	18,42	34,18	5,35	8,22	0,09	0,03	0,70	241,37	48,01	1,82	2,74	3,00
22/09/2014	0,31	22,45	18,47	34,35	4,92	8,10	1,10	0,16	3,25	228,92	50,88	2,01	2,89	6,88
05/11/2014	0,52	5,26	14,65	34,07	5,50	8,13	10,54	0,30	2,64	231,68	59,68	3,73	7,66	35,13
25/11/2014	0,44	5,11	12,52	34,21	5,73	8,12	12,85	0,40	4,87	192,80	55,19	3,91	2,59	32,13
04/12/2014	0,68	2,28	11,69	34,21	5,71	8,12	13,80	0,47	4,94	198,40	46,25	14,22	2,46	29,36
03/02/2015	0,29	6,78	7,18	34,01	6,46	8,09	16,88	0,37	8,58	102,30	33,27	15,43	1,60	45,62
18/02/2015	0,64	11,53	6,44	33,73	7,41	8,14	13,74	0,32	4,64	345,00	65,80	19,96	2,85	42,94
05/03/2015	0,61	15,33	7,06	33,17	7,65	8,26	15,80	0,14	1,16	369,90	43,80	18,95	13,13	112,86
07/04/2015	0,73	17,28	8,98	33,72	7,35	8,30	7,47	0,02	0,41	288,40	55,00	9,03	8,12	373,50
20/04/2015	0,57	30,54	10,31	33,96	8,18	8,38	1,35	0,08	0,38	928,84	180,53	4,97	9,14	16,88
04/06/2015	0,23	36,02	14,58	34,01	6,12	8,24	1,19	0,21	0,29	1125,80	103,80	3,58	4,64	5,67
18/06/2015	0,27	37,37	16,09	33,99	5,74	8,26	1,82	0,03	0,28	522,90	98,90	32,93	2,85	60,67
15/07/2015	0,25	28,39	17,89	34,19	5,07	8,09	0,19	0,06	0,89	355,60	75,90	22,62	1,33	3,17





## **CHAPITRE IV**

**LES PARASITES EUKARYOTES DANS UN SYSTEME COTIER MESO-EUTROPHE  
CARACTERISE PAR DES BLOOMS PHYTOPLANCTONIQUES MARQUES**



## **Parasitic Eukaryotes in a Meso-Eutrophic Coastal System with Marked *Phaeocystis globosa* Blooms**

Urania Christaki, Savvas Genitsaris, Sébastien Monchy, Luen L. Li, Sara Rachik, Elsa Breton and Téléphore Sime-Ngando

### **RESUME DU CONTEXTE ET OBJECTIFS**

Dans ce deuxième chapitre, nous nous sommes focalisés sur l'étude de la composante parasitaire des eucaryotes microbiens en Manche Orientale. L'étude des parasites a pris plus d'importance ces dernières années en raison de leur rôle important dans la structuration des communautés. Ils sont connus par leur faculté d'infecter les organismes à divers niveaux trophiques et d'augmenter la complexité des réseaux trophiques planctoniques en élargissant les chaînes alimentaires (*Grami et al. 2011*). Ces protistes parasites sont difficilement identifiables dans les échantillons environnementaux, en raison de l'absence de caractéristiques morphologiques et de leur mode de vie endosymbiotique. L'utilisation des techniques de séquençage à haut débit (e.g. pyroséquençage) pour leur identification a permis de découvrir de nouveaux groupes taxonomiques eucaryotes (par exemple *Lopez-Garcia et Moreira, 2008*) et dont certains appartenant à des symbiotes et des parasites intracellulaires potentiels (*Brate et al., 2012*).

Cette étude a été réalisée à la station côtière de la Manche Orientale. Cet écosystème méso-eutrophique, comme mentionné dans le chapitre précédent, connaît des proliférations annuelles de l'Haptophyte *Phaeocystis globosa* exerçant un impact sur la structuration et la succession des communautés planctoniques (par exemple, *Schapira et al. 2008* ; *Grattepanche et al., 2011a,b*). Une étude conduite en Manche Orientale sur les successions saisonnières (hiver-été) des protistes au moyen de l'approche de séquençage à haut débit du gène ARNr18S a suggéré que le concept de réseau trophique bien établi dans la Manche Orientale basé sur les interactions prédateur-proie est juste une partie de l'image écologique et que plusieurs organismes utilisant une variété de stratégies auparavant négligées, comme le parasitisme et la saprotrophie, étaient abondamment présents dans la communauté planctonique (*Christaki et al., 2014*).

Dans ce travail, nous nous sommes concentrés sur l'étude des protistes symbiotes/décomposeurs/parasitaires. Pour ceci, nous avons mené une étude de "métadonnée" pour les données obtenues pendant la période 2011-2013 (32 échantillons traités). Ainsi, les objectifs principaux de cette étude ont été d'explorer en premier temps la capacité des réseaux de corrélation pour mettre en évidence les relations entre les taxons symbiotes-décomposeurs et la succession saisonnière du microplancton, et dans un second temps, d'identifier les paramètres environnementaux affectant des groupes taxonomiques et trophiques spécifiques de la communauté du microplancton dans les eaux côtières tempérées de la Manche Orientale.



# Parasitic Eukaryotes in a Meso-Eutrophic Coastal System with Marked *Phaeocystis globosa* Blooms

Urania Christaki<sup>1\*†</sup>, Savvas Genitsaris<sup>1,2†</sup>, Sébastien Monchy<sup>1</sup>, Luen L. Li<sup>1</sup>, Sara Rachik<sup>1</sup>, Elsa Breton<sup>1</sup> and Télésphore Sime-Ngando<sup>3</sup>

<sup>1</sup> Laboratoire d'Océanologie et Géosciences, UMR Centre National de la Recherche Scientifique 8187, Université du Littoral Côte d'Opale, Wimereux, France, <sup>2</sup> School of Economics, Business Administration and Legal Studies, International Hellenic University, Thessaloniki, Greece, <sup>3</sup> Laboratoire Microorganismes: Génome et Environnement, UMR Centre National de la Recherche Scientifique 6023, Clermont Université Blaise Pascal, Aubière, France

## OPEN ACCESS

### Edited by:

Hongyue Dang,  
Xiamen University, China

### Reviewed by:

Esther Garcés,  
Consejo Superior de Investigaciones  
Científicas (CSIC), Spain  
Peter David Roopnarine,  
California Academy of Sciences,  
United States  
Toshi Nagata,  
The University of Tokyo, Japan

### \*Correspondence:

Urania Christaki  
urania.christaki@log.cnrs.fr

<sup>†</sup> These authors have contributed  
equally to this work.

### Specialty section:

This article was submitted to  
Aquatic Microbiology,  
a section of the journal  
Frontiers in Marine Science

Received: 02 July 2017

Accepted: 04 December 2017

Published: 19 December 2017

### Citation:

Christaki U, Genitsaris S, Monchy S,  
Li LL, Rachik S, Breton E and  
Sime-Ngando T (2017) Parasitic  
Eukaryotes in a Meso-Eutrophic  
Coastal System with Marked  
*Phaeocystis globosa* Blooms.  
Front. Mar. Sci. 4:416.  
doi: 10.3389/fmars.2017.00416

This study explored the potential interactions and relations of the putative Symbiotic/Decomposer/Parasitic community (SymbDec) with other taxa and predominant environmental parameters in the eastern English Channel (EEC) over a 2.5 years period (32 sampling dates). The EEC is a meso-eutrophic coastal system portrayed by intense reoccurring patterns in plankton succession with blooms of the Haptophyte *Phaeocystis globosa* and abundant populations of colonial diatoms and dinoflagellate grazers before and after the bloom. The study of the V2-V3 hypervariable region of the 18S rRNA gene allowed an overview of the planktonic community, which consisted of 32 high-level taxonomic groups. Forty-two percent of the identified OTUs belonged to symbionts, decomposers or other putative parasitic taxa belonging mainly to Syndiniales (MALV—MARine ALveolates), Fungi, Cercozoa, Perkinsea, and others, most of which were plankton parasites. We examined the Local Similarity Analysis (LSA) network of lag delayed correlations on the 142 most abundant OTUs (>0.1% of the total number of reads). LSA showed that 99 OTUs had highly significant connections, involving 26 OTUs characterized as potential parasites, and 2 decomposers. The parasitic network had “small world” characteristics, rendering the assemblage more robust to abiotic variations, but more exposed to the removal of highly associated taxa. The majority of the highly connected OTUs belonged to MALV, which were mainly connected with dinoflagellates, suggesting that it might play an important role on dinoflagellate top-down control. The bloom forming *P. globosa* was not observed in connection with parasitic OTUs in the network. Four environmental parameters (O<sub>2</sub>, N, T, and N/P) were strongly connected with only 11 OTUs, while correlations between microbes dominated the network.

**Keywords:** unicellular eukaryotes, 18S rRNA gene, *Phaeocystis*, MALV, coastal waters, network, local similarity analysis

## INTRODUCTION

Marine parasitic protists have been observed and described—along with other protists—from the early microbiologists over the last few centuries (see Caron et al., 2009; Skovgaard, 2014, for a review). Parasitic protists are not easily identified in environmental samples, due to their absence of morphological features and endosymbiotic lifestyle (e.g., De Bary, 1879), and therefore information about their diversity, abundance, and ecology had been scarce until quite recently (Scholz et al., 2016). During the last few decades, using the genetic diversity of the 18S rRNA gene, novel eukaryotic taxonomic groups have been revealed within the marine protistan plankton (e.g., López-García et al., 2001; López-García and Moreira, 2008) with great amount of them belonging to potential intracellular symbionts and parasites (e.g., Guillou et al., 2008; Bråte et al., 2012). The development of NGS (Next Generation Sequencing) tools has suggested that marine parasites are more frequent than the literature has reported (e.g., Lefèvre et al., 2008). Parasitism is without a doubt a relation, which enhances the density and extend foodwebs. This in turn enhances the connectivity between species and makes carbon transfers more efficient, particularly because parasitic taxa have complex life cycles and infect organisms that belong at different trophic levels (e.g., Laferty et al., 2008; Niquil et al., 2011). Thus, the effect of protistan parasites in the marine plankton varies greatly depending on parasite species (Skovgaard, 2014). Taking into consideration the complexity of most microbial interactions, there are substantial limitations in our efforts for direct insight into relating taxonomic datasets to rates of discrete microbial groups exploitable in biogeochemical cycles. In addition, in many microbial communities, recurrent organizational patterns cannot always associate with straightforward community roles and/or detailed dependencies between species owing to missing information on “overlooked” microbial interactions (for a review see Bik et al., 2012).

In an alternative approach, correlation networks have been proposed as a “means to approximate microbial interactions” in a way that these networks can be further used to build models describing ecosystem functioning (Steele et al., 2011; Faust and Raes, 2012; Berry and Widder, 2014; Weiss et al., 2016). They can also be used to reveal trophic relations, food-web connectance and biodiversity loss because of disturbance in the food chain (Dunne et al., 2002). In a seminal paper, Fuhrman and Steele (2008) summarized key approaches, successes, and challenges in this field. They addressed this issue directly by associating specific functions to individual organisms, and indirectly, by investigative the statistical relationships among organisms and environmental variables measured over different temporal and spatial scales (e.g., multivariate and network analysis). Several recent metagenomic studies of marine microbes have utilized network analysis to understand ecological processes. For example, Gilbert et al. (2012) with the use of networks described the connectivity between different eukaryotic and prokaryotic taxonomic groups, and between bacteria and environmental variables through a 6-year time-series dataset in a coastal area. Logares et al. (2014) used networks to identify patterns of rare

and abundant marine unicellular eukaryotes in surface coastal waters of six distinct sites in Europe.

The EEC is a meso-eutrophic coastal system portrayed by intense reoccurring patterns in plankton succession. The recurrent massive proliferation of the Haptophyte *Phaeocystis globosa* is associated to the surplus of nitrogen and silicate limitation at late winter (Peperzak et al., 1998). Abundant populations of colonial diatoms and dinoflagellate grazers are observed before and after the *P. globosa* bloom (e.g., Schapira et al., 2008; Grattepanche et al., 2011a,b and references therein). A winter-summer succession study of planktonic protists, by means of 18S rRNA gene high-throughput sequencing approach, suggested that the previously well-established food web concept in the EEC based on predator-prey (e.g., grazer-phytoplankton) interactions is “just a part of the ecological picture” and in fact several organisms using a variety of previously “overlooked” strategies, such as parasitism and saprotrophy, were abundant members of the planktonic community (Christaki et al., 2014). The objective of the present study was to focus on the putative Symbiotic/Decomposer/Parasitic community (SymbDec) in the EEC over a 2.5 years period (32 dates). The molecular signatures of protists belonging to SymbDec are included in numerous NGS datasets, indicating a major complementary role in top-down population control (Edgcomb, 2016). Thus, the two main objectives were (1) to examine connections between SymbDec taxa and other taxa using network analysis within the seasonal microplankton succession, and (2) to identify the underlying environmental parameters that affect specific taxonomic and trophic groups of the microplankton community; both in the temperate coastal waters of the EEC.

## MATERIALS AND METHODS

### Sample Collection

Subsurface sampling (2 m water depth) was carried out in the EEC at high tide at the coastal station C (50° 40' 75" N, 1° 31' 17" E; 20–25 maximum depth) belonging to the French monitoring network SOMLIT (<http://somalit.epoc.u-bordeaux1.fr/fr/>), from 07 March 2011 to 09 July 2013, on a biweekly basis when local weather conditions permitted. A total of 32 samples were collected in 5 L sterile polyethylene bottles for molecular analysis, kept in the dark at the *in situ* temperature, and filtered at laboratory within 2 h. Before filtering, the samples were screened with a 150 μm mesh to retain larger particles and most metazoa. Next, sequential filtration through 10, 3, and 0.6 μm nucleopore filters with 47 mm diameter, was performed using a peristaltic pump with a very low filtration pressure (15 rpm). The serial filtration was used in order to avoid filter clumping and minimize organism disruption.

The filters were immediately stored at –80°C until molecular analysis. In addition, a subsample of 250 mL was collected and fixed with 2% (final concentration) Lugol-glutaraldehyde solution for microscopic analysis.

### Physical-Chemical Variables and Phytoplankton Abundance

Seawater temperature (°C) and salinity were measured *in situ* using a conductivity-temperature-depth profiling system (CTD

Seabird SBE 25). The level of oxygen was immediately analyzed in triplicate with the Wrinkler method (Aminot and Chaussepied, 1983) according to potentiometric titration with a titrino848 (Methrom). Inorganic nutrient concentrations (nitrate  $[\text{NO}_3^-]$ , nitrite  $[\text{NO}_2^-]$ , phosphate  $[\text{PO}_4^{3-}]$ , and silicate  $[\text{SiO}_4^{4-}]$ ) were determined from 100 mL samples with an Alliance Integral Futura Autoanalyzer II based on Strickland and Parsons (1972) and Aminot and Kerouel (2004).

Abundance of *P. globosa* and diatoms was determined in sedimentation chambers (10–25 ml) according to the Utermöhl method using an inverted microscope (Nikon Eclipse TE 2000-S). Given the complex life-cycle of *P. globosa*, abundance of cells embedded in *P. globosa* colonies was determined according to biovolume measurements at magnification of X100 or X200 of all the colonies present in the sedimentation chamber (Rousseau et al., 1994), while at least 400 free cells were counted according to random fields at X400. For diatoms, at least 400 cells of the most abundant species were counted at X400 and 100 cells of the less abundant species at X200.

## DNA Extraction

The DNA of the planktonic organisms was extracted after collectively pooling the 10, 3, and the 0.6  $\mu\text{m}$  filters, using the PowerWater DNA isolation kit (MoBio Laboratories Inc., CA, USA), and following the manufacturer's protocol. The samples contained between 0.5 and 4.5  $\text{ng } \mu\text{L}^{-1}$  of DNA as measured by the Qubit 2.0 Fluorometer (Thermo Fischer Scientific Inc., Massachusetts, USA).

## PCR and Tag Pyrosequencing

The DNA samples were amplified using the eukaryotic primers 18S-28F (5'-GAACTGCGAATGGCTC-3') (López-García et al., 2003) and Euk-516r (5'-ACCAGACTTGCCCTCC-3') (Amann et al., 1990), which have been designed to amplify a domain around 470–480 bp corresponding to the hyper-variable V2 and V3 eukaryote 18S rRNA gene regions. These primers have been successfully used in previous studies of the protistan community in the area (Christaki et al., 2014; Genitsaris et al., 2015, 2016), and allow a comparison of the present results with previous findings. Polymerase Chain Reaction (PCR) was carried out according to standard conditions for Platinum Taq High-Fidelity DNA Polymerase (Invitrogen), with 5 ng of environmental DNA as a template, using the GeneAmp PCR System Apparatus (Applied Biosystems). Tag Pyrosequencing was carried out by the GenoScreen Company (Lille, France). The library was prepared following the procedures described by Roche (Basel, Switzerland) and used in one plate run on a 454 GS FLX Titanium Sequencer. Pyrosequences were submitted to GenBank-SRA under the accession number SRX768577.

## Tag Pyrosequencing Quality Filtering

All samples produced between 12,581 and 29,712 reads. Downstream read processing was performed using the mothur 1.28.0 software following the standard operating procedure (Schloss et al., 2009, 2011), described in Genitsaris et al. (2015). Briefly, the flowgrams of each sample were extracted, separated and denoised using the mothur implementation of

PyroNoise (Quince et al., 2009). Primer sequences, tag and key fragments were removed, and only reads above 200 bp long with homopolymers shorter than 8 bp were included. The dataset were dereplicated to the unique reads and aligned against the SILVA 108 database. The reads suspected of being chimeras were removed using the UCHIME software (Edgar, 2010). The dataset was normalized to the sample with the lowest number of reads using the subsample command in mothur, so that all samples contained 12,581 reads. These reads were clustered into Operational Taxonomic Units (OTUs) at 97% similarity threshold, using the average neighbor method. Finally, single singletons (i.e., unique amplicons that occurred only once in the whole dataset) were removed, as these are likely erroneous sequencing products (Kunin et al., 2010; Behnke et al., 2011).

## Data Analysis

After tag pyrosequencing filtering and normalization, 1,303 OTUs were produced. Taxonomic classification was assigned using BLASTN (Altschul et al., 1990), according to the PR2 curated database (Guillou et al., 2013). All reads affiliated to metazoan were removed from the dataset, thus 1,242 OTUs belonging exclusively to protists, remained. These OTUs, belonging to nine taxonomic super-groups were sorted into six groups according to their trophic status (Table 1), as in Genitsaris et al. (2016). The six trophic groups that were formed were: SymbDec, including symbionts, decomposers, and parasites; autotrophs; heterotrophs and mixotrophs (based on their metabolism); and micro-, nano-, and picograzers (based on the size of their prey). Briefly, the strategy applied in order to sort the OTUs into trophic groups, was to individually examine the 1,242 OTUs and annotate them to a trophic group using the highest level of information. For OTUs affiliated to micro- and nanoplankton (e.g., dinoflagellates, ciliates, diatoms, cryptophytes, chlorophytes, and the blooming *P. globosa*) the confidence about their trophic role was high, as they had been also detected by microscopy. Conversely, for OTUs that were affiliated with taxonomic groups impossible to detect with microscopy, consideration of their annotation to higher taxonomic groups (e.g., family level) was taken into account, which is considered reliable for the size of the reads that were obtained with pyrosequencing (Bachy et al., 2013; Santoferrara et al., 2014). Thus, the taxa belonging to the groups MALV (MARine ALveolates) and MAST (MARine STRamenopiles) were considered symbionts and nano-grazers, respectively, according to the existing literature (Massana et al., 2006; Skovgaard, 2014; respectively for MALV and MAST) (see Table 1).

Furthermore, the 1,242 OTUs were classified in relation to their total relative abundance. Specifically, the abundant OTUs were defined as those with relative abundances > 0.1% of the total number of reads, following previous studies on prokaryotes (e.g., Galand et al., 2009; Hugoni et al., 2013) and protists (e.g., Logares et al., 2014; Genitsaris et al., 2015). Overall, 142 OTUs were characterized as abundant. These alone, or in combination with the environmental variables, were used in network analysis. Two different correlation techniques were used in order to assess the significant relationships between OTUs, and between OTUs and environmental variables. It should be stated however, that here,

**TABLE 1** | Annotated major trophic role, OTUs number at the sampling station, and related references for the super group annotation.

Super groups	Groups	Number of OTUs	Major trophic role	References
Amoebozoa	Lobosa	5	Micrograzers & some parasites	Lesen et al., 2010 (and references within)
Apusozoa	Apusomonadidae	10	Picograzers (nanoflagellates)	Scheckenbach et al., 2006
	Hilomonadea	2		
Alveolata	Apicomplexa	13	Parasites	Skovgaard, 2014
	Litostomatea	15	Nanograzers & parasites	Lynn, 2008
	Oligohymenophorea	12		
	Phyllopharyngea	9		
	Spirotrichea	54		
	Colpodea	4		
	Other Ciliates	6		
	Dinophyceae	182	Micrograzers (Heterotrophs, Mixotrophs) & parasites	Stoecker, 1999; Sherr and Sherr, 2007
	MALV	154	Parasites	Guillou et al., 2008
	Perkinsea	7	Parasites	Park et al., 2004
Archaeplastida	Chlorophyta	59	Autotrophs	Not et al., 2012
	Rhodophyta	1		
Excavata	Euglenozoa	39	Picograzers (nanoflagellates), Symbionts, Parasites	Cavalier-Smith, 2002
Hacrobia	Centroheliozoa	5	Pico- nanograzers	Burki et al., 2009
	Cryptophyta	12	Autotrophs & mixotrophs	Not et al., 2012
	Haptophyta	30	Autotrophs & mixotrophs	Not et al., 2012
	Katablepharidophyta	8	Picograzers (nanoflagellates)	
	Picobilliphyta	11	Mixotrophs	Not et al., 2007
	Telonemia	29	Picograzers	Klaveness et al., 2005
Opisthokonta	Choanoflaggelida	31	Picograzers (nanoflagellates)	
	Fungi	103	Degradars & parasites	Manohar and Raghukumar, 2013
	Mesomycetozoa	20	Parasites	Marshall et al., 2008
Rhizaria	Cercozoa	94	Parasites	Chantangsi and Leander, 2010
Stramenopiles	Bacillariophyta	99	Autotrophs	Brown and Sorhannus, 2010
	Bicoecea	10	Picograzers	
	Bigyromonadea	1	Picograzers	Cavalier-Smith, 1997
	Bolidophyceae	13	Autotrophs	Brown and Sorhannus, 2010
	Chrysophyceae-Synurophyceae	12	Autotrophs & mixotrophs	Brown and Sorhannus, 2010
	Hyphochytriomyceta	1	Parasites & degraders	Kramarsky-Winter et al., 2006
	Labyrinthulea	41	Parasites	Raghukumar, 2002
	MAST	77	Picograzers (nanoflagellates)	Massana et al., 2006
	Oomycetes	35	Parasites	Park et al., 2004
	Pelagophyceae	1	Autotrophs	Brown and Sorhannus, 2010
	Pirsonia	7	Parasites	Skovgaard, 2014
Other protists		30		

The references cited present some of the most and are not exhaustive.

negative and positive correlations do not necessarily indicate co-occurrence or co-exclusion patterns, as they may result from indirect interaction with a third variable, whether biotic or

abiotic. The first analysis was done through MINE (Maximal Information-based Nonparametric Exploration) statistics by computing the Maximal Information Coefficient (MIC) between

each pair of OTUs, and between each pair and the environmental variables (Reshef et al., 2011). MIC is a non-parametric method that captures associations between data and provides a score that represents the strength of a pair relationship, independently of their function type (e.g., linear or exponential). The matrix of MIC-values corresponding to a  $P < 0.01$ , based on pre-computed  $P$ -values of various MIC scores at different sample sizes, was used (MIC > 0.6 in this case). Second, the Local Similarity correlations (LS), were determined by LS analysis (LSA; <https://bitbucket.org/charade/elsa/wiki/Home>), using a delay of one time point in order to identify time-lag correlations (Ruan et al., 2006; Steele et al., 2011; Xia et al., 2011, 2012). The correlations without any delay were also checked. The correlations with  $P < 0.001$  and  $R > 0.6$  were taken into consideration. LS analysis is optimized to detect non-linear, time-sensitive relationships and is used to construct correlation networks from time-series data, similar to the present dataset.

The networks of significant OTU-OTU pairs, or OTU-environmental parameter pairs as identified by MIC associations, and by both time-lag LS correlations and correlations without time-lag, were constructed using Cytoscape v3.5 (Smoot et al., 2011). Network Randomizer 1.1.2 (Tosadori et al., 2016) was used to generate random networks of the dataset. The “small-world” coefficient ( $\sigma$ ) of the network was calculated according to Humphries and Gurney (2008), as follows:

$$\sigma = \frac{C/C_r}{L/L_r} \quad (1)$$

where  $C$  and  $C_r$  are the observed and random Clustering Coefficients and  $L$  and  $L_r$  the observed and random Characteristic Path Lengths. If  $\sigma > 1$  ( $C \gg C_r$  and  $L \approx L_r$ ), then the network has “small world” characteristics (Watts and Strogatz, 1998; Humphries and Gurney, 2008). In an ecological context, small-world networks have specialized nodes or regions within themselves (e.g., a parasite on a certain microbe), and at the same time they can exhibit shared or distributed processing across all of the connected nodes (all connected microbes have trophic interactions, contributing to energy transfer in the community). The topological parameters of each network were calculated according to Cytoscape v3.5.

## RESULTS

### Environmental Parameters

The periods of the *P. globosa* blooms during the study period were determined based on microscopy observations and cell counts. Seawater temperature during the period of the study ranged from 5.4 to 18.5°C, while the salinity ranged from 33.3 to 34.7 (Figure 1). The highest value recorded for PO<sub>4</sub> was recorded during the *P. globosa* bloom in April 2012 (Figure 1). The highest values of inorganic nutrients were documented for all 3 years before the onset of the *P. globosa* proliferation, reaching 15.1 μM for NO<sub>3</sub> + NO<sub>2</sub> in January 2012, and 7.7 μM for SiOH<sub>4</sub>, in February 2013. Furthermore, the N/P ratio varied from 0.1 to 43.2 (coefficient of variation CV = 109%) with values higher before and lower after the *P. globosa* blooms in all 3 years (Figure 1).

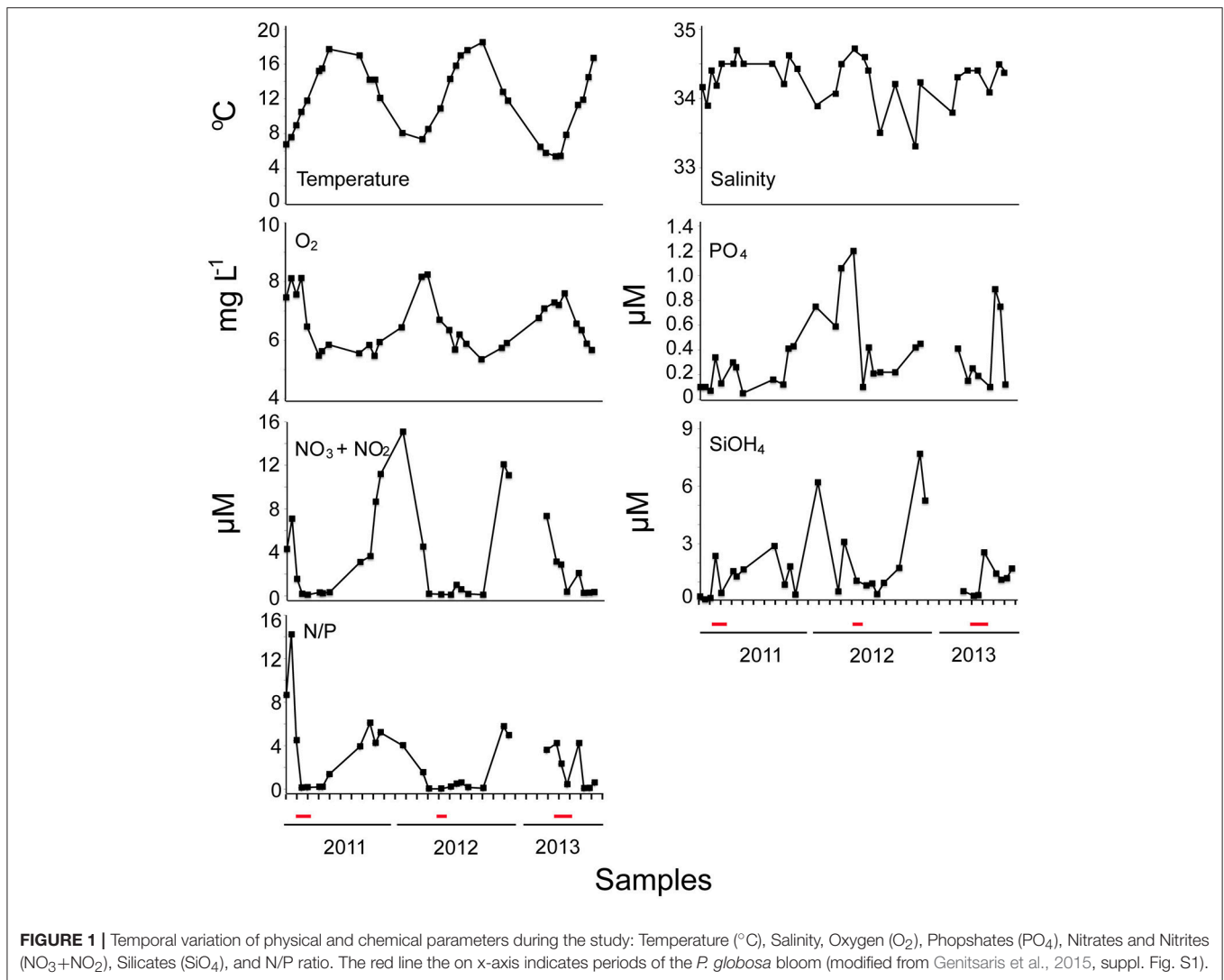
### Protistan Community and Potentially Parasitic Community

In the present study we identified 1,242 OTUs. Rarefaction curves of the samples were calculated, reaching a plateau in all cases, suggesting a good coverage of the protistan diversity in all samples (Figure S1). All OTUs were classified into 9 taxonomic super-groups, which were further sorted into 32 high-level taxonomic groups (Table 1). The microscopic counts of *P. globosa*, diatoms, and dinoflagellates during the sampling period showed a similar succession with read data (Figures 2A–C). The most abundant OTUs in the dataset belonged to the above 3 groups. Among the 99 diatom-related OTUs, an OTU identified as *Leptocylindrus danicus* (100% similarity; OTU003) showed an increase during late spring and summer (05/2011, 07/2012 and 05 to 07 2013, Figure 2B). Dinophyceae was the most diverse group consisting of 182 OTUs. Although many trophic types were represented in this group (Table 1), the majority were heterotrophic dinoflagellates (mainly phytoplankton grazers). The genus *Gyrodinium* dominated the dinoflagellate community (Figure 2C). In particular, 42 OTUs had as closest relative a dinoflagellate identified as *Gyrodinium spirale*, and 9 of them were abundant (> 0.1% of the total number of reads) and were included in the network analysis. *Gyrodinium* was present in all samples, and particularly abundant during spring and early summer coinciding with the *P. globosa* blooms (Figures 2A,C). *P. globosa* was also among the most abundant taxa and was represented by an OTU with a close relation (100% similarity). Finally, high abundance values were annotated to the potentially parasitic groups of MALV (Figure 2C), Fungi and Cercozoa throughout the study (Figure 2D).

Among the 32 higher taxonomic groups, >10 were totally composed of or included potentially parasitic, symbiotic, or decomposers' taxa (SymbDec), making up 42% of the total number of OTUs (Figure 3). Most of the potentially parasitic groups were plankton parasites, while around 30% among them were fish/molluscs parasites. SymbDec OTUs showed the highest richness in most of the samples, reaching 130 OTUs on 03 October 2012 (Figure 4A). The lowest number of SymbDec OTUs (10 OTUs) was recorded in the sample with the lowest OTU richness on 05 April 2012 (Figure 4A). In all samples, SymbDec contributed >20% of the total number of OTUs, and in more than the half samples, they contributed >30% of the total number of OTUs (Figure 4B). It is noteworthy that in one sample (04 April 2011), the SymbDec group comprised of 58% of the total number of OTUs (Figure 4B). The trophic group following SymbDec in terms of OTUs richness (reaching 80 OTUs in one sample) and contribution (generally between 10 and 30% of the total number of OTUs) were OTUs identified as picograzers (picoG, Figure 4B), while the least represented in terms of OTUs was trophic group ciliates, identified as grazers of nanoplankton (nanoG, Table 1), which contributed to about 10% of the total OTUs in most samples (Figure 4B), and in a unique case (04/07/2012) only one OTU was present.

The most diverse taxonomic group belonging to SymbDec was Syndiniales (MALV), comprising of 12% of the total number of OTUs (Figure 3, Table 1). MALVs were always present and showed several peaks (Figure 2C), the most pronounced of which



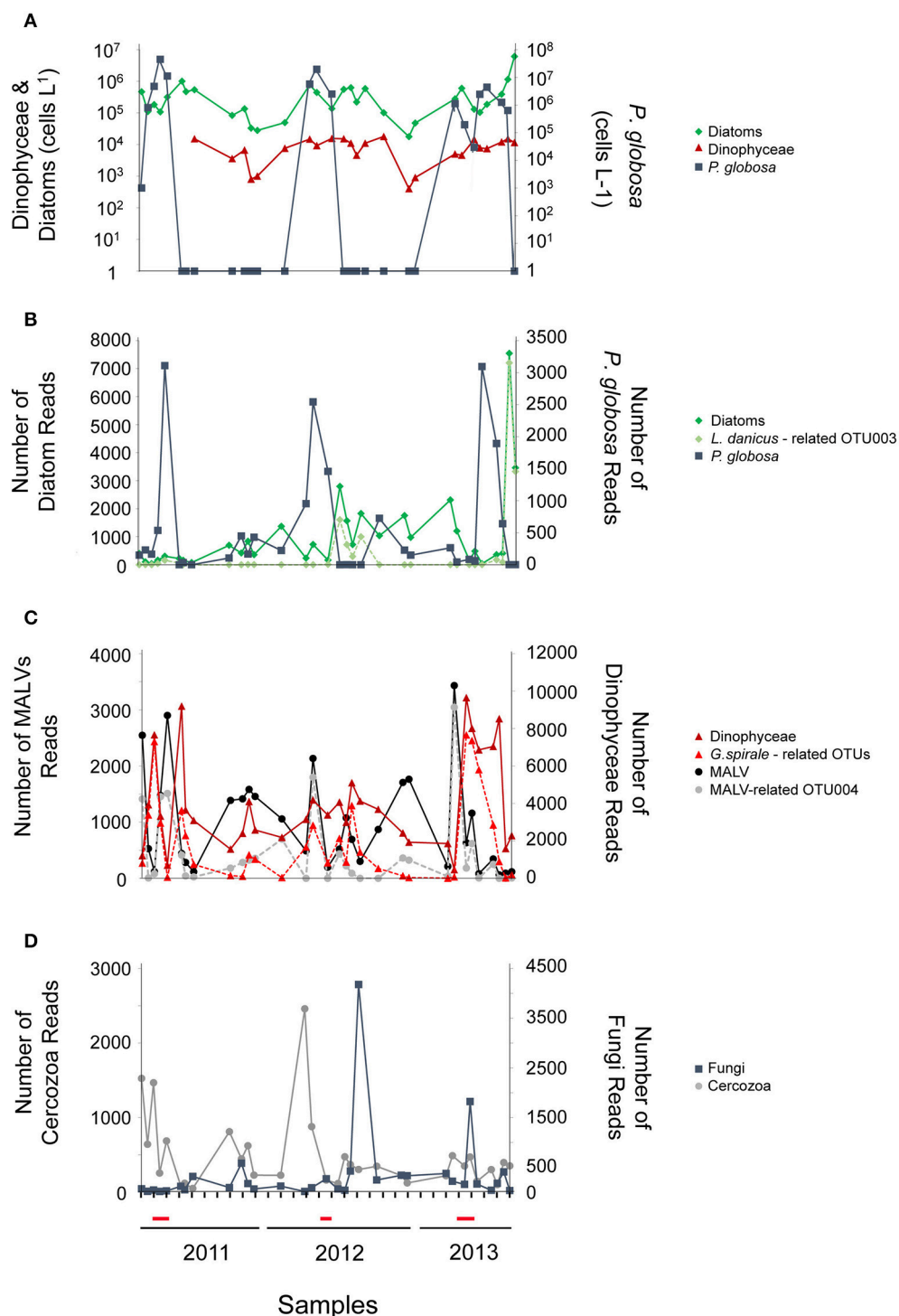


were recorded in March (2011, 2013), April (2012), and May (2011) (**Figure 2C**). These four peaks were primarily attributed to the same OTU (**Figure 2C**, OTU004 with closest relative MALV-I with 100% similarity), which accounted from 52 to 89% of the MALV community in these 4 samples and 45% of the overall MALV community. The MALVs' seasonal dynamic seemed visually opposite in many cases with the heterotrophic dinoflagellates (**Figure 2C**, e.g., March to September 2011, March-April 2012, September-December 2012, and March to May, 2013) with a significant negative correlation calculated between them ( $r = -0.49$ ,  $p < 0.0001$ ). MALVs were followed in terms of OTUs richness by Fungi (8% of the total number of OTUs), and Cercozoa (8%; **Figure 3**). The most abundant Fungi-related OTU in terms of number of reads was an Ascomycota identified as *Tritirachium* sp. (100% similarity, OTU020) which accounted for 24 and 75% of the two major fungal peaks occurred in July 2012 and April 2013, respectively (**Figure 2D**). The second most abundant fungi-related OTU was a Basidiomycota, which accounted for 62% of the fungi-related sequences in the July 2012

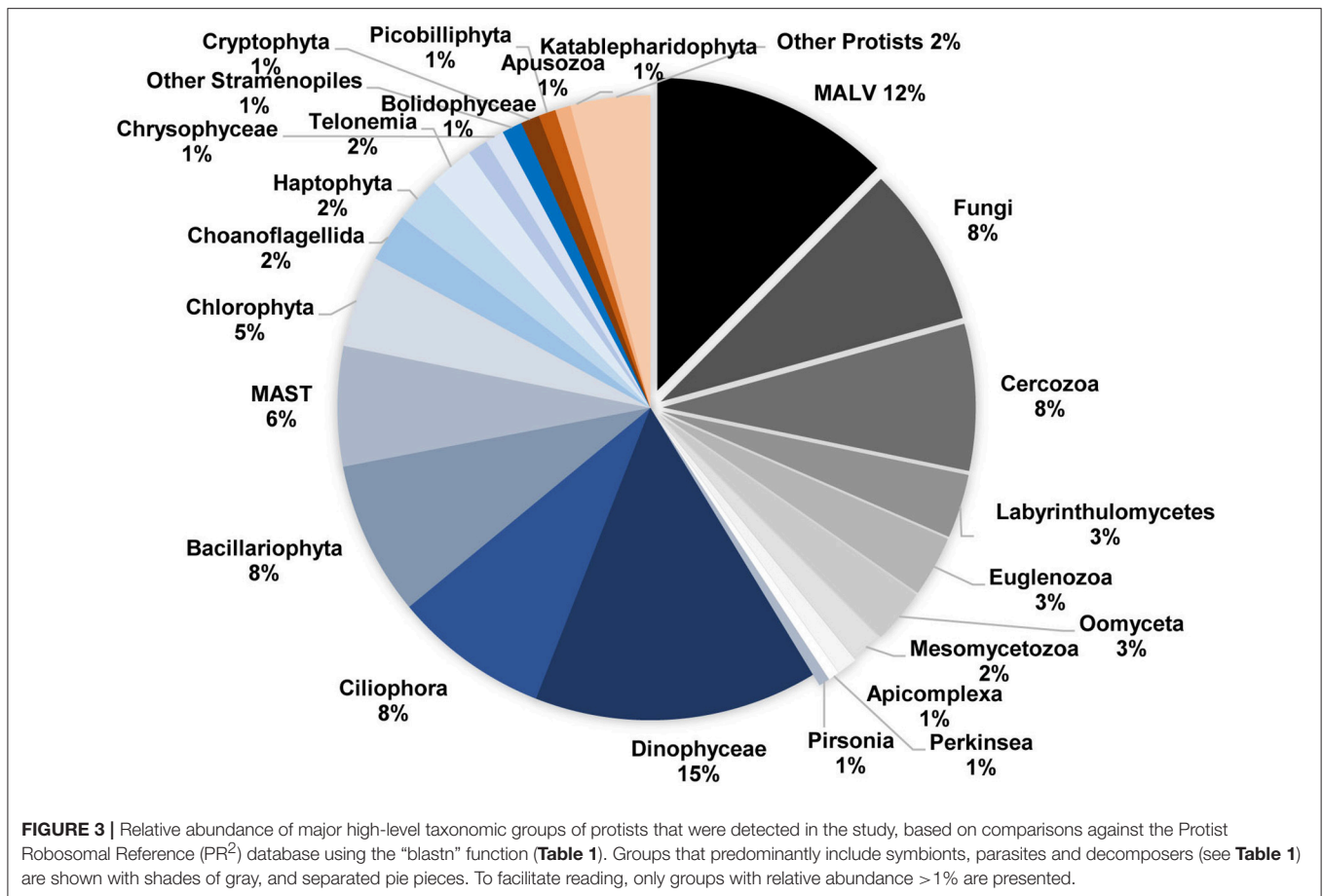
peak. Among Fungi, 24 OTUs belonged to Chytridiomycota, 3 of which were Rhizophidiales; however, chytrid-related OTUs were rare and represented <0.1% of total abundance. Cercozoa showed a pronounced peak in March 2012, which was dominated by OTUs related to an uncultured clone (OTU028; 75%) and *Cryothecomonas* sp. (OTU045; 14%). These two OTUs peaked also together in April 2011 (**Figure 2D**).

The other taxonomic groups belonging to SymbDec were Labyrinthulomycetes (3% of the total number of OTUs), Oomycetes (3%), Euglenozoa (3%), Mesomycetozoa (2%), Apicomplexa (1%), Perkinsea (1%), and Pirsonia (1%) (**Figure 3**, **Table 1**). Perkinsea, were represented by 7 OTUs including OTUs closely related to *Parviculifera* sp. and *Perkinsus* sp., and Apicomplexa were represented by 13 OTUs, including several Gregarines, but as chytrids above, they all were found at very low abundances (<0.1%).

The plot of the cumulative number of OTUs indicated that the number of newly introduced OTUs was always higher during summer and autumn for each year, and the highest number



**FIGURE 2** | Microscopic counts (cells L<sup>-1</sup>) of diatoms (green line), dinoflagellates (red line) and *Phaeocystis globosa* (blue line) **(A)**; Number of reads of diatoms (green line), the most abundant diatom OTU003 related to *Leptocylindrus danicus* (dotted light green line), and *P. globosa* (blue line) **(B)**; Number of reads of heterotrophic dinoflagellates (red line), the most abundant dinoflagellate OTUs related to *Gyrodinium spirale* (dotted light red line), MALV (black line), and the most abundant MALV-related OTU004 (dotted gray line) **(C)**; Number of reads of Fungi (blue line) and Cercozoa (gray line) **(D)**. The red line on the x-axis indicates periods of the *P. globosa* bloom.



was recorded in July 2011 (98 OTUs, Figure 5A). The influx of total OTUs accelerated from September to November 2011 and 2012, and decelerated at the onset and during the *P. globosa* bloom (Figure 5A). A parallel but hampered trend was observed for the newly introduced SymbDec OTUs, while the MALV-related newly introduced OTUs leveled out by autumn 2011. MALV data revealed the presence of 154 species-level operational taxonomic units (OTUs) falling into three of the five well-established Syndiniales clades (Table 1). Based on the newly arriving OTUs between two consecutive dates, a highly significant positive relation was observed between the cumulative number of OTUs not belonging in SymbDec and the cumulative number of SymbDec OTUs (Figure 5B).

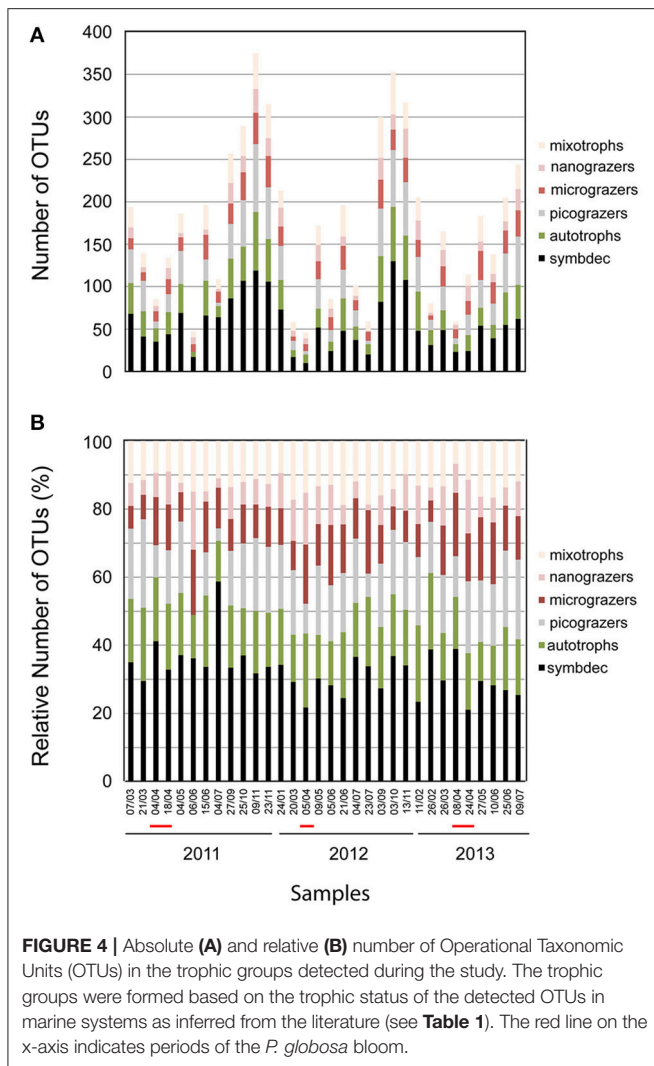
## Network Analyses

Two correlation techniques, MIC and LSA, that detect ecological and time-series relationships between 18S rRNA gene reads were selected in order to implement network analysis on the present time-series dataset.

According to the MIC, a total of 93 out of the 142 most abundant OTUs (i.e., >0.1% abundance) were found to have 254 strong correlations (edges) with each other. Fifty-seven of the correlations were negative (Table S1). The environmental parameters found to have strong correlations with OTUs were O<sub>2</sub>, N, T, and N/P (data not shown), however, only 11 OTUs were

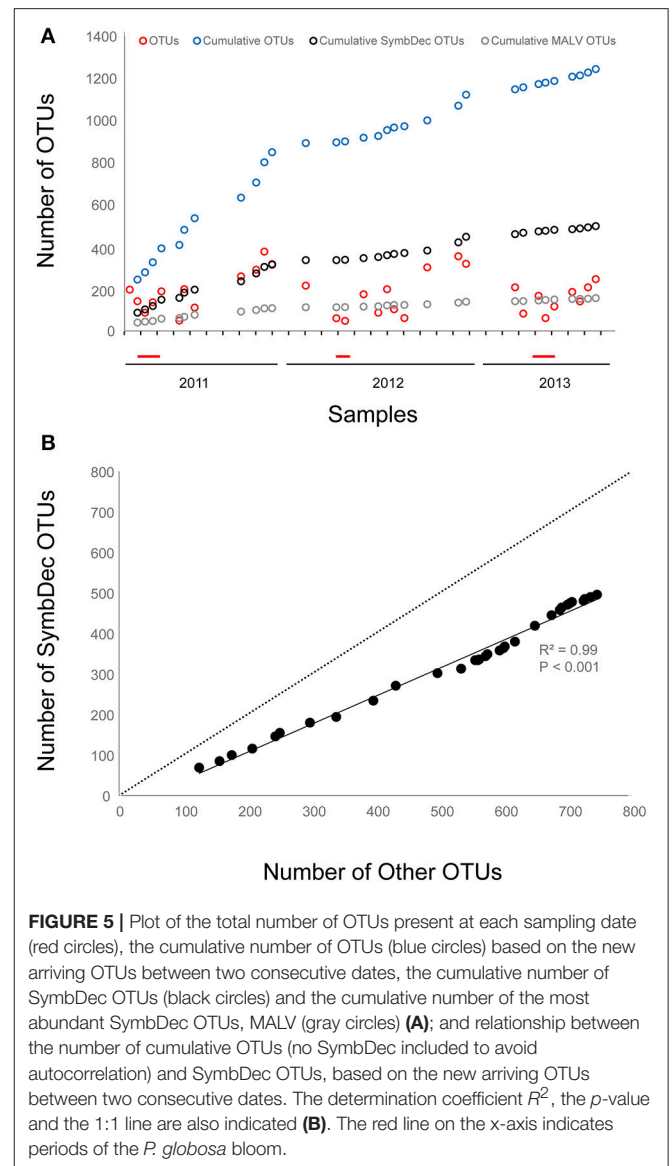
found to be connected with them. The 60 OTUs present in the MIC network when only connections of SymbDec OTUs were included were involved in 86 connections (12 negative, Table S1).

LSA with no lag showed a relatively higher level of participation of OTUs in the respective networks than MIC (Table 2). In order to explore possible correlations that could be hidden due to delayed reaction to specific changes between environmental parameters and the abundant OTUs, and OTUs of the SymbDec group and the other OTUs, a one-time point lag delay was also implemented to the LSA. In fact, the lag delayed analysis revealed a higher number of connected OTUs (103 OTUs in total, 69 when only SymbDec connections were included, and 22 connected to O<sub>2</sub>, N, T, and N/P) and significant correlations (394 total connections, 143 when only SymbDec connections were included, 28 with environmental parameters) (Figure 6A, Table 2). The topological parameters of the networks suggested that the lag-delayed analysis provided a denser network structure between the most abundant OTUs and in particular for the SymbDec OTUs. The Clustering Coefficient, the Characteristic Path Length, the Average Number of Neighbors, and the Network Density of the SymbDec lag network were all higher than in MIC and no lag networks (Table 2). Furthermore, topological parameters, such as the Clustering Coefficient, Centralization, and Heterogeneity were ~x5, x7, and x3 times higher, respectively, for the observed



relative to a random network of the same size (Table 2, see Steele et al., 2011). On the other hand, the Characteristic Path Length of the 2 networks was similar (Table 2), resulting a “small-world” coefficient  $\sigma = 5.51$ .

Here the lag-delayed LSA network of only SymbDec connections is shown in Figure 6A. Among the 26 potentially parasitic OTUs, the most connected belonged mainly to MALVs, Pirsonia and Cercozoa (Table 3), and 29 out of the 143 connections involved connections between SymDecs. The majority of the correlations involved in the network were positive, while only 3 negative correlations were detected among OTUs (Figure 6A). The rest of the negative edges (six) in the network were OTUs—environmental parameters pairs. Among the 4 environmental parameters ( $O_2$ , N, T, and N/P), temperature (T) had the highest number of connections (eight), while half of them were negative connections with MALV-related OTUs. In order to examine the significant correlations of the environmental parameters with all OTUs in addition to the parasitic-related, a network that included only the lag-delayed connections of the environmental parameters is shown



(Figure 6B). In this case, Temperature (T) was negatively connected with oxygen ( $O_2$ ), while  $O_2$  was the environmental parameter with the highest number of connections (12 overall, eight negative) with a variety of taxonomically and functionally different OTUs. Among them, six were with parasitic OTUs (four with MALVs). Furthermore, Nitrogen (N) had exclusively positive connections (10) with a variety of taxonomically and functionally different OTUs (Figure 6B). Finally, it is noteworthy that the bloom forming *P. globosa* was not present in any of the networks constructed.

## DISCUSSION

These results have showed that the importance of SymbDec OTUs in the EEC is extended throughout the whole year, as SymbDecs were found to comprise 42% of the total number

**TABLE 2** | Network statistics of the LS analysis without lag and with lag delay of 1 time point, (a) for the total community including environmental variables, (b) for the connections only with environmental variables, and (c) the connections of SymbDec.

Network Statistics	LSA (Spearman's Rank; no lag)			LSA (Spearman's Rank; lag 1 delay)			Random network
	Total	Environmental	SymbDec	Total	Environmental	SymbDec	SymbDec
Nodes	96	17	64	103	26	69	69
Edges (pairs)	282	18	118	394	28	143	143
Negative edges	13	7	4	33	9	9	–
Clustering coefficient	0.324	0.243	0.281	0.317	0.086	0.332	0.062
Connected components	8	2	5	3	1	4	1
Network diameter	8	3	6	9	5	8	7
Network radius	1	1	1	1	3	1	4
Network centralization	0.184	0.417	0.316	0.233	0.470	0.392	0.058
Shortest paths	5,758 (63%)	152 (55%)	2,142 (53%)	9,514 (90%)	650 (100%)	3,792(80%)	4,692(100%)
Characteristic path length	3.299	2	2.497	3.266	2.849	3.080	3.162
Avg. number of neighbors	5.875	2.118	3.688	7.650	2.154	4.145	4.145
Network density	0.062	0.132	0.059	0.075	0.086	0.061	0.061
Network heterogeneity	0.983	0.885	1.315	0.950	1.379	1.324	0.476

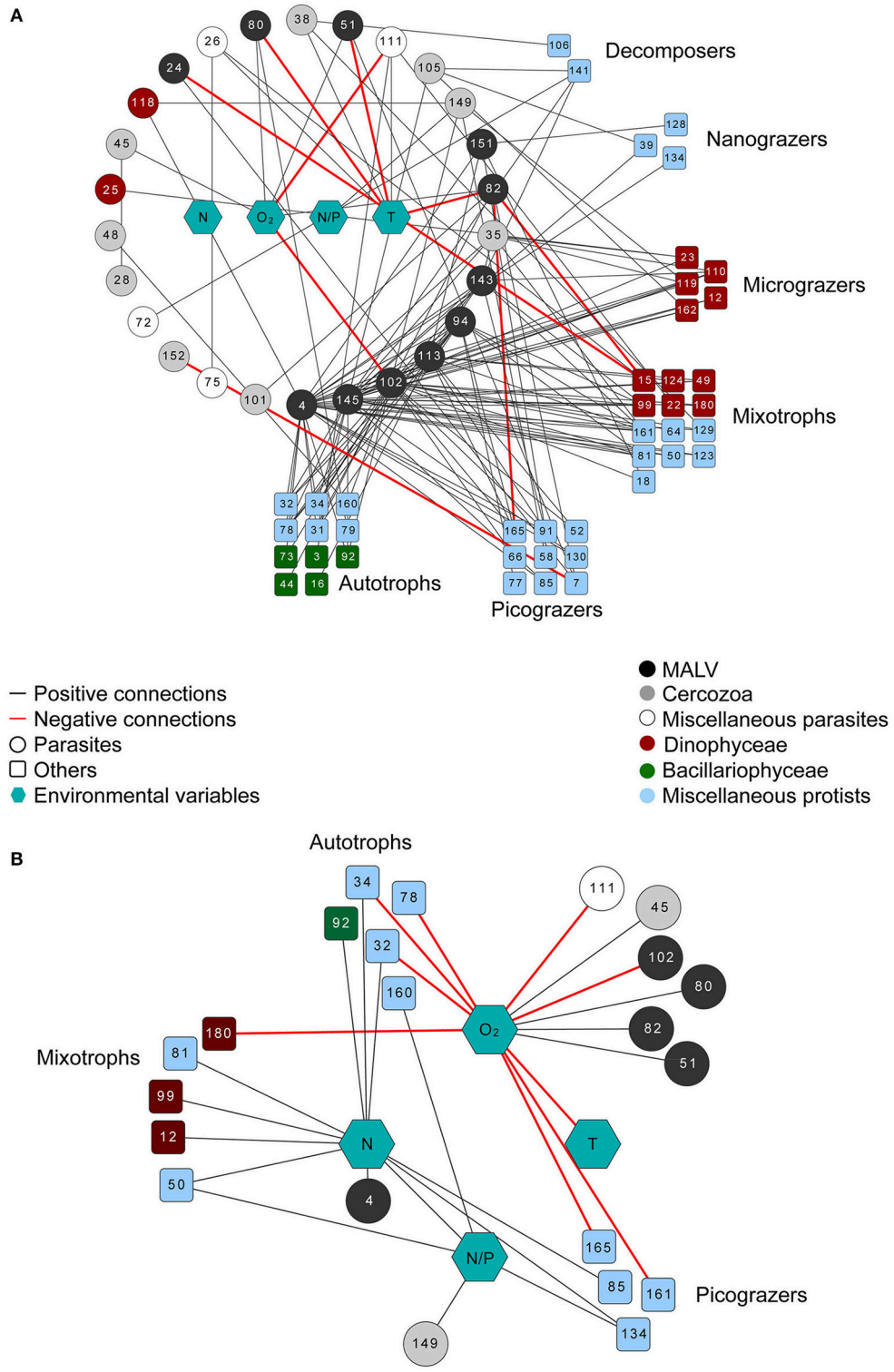
The connections with  $p < 0.001$  and Spearman's rank coefficient  $R > 0.6$  are shown. Also the topological parameters of a random network of the same size as LSA lag-delayed SymbDec network is presented. The random network was constructed using Network Randomizer 1.1.12 (Tosadori et al., 2016), and all topological parameters were calculated with Cytoscape v3.5. (Smoot et al., 2011).

of OTUs in the entire study (Figure 3), and at least 20% of the total number of OTUs per sample (Figure 4). Accumulating evidence shows that biotic interactions among eukaryotes, as well as between bacteria and certain algae, can shape the composition and dynamics of microbial communities and algal blooms (Dang and Lovell, 2016). Although sequencing studies are not quantitative, the relative abundance of reads belonging to groups connected to parasitism or symbiosis has attracted scientific attention in virtually all marine environments (e.g., Georges et al., 2014; Duret et al., 2015; Massana et al., 2015). According to the existing literature on cell-cell interactions among the sea-surface microbiota (e.g., Dang and Lovell, 2016), it is now possible to draw a more complete picture about the microbial interactions in coastal water communities. In the present study, a network analysis approach was conducted in order to explore the aptitude of correlation networks, highlight the strong connections of SymbDec taxa within the seasonal microplankton succession and identify the underlying environmental parameters that affect specific groups of the microplankton community. The detection of ecological relationships between microbial taxa is challenging for sparse datasets and filtering out the rare biosphere in microbial datasets is recommended prior to utilizing these tools (Weiss et al., 2016). The comparison of performance of eight correlation techniques commonly used to describe microbial interactions inferred from high-throughput sequence data suggested that the best tools to address sparsity of data were MIC and LSA (Weiss et al., 2016). The authors concluded that although these tools reflect relationships that involve commensalism and mutualism, they have better fidelity in reflecting parasitic relationships than the other correlation techniques tested. Furthermore, in the case of time series data, LSA best describes the ecological relations between taxa. In particular, LSA can also identify

time-lag correlations (Ruan et al., 2006; Xia et al., 2011, 2012).

For all these reasons, the network that was chosen to be presented in the paper was constructed according to LSA (Figures 6A,B). The LSA with time-lag network displayed higher connectivity than the MIC network, based on the topological parameters that show density (Clustering Coefficient, Characteristic Path Length, and Average Number of Neighbours) (Table 2, Faust and Raes, 2012). Furthermore, the higher Clustering Coefficient, Centralization, and Heterogeneity values in comparison to the clustering coefficient of a random network of the same size (Table 2) showed that the parasitic assemblage in the EEC were more connected than a random network of similar size. As the Clustering Coefficient was relatively high and the Characteristic Path Length relatively short, the parasitic network had “small world” characteristics (Watts and Strogatz, 1998; Humphries and Gurney, 2008), rendering the assemblage more robust to environmental change, but more susceptible to removal of highly associated taxa (Dunne et al., 2002). The “small world” model “of few highly connected nodes, as opposed to an even distribution of connectivity, makes the network more robust to change (Albert et al., 2000; Montoya et al., 2006) with an important caveat: if highly connected nodes are lost, the network would change dramatically”. According to Steele et al. (2011), the ‘highly connected nodes’ can be considered as equivalent to microbial “keystone species.”

However, it is important to consider here that network structure alone is incapable of capturing the processes that occur within a dynamic community of microbes. For instance, the presence of hub nodes could indeed confer greater robustness or resilience on a community, but considerations of bottom-up disturbances, i.e., disturbances that propagate in the direction of energy transfer, need also to be taken into account. Also, on



**FIGURE 6 |** Network diagram of highly significant connections ( $R > 0.6$  and  $p < 0.001$ ) between the SymbDec Operational Taxonomic Units (OTUs) and other dominant OTUs, as well as the environmental variables (nodes) as determined by LS analysis (LSA) using a delay of one time point. To facilitate reading all the potentially parasitic nodes are structured as a circle which can be read anticlockwise starting from the OTU4, which has the maximum number of connections (**A**); Network diagram of highly significant connections ( $R > 0.6$  and  $p < 0.001$ ) between environmental parameters and dominant OTUs (nodes) as determined by LS analysis using a delay of one time point (**B**). The exagons represent the environmental variables, the circles represent parasitic taxa, and the rectangles all other taxa. Different colors: different taxonomic groups with emphasis on Dinophyceae (red) and Bacillariophyceae (green). Black lines (edges) represent positive connections and red lines represent negative connections. See also **Table 3** for taxonomic details of the OTUs presented on this figure.

**TABLE 3** | Putative affiliation, closest relatives according to PR2 database BLAST results, and connected OTUs based on LSA log correlation coefficient of the OTUs included in the network.

OTUs	Putative Affiliation	Closest Relative (% similarity) [Isolation Source]	Number of Connections	Connected OTUs
<b>PARASITIC OTUS</b>				
OTU004	MALV-II	Uncultured clone (99%) [Mediterranean Sea]	30	OTU012, OTU031, OTU032, OTU034, OTU035, OTU049, OTU050, OTU073, OTU078, OTU081, OTU085, OTU091, OTU092, OTU094, OTU099, OTU102, OTU110, OTU113, OTU118, OTU119, OTU123, OTU129, OTU130, OTU134, OTU143, OTU145, OTU151, OTU161, OTU162, OTU165
OTU145	MALV-II	Uncultured clone (99%) [Mediterranean Sea]	27	OTU012, OTU023, OTU031, OTU032, OTU034, OTU035, OTU039, OTU049, OTU050, OTU078, OTU081, OTU085, OTU004, OTU092, OTU094, OTU099, OTU102, OTU105, OTU110, OTU113, OTU123, OTU130, OTU141, OTU143, OTU161, OTU162, OTU180
OTU102	MALV-III	Uncultured clone (100%) [Ocean water]	23	OTU018, OTU031, OTU032, OTU034, OTU052, OTU064, OTU073, OTU078, OTU081, OTU004, OTU091, OTU094, OTU099, OTU110, OTU113, OTU129, OTU143, OTU145, OTU160, OTU161, OTU165, OTU180
OTU113	MALV-II	Uncultured clone (93%) [Coastal water]	14	OTU016, OTU049, OTU052, OTU004, OTU091, OTU094, OTU102, OTU110, OTU124, OTU143, OTU145, OTU151, OTU161, OTU165
OTU094	MALV-II	Uncultured clone (100%) [Mediterranean Sea]	13	OTU031, OTU032, OTU034, OTU077, OTU078, OTU081, OTU004, OTU099, OTU102, OTU113, OTU145, OTU161, OTU165
OTU143	MALV-II	Uncultured clone (97%) [Sargasso Sea]	11	OTU035, OTU078, OTU004, OTU091, OTU099, OTU102, OTU110, OTU113, OTU141, OTU145, OTU161
OTU035	Pirsonia	<i>Pirsonia guinardiae</i> (100%) [North Sea]	7	OTU023, OTU004, OTU110, OTU119, OTU143, OTU145, OTU161
OTU082	MALV-I	Uncultured clone (99%) [Unknown]	5	OTU007, OTU015, OTU044, OTU101, OTU165
OTU151	MALV-II	Uncultured clone (100%) [Sargasso Sea]	5	OTU018, OTU004, OTU091, OTU113, OTU128
OTU149	Cercozoa	<i>Cryothecomonas</i> sp. (100%) [Anoxic sediment]	4	OTU058, OTU092, OTU118 OTU119
OTU105	Cercozoa	Uncultured clone (87%) [Rhizosphere]	4	OTU039, OTU141, OTU145, OTU162
OTU111	Labyrinthulea	Uncultured clone (91%) [Paeonic Bay]	2	OTU003, OTU022
OTU038	Cercozoa	<i>Cryothecomonas aestivalis</i> (100%) [Unknown]	2	OTU106, OTU129
OTU026	Euglenozoa	<i>Ichthyobodo</i> sp. (89%) [Gills]	2	OTU015, OTU075
OTU024	MALV-II	Uncultured clone (98%) [Pacific Ocean]	2	OTU015, OTU007
OTU118	Dinophyceae	<i>Piscinodinium</i> sp. (91%) [Culture]	2	OTU004, OTU149
OTU051	MALV-II	Uncultured clone (100%) [Mediterranean Sea]	1	OTU007
OTU080	MALV-I	Uncultured clone (98%) [Mariager Fiord]	1	OTU079
OTU045	Cercozoa	<i>Cryothecomonas</i> sp. (100%) [Arctic waters]	1	OTU028
OTU025	Dinophyceae	<i>Blastodinium galatheanum</i> (97%) [Host]	1	OTU023
OTU048	Cercozoa	Uncultured clone (100%) [Arctic waters]	1	OTU160
OTU028	Cercozoa	Uncultured clone (100%) [Sediment]	1	OTU045
OTU072	Oomyeeta	Uncultured clone (98%) [Coastal surface water]	1	OTU141
OTU152	Cercozoa	Uncultured clone (99%) [Sediment]	1	OTU007
OTU075	Euglenozoa	<i>Ichthyobodo</i> sp. (89%) [Sediment]	1	OTU026
OTU101	Cercozoa	<i>Cryothecomonas</i> sp. (100%) [Arctic waters]	1	OTU082
<b>DECOMPOSERS</b>				
OTU106	Fungi	<i>Coccoldinium bartschii</i> (100%) [Culture]		
OTU141	Fungi	<i>Sporisorium relianum</i> (100%) [Culture]		
<b>NANOGRAZERS</b>				
OTU039	Ciliophora	Uncultured clone (99%) [Sargasso Sea]		

(Continued)

TABLE 3 | Continued

OTUs	Putative Affiliation	Closest Relative (% similarity) [Isolation Source]	Number of Connected OTUs	Connections
OTU128	Ciliophora	Uncultured clone (100%) [Arctic Sea]		
OTU134	Telonemia	Uncultured clone (100%) [Marine]		
<b>MICROGRAZERS</b>				
OTU012	Dinophyceae	<i>Warnowia</i> sp. (97%) [Culture]		
OTU023	Dinophyceae	<i>Lessardia elongata</i> (94%) [Culture]		
OTU110	Dinophyceae	<i>Gyrodinium spirale</i> (96%) [Otaru Bay]		
OTU119	Dinophyceae	<i>Gyrodinium rubrum</i> (91%) [Otaru Bay]		
OTU162	Dinophyceae	<i>Polykrikos kofoidii</i> (99%) [Culture]		
<b>MIXOTROPHS</b>				
OTU015	Dinophyceae	Uncultured clone (95%) [Seawater]		
OTU018	Prymnesiophyceae	<i>Chrysochromulina strobilus</i> (100%) [Marine]		
OTU022	Dinophyceae	<i>Warnowia</i> sp. (95%) [Culture]		
OTU049	Dinophyceae	<i>Gymnodinium aureolum</i> (94%) [Culture]		
OTU050	Picobiliphyta	Uncultured clone (99%) [English Channel]		
OTU064	Prymnesiophyceae	<i>Prymnesium kappa</i> (100%) [Culture]		
OTU081	Cryptophyta	<i>Teleaulax amphioxeia</i> (100%) [Culture]		
OTU099	Dinophyceae	Uncultured clone (93%) [Seawater]		
OTU123	Picobiliphyta	Uncultured clone (100%) [English Channel]		
OTU124	Dinophyceae	<i>Gymnodinium aureolum</i> (92%) [Culture]		
OTU129	Prymnesiophyceae	<i>Chrysochromulina</i> sp. (100%) [Seawater]		
OTU161	Picobiliphyta	Uncultured clone (100%) [English Channel]		
OTU180	Dinophyceae	<i>Prorocentrum micans</i> (96%) [Jiaozhou Bay]		
<b>PICOGRAZERS</b>				
OTU007	MAST-1	Uncultured clone (100%) [Ocean water]		
OTU052	MAST-4	Uncultured clone (100%) [Pacific Ocean]		
OTU058	MAST-6	Uncultured clone (96%) [Ocean water]		
OTU066	MAST-3	Uncultured clone (94%) [Sargasso Sea]		
OTU077	MAST-3	Uncultured clone (96%) [Sargasso Sea]		
OTU085	MAST-4	Uncultured clone (100%) [Coastal water]		
OTU091	MAST-3	Uncultured clone (100%) [Coastal water]		
OTU130	MAST-4	Uncultured clone (97%) [Sargasso Sea]		
OTU165	MAST-12	Uncultured clone (96%) [Lake Namako-ike]		
<b>AUTOTROPHS</b>				
OTU003	Bacillariophyceae	<i>Leptocyindrus danicus</i> (98%) [Culture]		
OTU016	Bacillariophyceae	<i>Rhizosolenia setigera</i> (99%) [Culture]		
OTU031	Chlorophyta	<i>Ostreococcus</i> sp. (100%) [Coastal water]		
OTU032	Chlorophyta	<i>Micromonas</i> sp. (100%) [Marine]		

(Continued)



TABLE 3 | Continued

OTUs	Putative Affiliation	Closest Relative (% similarity) [Isolation Source]	Number of Connected OTUs Connections
<b>OTU034</b>	Chlorophyta	<i>Bathycoccus prasinos</i> (100%) [Coastal water]	
<b>OTU044</b>	Bacillariophyceae	<i>Minidiscus trioculatus</i> (100%) [Culture]	
<b>OTU073</b>	Bacillariophyceae	<i>Detonula pumila</i> (98%) [Culture]	
<b>OTU078</b>	Chlorophyta	<i>Micromonas pusilla</i> (100%) [Marine]	
<b>OTU079</b>	Bolidophyceae	Uncultured clone (91%) [South Atlantic]	
<b>OTU092</b>	Bacillariophyceae	<i>Chaetoceros rostratus</i> (97%) [Culture]	
<b>OTU160</b>	Bolidophyceae	<i>Bolidomonas pacifica</i> (98%) [Ocean waters]	

OTUs coding colors represent different taxonomic groups according to network coloring (see **Figure 5A**). Black represents MALV, dark gray Cercozoa, light gray miscellaneous parasites, red Dinophyceae, green Bacillariophyceae and blue miscellaneous protists.

a complementary note, the removal of nodes at higher trophic levels may not affect nodes at lower levels, and in this case it is challenging to capture any top-down effects or cascades. While this can be done in a network context, one has to add dynamic processes to the links, and go beyond considering network structure alone. For these reasons, in the present study we attempted to interpret network structure keeping in mind the community dynamics as revealed by both NGS and microscopic data, and including existing knowledge from previous studies of the plankton community in the area.

In the parasitic network of the present study, several MALVs and Cercozoa exhibit characteristics of “hubs,” as both taxonomic groups contribute with highly connected taxa, and also include few taxa with low number of connections that are traced back to the highly connected taxa of their respective groups. Many different parasites were connected; for example the most abundant cercozoan (OTU028), which showed a pronounced peak, was connected to the second most abundant cercozoan closely related to *Cryothecomonas* (OTU045). Cercozoans showed connections with fungi and between them and in one case, with a *Chaetoceros*-like diatom. According to Steele et al. (2011) several of the co-occurrences may indicate associations of taxa with analogous or matching functions, and others co-occur because they share similar optimal environmental conditions. Recent findings suggest that highly connected networks promote the stability of the community under changing environments, possibly because of direct and indirect mutualistic support (Lever et al., 2014). In other words, when harsh conditions drive the disappearance of a central highly connected network taxon, another one of the same group directly connected to the first, might survive or even thrive and take its place as if parasites were “holding the ecosystem together” (Windsor, 1998). This may be part of the explanation for the frequent and persistent occurrence of the parasitic guild, and especially of the MALVs and Cercozoa, throughout the year in the EEC, under a variety of conditions (Christaki et al., 2014; this study), based on network structure. Indeed, among SymbDec groups, MALVs show overwhelming occurrences representing up to 50% of sequences retrieved, from coastal waters to deep hydrothermal vents (Guillou et al., 2008).

As previously observed in the studied area (Christaki et al., 2014), MALVs were identified as the predominant group of potential parasites. While the influx of the total number of OTUs continued to increase after 32 sampling dates, the cumulative plot of newly arriving MALV OTUs leveled out before the end of the first year of sampling (see **Figure 5**). This result may point toward a wide host range of this group. It has been suggested that MALV groups, especially MALV-II, because of their considerable abundance and diversity in marine environments, interact mainly with dinoflagellates and also with other various hosts that include radiolarians, ciliates, as well as copepods and fish (Skovgaard et al., 2005; Harada et al., 2007; Guillou et al., 2008; Bråte et al., 2012). The relationship between MALVs and heterotrophic dinoflagellates in the present study ( $P < 0.001$ ) suggests that most of these MALVs were dinoflagellate related ones. Also, according to network analysis, five out of the six most connected MALV-related

OTUs (closely affiliated to MALV-II) were mostly connected with OTUs affiliated to dinoflagellates (e.g., *Gyrodinium* spp., *Warnowia* sp., *Lessardia elongata*). Planktonic dinoflagellates in general and the ones of the genus *Gyrodinium* in particular are infected by Syndiniales (e.g., Coats et al., 1996; Coats, 1999; Park et al., 2004). The genus *Gyrodinium* was represented by 42 OTUs of which 9 were abundant (>0.1%), and was always present in the samples and particularly abundant during spring and summer. Although dinoflagellate cells have been found to contain multiple 18S rRNA gene copies, thus potentially affecting abundance data by oversaturating the sequencing samples (Stoeck et al., 2010; Pirreda et al., 2016), our microscope observations of the samples confirmed the overall dominant presence of *Gyrodinium* species. To date, the phenomenon of massive killing of dinoflagellates by MALV (e.g., Coats et al., 1996; Park et al., 2004; Guillou et al., 2008) has been the subject of scientific interest as a possible means to limit toxic dinoflagellate blooms (Chambouvet et al., 2008). In addition, according to the overall connections of MALV-related OTUs as revealed by network analysis (Figure 6A), a larger host range and variability than previously thought was indicated, as MALVs are highly connected to diatoms (e.g., *Rhizosolenia setigera*, *Minidiscus* sp.), ciliates (*Strombidium* spp.), MAST groups, and taxa belonging to Picobiliphyta. The most abundant MALV OTU, which was also mainly responsible for the four most important MALV peaks during the study (Figure 2C), was the most highly connected; with connections to a variety of organisms, including MAST groups, diatoms and a small number of dinoflagellates.

MALV were followed in richness and abundance in the SymbDec assemblage, by Fungi and Cercozoa, which showed several peaks but did not show any particular seasonal dynamic (Figure 2D). Fungi and Cercozoa were represented by 103 and 94 OTUs, respectively. In a previous study carried out in the EEC, these groups were mostly connected with diatoms (Christaki et al., 2014). In another study in contrasting Southern Ocean waters, Fungi and Cercozoa were associated with high diatom biomasses present in iron-fertilized waters; but almost absent in the High Nutrient Low Chlorophyll waters dominated by small sized phytoplankton (Georges et al., 2014). Relative to sequence abundances the two highest fungal peaks in July 2012 and April 2013 were mainly attributed to one Basidiomycetes (OTU021) and one Ascomycota (OTU020) OTU respectively. Their peaks followed diatom declines in July 2012 and March 2013, observed in both microscopy and sequencing data (Figures 2A,B,D), suggesting that they were related to phytoplankton exudates such as polysaccharides which are known to promote the growth of Fungi (Myklestad, 1995; Kimura and Naganuma, 2001; and references therein). Still, at this level of taxonomic identification and the lack of any knowledge of their ecology does not allow us to make any further hypothesis about their role. For example, many members of zoosporic fungi are hyperparasites having complex tripartite interactions and are consequently difficult to monitor (Gleason et al., 2014). Among Fungi, Chytridiomycota gain more and more attention since a large number of phytoplankton and invertebrate taxa are affected by chytrid parasites. Chytrids play also an important role in the

food web since they recycle dead organic matter which can be further used by other organisms (see Gleason et al., 2011 for a review). In this study, chytrids accounted for about the 25% of fungal OTUs (24 OTUs) but, as also observed before at the same station their number of sequences were low and without apparent relation to the phytoplankton blooms in the area (Monchy et al., 2012; Christaki et al., 2014). Most of the OTUs found in the data set are currently reported as freshwater and soil inhabiting chytrids and it cannot be excluded that they arrived from rain or freshwater inputs at this coastal station. However, four OTUs were affiliated to the genus *Rhizophyidium*, which is reported to include parasites of chlorophytes and diatoms ((Gleason et al., 2011) and references therein) but they were also low in abundance (< 0.1%) and no relation with other OTUs could be established here.

Cercozoa showed a pronounced peak in March 2012, which was dominated by an uncultured clone (OTU028; 75%) and a *Cryothecomonas*-related OTU (OTU045 14%). These two OTUs peaked also together in April 2011 (Figure 2D). Cercozoa are parasites of a variety of marine organisms, including diatoms (e.g., Tillmann et al., 1999; Schnepf and Kühn, 2000), for example *Cryothecomonas* sp. has been identified as a parasite of the diatom *Guinardia delicatula* (Drebes et al., 1996).

The same pattern of large peaks attributed almost exclusively to one distinct MALV, Cercozoa or Fungi OTU has been previously observed at the EEC SOMLIT station (Christaki et al., 2014). We compared these OTUs sampled from February to June 2009 and the ones found here and concluded that the MALV (OTU004), the Ascomycota (OTU020) and the Cercozoa (OTU045) related OTUs were identical or almost identical (Table S2) with the ones that were found to peak in the 2009 samples (Christaki et al., 2014). These OTUs are most likely persistent and important members of the SymbDec community and they deserve further attention.

From the 99 diatom OTUs identified in the present data set, 12 were considered as abundant (>0.1%) and only two OTUs affiliated to *L. danicus* and *Rhizosolenia* sp. showed significant connections (Figure 6A, Table 3). During a 3 year microscopic survey (from May 2011 to June 2014, 47 samples) at a neighboring EEC station *L. danicus*, *Rhizosolenia* sp., and *G. delicatula* were found in 51%, 74% and 70% of the samples (Breton et al., 2017). *L. danicus* showed the highest mean biomass of all the 58 diatoms identified by microscopy and was followed in biomass by *Rhizosolenia* sp. and *G. delicatula* (Breton et al., 2017). In the present data set a Labyrinthulea-related OTU (OTU111) was connected with the *L. danicus*-related OTU (OTU003). Labyrinthulomycetes are common osmoheterotrophic marine unicellular eukaryotes (López-García et al., 2001) that may have a variety of relationships with their hosts including parasitism, commensalism and mutualism. OTU111 showed several peaks in spring, winter and summer; the two highest peaks occurred simultaneously with higher numbers of *L. danicus* sequences in July 2012 and July 2013 (Labyrinthulea data not drawn).

The studied area is subject to *P. globosa* proliferation, which may account up to 90% of the algal biomass (Grattepanche et al., 2011a,b) but there is no record up to date of any harmful bloom. Microscopy and experimental data at the same site

have recognized certain heterotrophic dinoflagellates and in particular those of the genus *Gyrodinium* as the main predators of diatoms and *P. globosa* colonies (Grattepanche et al., 2011a,b). These studies proposed that because copepods do not consume *P. globosa* colonies, they intensively graze on heterotrophic dinoflagellates (top-down control) and in this way they promote the buildup of *P. globosa* biomass. Based on the results of the present study, parasitism on potential *P. globosa* grazers such as *Gyrodinium* taxa and competitors, might further promote *P. globosa* biomass buildup. Furthermore, the *P. globosa*-related OTU, albeit being one of the 5 most abundant OTUs in the entire dataset, did not show any significant correlation with parasitic or/and other, possibly indicating no parasitic or antagonistic regulation. Also, at the best of our knowledge there have been no reports about eukaryotic *P. globosa* parasites. However, it is well-documented that viruses play an important ecological role in population regulation of free *P. globosa* cells (Brussard et al., 2007). The presence of a “skin,” which effectively protects from grazing and infection by pathogens, is most likely the key reason why *P. globosa* colonies have reduced mortality relatively to singular cells and can create large blooms in various areas of the world’s oceans (Hamm et al., 1999). The combination of the dominance of *P. globosa* and its absence of connections with parasitic taxa is intriguing. Moreover, in comparison to pre-bloom and post-bloom periods, the lower rate of newly-introduced OTUs was observed during all three *P. globosa* blooms (Figure 5).

Finally, network inference can also incorporate measurements of environmental traits by assuming they are additional taxa (e.g., Steele et al., 2011). All the Nitrogen connections were positive, while those of O<sub>2</sub> were mostly negative due to highest N concentrations being before and lowest O<sub>2</sub> concentrations being during the phytoplankton bloom, respectively (Figure 1). In comprehensive studies, strong associations between microbes and concentrations of nutrients such as nitrites and nitrates, can identify taxa that contribute in biogeochemical cycling (Falkowski et al., 2008; Fuhrman, 2009). In the EEC the light increase and the limitation in nutrients during early spring, was associated to an increase of ‘functional dominance’ of phytoplankton (diatoms and *P. globosa*), promoting nitrogen use efficiency (Breton et al., 2017). However, this was not clearly reflected in the network connections as *P. globosa* was virtually missing and few connections of nutrients concerned diatoms. Not conclusive network connections allowing predictions of relationships between taxa and environment have also been reported in Steele et al. (2011).

## REFERENCES

- Albert, R., Jeong, H., and Barabási, A. L. (2000). Error and attack tolerance of complex networks. *Nature* 406, 378–382. doi: 10.1038/35019019
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2
- Amann, R. L., Blinder, B. J., Olson, R. J., Chrisholm, S. W., Devereux, R., and Stahl, D. A. (1990). Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. *Appl. Environ. Microbiol.* 56, 1919–1925.
- Aminot, A., and Chaussepied, M. (1983). *Manuel des Analyses Chimiques en Milieu Marin*. Brest: CNEXO.
- Aminot, A., and Kerouel, R. (2004). Dissolved organic carbon, nitrogen, and phosphorus in the N-E Atlantic and N-W Mediterranean with particular reference to non-refractory fractions and degradation. *Deep Sea Res. I* 51, 1975–1999. doi: 10.1016/j.dsr.2004.07.016

Generally, significant relations between microbial taxa dominated the network, while significant relations between microbes and abiotic variables were scarce (Figure 6) suggesting that changes in community composition were determined more by biological relationships than by major changes in the physical and chemical environment (Steele et al., 2011; Genitsaris et al., 2015). However, since correlations are not necessarily causative agents of the observed biological relations, network data connections among a variety of different taxa can be interpreted by different, but possibly interrelating, explanations. Because of the sampling frequency applied here, the picture consists of a series of snapshots giving a broad idea of extremely complex microbial interactions. It is encouraging though that many of the results of this study support previous evidence and naturally lead to some new hypotheses that deserve further research. For example future studies may aim at investigating if *P. globosa* colonies are indeed free from parasites; which is the role of the MALV taxa in many marine systems in the C flux, and if it can favor phytoplankton accumulation through dinoflagellate control.

## AUTHOR CONTRIBUTIONS

UC: study conception and design, acquisition of data, analysis and interpretation of data, drafting of manuscript; SG: acquisition of data, analysis and interpretation of data, drafting of manuscript; EB: acquisition of data, analysis of data; SM: analysis of data; LL, SR: critical revision. TS-N: study conception and design.

## ACKNOWLEDGMENTS

This study was supported by the “Nord-Pas de Calais” FRB-DEMO (FRB 2013), the ANR-ROME (ANR 12 BSV7 0019) and the SOMLIT network. The authors thank Eric Lecuyer for his help during the sampling. We thank [www.englisheditor.webs.com](http://www.englisheditor.webs.com) for English proofing. We are thankful to three reviewers for their constructive comments that have helped improve the original manuscript.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmars.2017.00416/full#supplementary-material>

- Bachy, C., Dolan, J. R., López-García, P., Deschamps, P., and Moreira, D. (2013). Accuracy of protest diversity assessments: morphology compared with cloning and direct pyrosequencing of the 18S rRNA genes and ITS regions using the conspicuous tintinnid ciliates as a case study. *ISME J.* 7, 244–255. doi: 10.1038/ismej.2012.106
- Behnke, A., Engel, M., Christen, R., Nebel, M., Klein, R. R., and Stoeck, T. (2011). Depicting more accurate pictures of protistan community complexity using pyrosequencing of hypervariable SSU rRNA gene regions. *Environ. Microbiol.* 13, 340–349. doi: 10.1111/j.1462-2920.2010.02332.x
- Berry, D., and Widder, S. (2014). Deciphering microbial interactions and detecting keystone species with co-occurrence networks. *Front. Microbiol.* 5:219. doi: 10.3389/fmicb.2014.00219
- Bik, H. M., Porazinska, D. L., Creer, S., Caporaso, J. G., Knight, R., and Thomas, W. K. (2012). Sequencing our way towards understanding global eukaryotic biodiversity. *Trends Ecol. Evol.* 27, 233–243. doi: 10.1016/j.tree.2011.11.010
- Bråte, J., Krabberot, A. K., Dolven, J. K., Ose, R. F., Kristensen, T., Bjorklund, K. R., et al. (2012). Radiolaria associated with large diversity of marine alveolates. *Protist* 163, 767–777. doi: 10.1016/j.protis.2012.04.004
- Breton, E., Christaki, U., Bonato, S., Didry, M., and Artigas, L. F. (2017). Phytoplankton trait variation and nitrogen use efficiency in temperate coastal phytoplankton. *Mar. Ecol. Prog. Ser.* 563, 35–49. doi: 10.3354/meps11974
- Brown, J. W., and Sorhannus, U. (2010). A molecular genetic timescale for the diversification of autotrophic stramenopiles (Ochrophyta): substantive underestimation of putative fossil ages. *PLoS ONE* 5:e12759. doi: 10.1371/journal.pone.0012759
- Brussard, C. P. D., Bratbak, G., Baudoux, A. C., and Ruardij, G. (2007). *Phaeocystis* and its interaction with viruses. *Biogeochemistry* 83, 201–215. doi: 10.1007/s10533-007-9096-0
- Burki, F., Inagaki, Y., Bråte, J., Archibald, J. M., Keeling, P. J., Cavalier-Smith, T., et al. (2009). Large-scale phylogenomic analyses reveal that two enigmatic protist lineages, telonemia and centroheliozoa, are related to photosynthetic chromalveolates. *Genome Bio. Evol.* 1, 231–238. doi: 10.1093/gbe/evp022
- Caron, D. A., Worden, A. Z., Countway, P. D., Demir, E., and Heidelberg, K. B. (2009). Protists are microbes too: a perspective. *ISME J.* 3, 4–12. doi: 10.1038/ismej.2008.101
- Cavalier-Smith, T. (1997). Sagenista and bigyra, two phyla of heterotrophic heterokont chromists. *Arch. Protistenkd.* 148, 253–267. doi: 10.1016/S0003-9365(97)80006-1
- Cavalier-Smith, T. (2002). The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa. *Int. J. Syst. Evol. Micr.* 52, 297–354. doi: 10.1099/00207713-52-2-297
- Chambouvet, A., Morin, P., Marie, D., and Guillou, L. (2008). Control of toxic marine dinoflagellate blooms by serial parasitic killers. *Science* 322, 1254–1257. doi: 10.1126/science.1164387
- Chantangsi, C., and Leander, B. S. (2010). An SSU rDNA barcoding approach to the diversity of marine interstitial cercozoans, including descriptions of four novel genera and nine novel species. *Int. J. Syst. Evol. Micr.* 60, 1962–1977. doi: 10.1099/ijss.0.013888-0
- Christaki, U., Kormas, K. A., Genitsaris, S., Georges, C., Sime-Ngando, T., Viscogliosi, E., et al. (2014). Winter-summer succession of unicellular eukaryotes in a meso-eutrophic coastal system. *Microb. Ecol.* 67, 13–23. doi: 10.1007/s00248-013-0290-4
- Coats, D. W. (1999). Parasitic life styles of marine dinoflagellates. *J. Eukaryot. Microbiol.* 46, 402–409. doi: 10.1111/j.1550-7408.1999.tb04620.x
- Coats, D. W., Adam, E. J., Gallegos, C. L., and Hedrick, S. (1996). Parasitism of photosynthetic dinoflagellates in a shallow subestuary of Chesapeake Bay, USA. *Aquat. Microb. Ecol.* 11, 1–9. doi: 10.3354/ame011001
- Dang, H., and Lovell, C. R. (2016). Microbial surface colonization and biofilm development in marine environments. *Microbiol. Mol. Biol. R.* 80, 91–138. doi: 10.1128/MMBR.00037-15
- De Bary, A. (1879). *Die Erscheinung der Symbiose*. Strasbourg: Karl J. Trubner.
- Drebes, G., Kühn, S. F., Gmelch, A., and Schnepf, E. (1996). *Cryothecomonas aestivalis* sp. nov., a colourless nanoflagellate feeding on the marine centric diatom *Guinardia delicatula* (Cleve). *Hasle. Helg. Mar. Res.* 50, 497–515.
- Dunne, J. A., Williams, R. J., and Martinez, N. D. (2002). Network structure and biodiversity loss in food webs: robustness increases with connectance. *Ecol. Lett.* 5, 558–567. doi: 10.1046/j.1461-0248.2002.00354.x
- Duret, M. T., Pachiadaki, M. G., Stewart, F. J., Sarode, N., Christaki, U., Monchy, S., et al. (2015). Size-fractionated diversity of eukaryotic microbial communities in the Eastern Tropical North Pacific oxygen minimum zone. *FEMS Microbiol. Ecol.* 91:fiv037. doi: 10.1093/femsec/fiv037
- Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26, 2460–2461. doi: 10.1093/bioinformatics/btq461
- Edgcomb, V. P. (2016). Marine protist associations and environmental impacts across trophic levels in the twilight zone and below. *Curr. Opin. Microbiol.* 31, 169–175. doi: 10.1016/j.mib.2016.04.001
- Falkowski, P. G., Fenchel, T., and Delong, E. F. (2008). The microbial engines that drive earth's biogeochemical cycles. *Science* 320, 1034–1039. doi: 10.1126/science.1153213
- Faust, K., and Raes, J. (2012). Microbial interactions: from networks to models. *Nat. Rev. Microbiol.* 10, 538–550. doi: 10.1038/nrmicro2832
- Fuhrman, J. A. (2009). Microbial community structure and its functional implications. *Nature* 459, 193–199. doi: 10.1038/nature08058
- Fuhrman, J. A., and Steele, J. A. (2008). Community structure of marine bacterioplankton: patterns, networks, and relationships to function. *Aquat. Microb. Ecol.* 53, 69–81. doi: 10.3354/ame01222
- Galand, P. E., Casamayor, E. O., Kirchman, D. L., and Lovejoy, C. (2009). Ecology of the rare microbial biosphere of the Arctic Ocean. *Proc. Natl. Acad. Sci. U.S.A.* 106, 22427–22432. doi: 10.1073/pnas.0908284106
- Genitsaris, S., Monchy, S., Breton, E., Lecuyer, E., and Christaki, U. (2016). Small-scale variability of protistan planktonic communities relative to environmental pressures and biotic interactions at two adjacent coastal stations. *Mar. Ecol. Prog. Ser.* 548, 61–75. doi: 10.3354/meps11647
- Genitsaris, S., Monchy, S., Viscogliosi, E., Sime-Ngando, T., Ferreira, S., and Christaki, U. (2015). Seasonal variations of marine protest community structure based on taxon-specific traits using the eastern English Channel as a model coastal system. *FEMS Microbiol. Ecol.* 91:fiv034. doi: 10.1093/femsec/fiv034
- Georges, C., Monchy, S., Genitsaris, S., and Christaki, U. (2014). Protist community composition during early phytoplankton blooms in the naturally iron-fertilized Kerguelen area (Southern Ocean). *Biogeosciences* 11, 5847–5863. doi: 10.5194/bg-11-5847-2014
- Gilbert, J. A., Steele, J. A., Caporaso, J. G., Steinbrück, L., Reeder, J., Temperton, B., et al. (2012). Defining seasonal marine microbial community dynamics. *ISME J.* 6, 298–308. doi: 10.1038/ismej.2011.107
- Gleason, F. H., Küpper, F. C., Amon, J. P., Picard, K., Gachon, C. M. M., Marano, A. V., et al. (2011). Zoospore true fungi in marine ecosystems: a review. *Mar. Freshw. Res.* 62, 383–393. doi: 10.1071/MF10294
- Gleason, F. H., Lilje, O., Marano, A. V., Sime-Ngando, T., Sullivan, B. K., Kirchman, M., et al. (2014). Ecological functions of zoospore hyperparasites. *Front. Microbiol.* 5:244. doi: 10.3389/fmicb.2014.00244
- Gratapanche, J. D., Breton, E., Brylinski, J. M., Lecuyer, E., and Christaki, U. (2011a). Succession of primary producers and micrograzers in a coastal system dominated by *Phaeocystis globosa* blooms. *J. Plankton Res.* 33, 37–50. doi: 10.1093/plankt/fbq097
- Gratapanche, J. D., Vincent, D., Breton, E., and Christaki, U. (2011b). Phytoplankton growth and microzooplankton grazing during a spring bloom in the eastern English Channel. *J. Exp. Mar. Biol. Ecol.* 404, 87–97. doi: 10.1016/j.jembe.2011.04.004
- Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., et al. (2013). The Protist Ribosomal Reference database (PR<sup>2</sup>): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. *Nucleic Acids Res.* 41, D597–D604. doi: 10.1093/nar/gks1160
- Guillou, L., Viprey, M., Chambouvet, A., Welsh, R. M., Kirkham, A. R., Massana, R., et al. (2008). Widespread occurrence and genetic diversity of marine parasitoids belonging to *Syndiniales* (*Alveolata*). *Environ. Microbiol.* 10, 3349–3365. doi: 10.1111/j.1462-2920.2008.01731.x
- Hamm, C. E., Simson, D. A., Merkel, R., and Smetacek, V. (1999). Colonies of *Phaeocystis globosa* are protected by a thin but tough skin. *Mar. Ecol. Prog. Ser.* 187, 101–111. doi: 10.3354/meps187101
- Harada, A., Ohtsuka, S., and Horiguchi, T. (2007). Species of the parasitic genus *Duboscquella* are members of the enigmatic marine alveolate group, I. *Protist* 158, 337–347. doi: 10.1016/j.protis.2007.03.005

- Hugoni, M., Taib, N., Debroas, D., Domaizon, I., Dufournel, I. J., Bronner, G., et al. (2013). Structure of the rare archaeal biosphere and seasonal dynamics of active ecotypes in surface coastal waters. *Proc. Natl. Acad. Sci. U.S.A.* 110, 6004–6009. doi: 10.1073/pnas.1216863110
- Humphries, M. D., and Gurney, K. (2008). Network “small-world-ness”: a quantitative method for determining canonical network equivalence. *PLoS ONE* 3:e0002051. doi: 10.1371/journal.pone.0002051
- Kimura, H., and Naganuma, T. (2001). Thraustochytrids: a neglected agent of the marine microbial food chain. *Aquat. Ecosyst. Health* 4, 13–18. doi: 10.1080/146349801753569243
- Klavness, D., Shalchian-Tabrizi, K., Thomsen, H. A., Eikrem, W., and Jakobsen, K. S. (2005). *Telonema antarcticum* sp. nov., a common marine phagotrophic flagellate. *Int. J. Syst. Evol. Micr.* 55, 2595–2604. doi: 10.1099/ij.s.0.63652-0
- Kramarsky-Winter, E., Harel, M., Siboni, N., Ben Dov, E., Brickner, I., Loya, Y., et al. (2006). Identification of a protist-coral association and its possible ecological role. *Mar. Ecol. Prog. Ser.* 317, 67–73. doi: 10.3354/meps317067
- Kunin, V., Engelbrekton, A., Ochman, H., and Hugenholtz, P. (2010). Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. *Environ. Microbiol.* 12, 118–123. doi: 10.1111/j.1462-2920.2009.02051.x
- Laferty, K. D., Allesina, S., Arim, M., Briggs, C. J., De Leo, G., Dobson, A. P., et al. (2008). Parasites in food webs: the ultimate missing links. *Ecol. Lett.* 11, 533–546. doi: 10.1111/j.1461-0248.2008.01174.x
- Lefèvre, E., Roussel, B., Amblard, C., and Sime-Ngando, T. (2008). The molecular diversity of freshwater picoeukaryotes reveals high occurrence of putative parasites in the plankton. *PLoS ONE* 3:e2324. doi: 10.1371/journal.pone.0002324
- Lesen, A. E., Juhl, A. R., and Anderson, O. R. (2010). Heterotrophic microplankton in the lower Hudson River Estuary: potential importance of naked, planktonic amoebas for bacterivory and carbon flux. *Aquat. Microb. Ecol.* 61, 45–56. doi: 10.3354/ame01434
- Lever, J. J., Van Nes, E. H., Scheffer, M., and Bascompte, J. (2014). The sudden collapse of pollinator communities. *Ecol. Lett.* 17, 350–359. doi: 10.1111/ele.12236
- Logares, R., Audic, S., Bass, D., Bittner, L., Boutte, C., Christen, R., et al. (2014). Patterns of rare and abundant marine microbial eukaryotes. *Curr. Biol.* 24, 813–821. doi: 10.1016/j.cub.2014.02.050
- López-García, P., and Moreira, D. (2008). Tracking microbial biodiversity through molecular and genomic ecology. *Res. Microbiol.* 159, 67–73. doi: 10.1016/j.resmic.2007.11.019
- López-García, P., Philippe, H., Gail, F., and Moreira, D. (2003). Autochthonous eukaryotic diversity in hydrothermal sediment and experimental microcolonizers at the Mid-Atlantic Ridge. *Proc. Natl. Acad. Sci. U.S.A.* 100, 697–702. doi: 10.1073/pnas.0235779100
- López-García, P., Rodríguez-Valera, F., Pedrós-Alió, C., and Moreira, D. (2001). Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. *Nature* 409, 603–607. doi: 10.1038/35054537
- Lynn, D. (2008). *The Ciliated Protozoa. Characterization, Classification and Guide to the Literature, 3rd Edn.* New York, NY: Springer.
- Manohar, C. S., and Raghukumar, C. (2013). Fungal diversity from various marine habitats deduced through culture-independent studies. *FEMS Microbiol. Lett.* 341, 69–78. doi: 10.1111/1574-6968.12087
- Marshall, W. L., Celio, G., McLaughlin, D. J., and Berbee, M. L. (2008). Multiple isolations of a culturable, motile Ichthyosporan (Mesomycetozoa, Opisthokonta), *Creolimax fragrantissima* n. gen., n. sp., from marine invertebrate digestive tracts. *Protist* 159, 415–433. doi: 10.1016/j.protis.2008.03.003
- Massana, R., Gobet, A., Audic, S., Bass, D., Bittner, L., Boutte, C., et al. (2015). Marine protist diversity in European coastal water and sediments as revealed by high-throughput sequencing. *Environ. Microbiol.* 10, 4035–4049. doi: 10.1111/1462-2920.12955
- Massana, R., Terrado, R., Forn, I., Lovejoy, C., and Pedrós-Alió, C. (2006). Distribution and abundance of uncultured heterotrophic flagellates in the world oceans. *Environ. Microbiol.* 8, 1515–1522. doi: 10.1111/j.1462-2920.2006.01042.x
- Monchy, S., Grattepance, J. D., Breton, E., Meloni, D., Sancier, G., Chabé, M., et al. (2012). Microplanktonic community structure in a coastal system relative to a *Phaeocystis* bloom inferred from morphological and tag pyrosequencing methods. *PLoS ONE* 7:e39924. doi: 10.1371/journal.pone.0039924
- Montoya, J. M., Pimm, S. L., and Solé, R. V. (2006). Ecological networks and their fragility. *Nature* 442, 259–264. doi: 10.1038/nature04927
- Mykkestad, S. M. (1995). Release of extracellular products by phytoplankton with special emphasis on polysaccharides. *Sci. Total Environ.* 165, 155–164. doi: 10.1016/0048-9697(95)04549-G
- Niquil, N., Kagami, M., Urabe, J., Christaki, U., Viscogliosi, E., and Sime-Ngando, T. (2011). Potential role of fungi in plankton food web functioning and stability: a simulation analysis based on Lake Biwa inverse model. *Hydrobiologia* 659, 65–79. doi: 10.1007/s10750-010-0308-6
- Not, F., Gausling, R., Azam, F., Heidelberg, J. F., and Worden, A. Z. (2007). Vertical distribution of picoeukaryotic diversity in the Sargasso Sea. *Environ. Microbiol.* 9, 1233–1252. doi: 10.1111/j.1462-2920.2007.01247.x
- Not, F., Siano, R., Kooistra, W. H. C. F., Simon, N., Vaulot, D., and Probert, I. (2012). Diversity and ecology of eukaryotic marine phytoplankton. *Adv. Bot. Res.* 64, 1–53. doi: 10.1016/B978-0-12-391499-6.00001-3
- Park, M. G., Yih, W., and Coats, D. W. (2004). Parasites and phytoplankton, with special emphasis on dinoflagellate infections. *J. Eukaryot. Microbiol.* 51, 145–155. doi: 10.1111/j.1550-7408.2004.tb00539.x
- Peperzak, L., Colijn, F., Gieskes, W. W. C., and Peeters, J. C. H. (1998). Development of the diatom – *Phaeocystis* spring bloom in the Dutch coastal zone of the North Sea: the silicon depletion versus the daily irradiance threshold hypothesis. *J. Plankton Res.* 20, 517–537. doi: 10.1093/plankt/20.3.517
- Pirreda, R., Tomasino, M. P., D’Erchia, A. M., Manzari, C., Pesole, G., Montresor, M., et al. (2016). Diversity and temporal patterns of planktonic protist assemblages at a Mediterranean long term ecological research site. *FEMS Microbiol. Ecol.* 93:fiw200. doi: 10.1093/femsec/fiw200
- Quince, C., Lanzén, A., Curtis, T. P., Davenport, R. J., Hall, N., Head, I. M., et al. (2009). Accurate determination of microbial diversity from 454 pyrosequencing data. *Nat. Methods* 6, 639–641. doi: 10.1038/nmeth.1361
- Raghukumar, S. (2002). Ecology of the marine protists, the Labyrinthulomycetes (Thraustochytrids and Labyrinthulids). *Eur. J. Protistol.* 38, 127–145. doi: 10.1078/0932-4739-00832
- Reshef, D. N., Reshef, Y. A., Finucane, H. K., Grossman, S. R., McVean, G., Turnbaugh, P. J., et al. (2011). Detecting novel associations in large data sets. *Science* 334, 1518–1524. doi: 10.1126/science.1205438
- Rousseau, V., Vaulot, D., Casotti, R., Cariou, V., Lenz, J., Gunkel, J., et al. (1994). The life cycle of *Phaeocystis* (*Prymnesiophyceae*): evidence and hypotheses. *J. Marine Syst.* 5, 23–39. doi: 10.1016/0924-7963(94)90014-0
- Ruan, Q., Dutta, D., Schwalbach, M. S., Steele, J. A., Fuhrman, J. A., and Sun, F. (2006). Local similarity analysis reveals unique associations among marine bacterioplankton species and environmental factors. *Bioinformatics* 22, 2532–2538. doi: 10.1093/bioinformatics/btl417
- Santoferrara, L. F., Grattepance, J. D., Katz, L. A., and McManus, G. B. (2014). Pyrosequencing for assessing diversity of eukaryotic microbes: analysis of data on marine planktonic ciliates and comparison with traditional methods. *Environ. Microbiol.* 16, 2752–2763. doi: 10.1111/1462-2920.12380
- Schapiro, M., Vincent, D., Gentilhomme, V., and Seuront, L. (2008). Temporal patterns of phytoplankton assemblages, size spectra and diversity during the wane of a *Phaeocystis globosa* spring bloom in hydrologically contrasted coastal waters. *J. Mar. Biol. Assoc. U.K.* 88, 649–662. doi: 10.1017/S0025315408001306
- Scheckenbach, F., Wylezich, C., Mylnikov, A. P., Weitere, M., and Arndt, H. (2006). Molecular comparisons of freshwater and marine isolates of the same morphospecies of heterotrophic flagellates. *Appl. Environ. Microb.* 72, 6638–6643. doi: 10.1128/AEM.02547-05
- Schloss, P. D., Gevers, D., and Westcott, S. L. (2011). Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. *PLoS ONE* 6:e27310. doi: 10.1371/journal.pone.0027310
- Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., et al. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* 75, 7537–7541. doi: 10.1128/AEM.01541-09
- Schnepf, E., and Kühn, S. F. (2000). Food uptake and fine structure of *Cryothecomonas longipes* sp. nov., a marine nanoflagellate incertae sedis feeding phagotrophically on large diatoms. *Helgol. Mar. Res.* 54, 18–32. doi: 10.1007/s101520050032

- Scholz, B., Guillou, L., Marano, A. V., Neuhauser, S., Sullivan, B. K., Karsten, U., K pper, F. C., et al. (2016). Zoosporic parasites infecting marine diatoms – A black box that needs to be opened. *Fungal Ecol.* 19, 59–76. doi: 10.1016/j.funeco.2015.09.002
- Sherr, E. B., and Sherr, B. F. (2007). Heterotrophic dinoflagellates: a significant component of microzooplankton biomass and major grazers of diatoms in the sea. *Mar. Ecol. Prog. Ser.* 352, 187–197. doi: 10.3354/meps07161
- Skovgaard, A. (2014). Dirty tricks in the plankton: diversity and role of marine parasitic protists. *Acta Protozool.* 20, 130–134. doi: 10.4467/16890027AP.14.006.1443
- Skovgaard, A., Massana, R., Balague, V., and Saiz, E. (2005). Phylogenetic position of the copepod-infesting parasite *Syndinium turbo* (Dinoflagellata, Syndinea). *Protist* 156, 413–423. doi: 10.1016/j.protis.2005.08.002
- Smoot, M. E., Ono, K., Ruschinski, J., Wang, P. L., and Ideker, T. (2011). Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* 27, 431–432. doi: 10.1093/bioinformatics/btq675
- Steele, J. A., Countway, P. D., Xia, L., Vigil, P. D., Beman, J. M., Kim, D. Y., et al. (2011). Marine bacterial, archaeal and protistan association networks reveal ecological linkages. *ISME J.* 5, 1414–1425. doi: 10.1038/ismej.2011.24
- Stoecker, D. K. (1999). Mixotrophy among Dinoflagellates. *J. Eukaryot. Microbiol.* 46, 397–401. doi: 10.1111/j.1550-7408.1999.tb04619.x
- Stoeck, T., Bass, D., Nebel, M., Christen, R., Jones, M. D. M., Breiner, H.-W., et al. (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
- Strickland, J., and Parsons, T. A. (1972). *A Practical Handbook of Seawater Analysis*. Ottawa, ON: Fisheries Research Board of Canada.
- Tillmann, U., Hesse, K. J., and Tillmann, A. (1999). Large-scale parasitic infection of diatoms in the Northfrisian Wadden Sea. *J. Sea Res.* 42, 255–261. doi: 10.1016/S1385-1101(99)00029-5
- Tosadori, G., Bestvina, I., Spoto, F., Laudanna, C., and Scardoni, G. (2016). Creating, generating and comparing random network models with Network Randomizer. *F1000Res.* 5:2524. doi: 10.12688/f1000research.9203.1
- Watts, D. J., and Strogatz, S. H. (1998). Collective dynamics of “small-world” networks. *Nature* 393, 440–442.
- Weiss, S., Van Treuren, W., Lozupone, C., Faust, K., Friedman, J., Deng, Y., et al. (2016). Correlation detection strategies in microbial data sets vary widely in sensitivity and precision. *ISME J.* 10, 1669–1681. doi: 10.1038/ismej.2015.235
- Windsor, D. A. (1998). Controversies in parasitology, most of the species on earth are parasites. *Int. J. Parasitol.* 28, 1939–1941. doi: 10.1016/S0020-7519(98)00153-2
- Xia, L. C., Ai, D., Cram, J., Fuhrman, J. A., and Sun, F. (2012). Efficient statistical significance approximation for local similarity analysis of high-throughput time series data. *Bioinformatics* 29, 230–237. doi: 10.1093/bioinformatics/bts668
- Xia, L. C., Steele, J. A., Cram, J. A., Cardon, Z. G., Simmons, S. L., Vallino, J. J., et al. (2011). Extended local similarity analysis (eLSA) of microbial community and other time series data with replicates. *BMC Syst. Biol.* 5(Suppl. 2):S15. doi: 10.1186/1752-0509-5-S2-S15

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright   2017 Christaki, Genitsaris, Monchy, Li, Rachik, Breton and Sime-Ngando. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Additional tables

**Table S1:** Network statistics of the MIC, (a) for the total community including environmental variables, (b) for the connections only with environmental variables, and (c) the connections of SymbDec. The connections with p-value < 0.01 and MIC > 0.6 are shown.

Network Statistics	MIC		
	Total	Environmental	SymbDec
Nodes	93	15	60
Edges (pairs)	254	18	86
Negative edges	57	8	12
Clustering coefficient	0.269	0.298	0.172
Connected components	4	1	6
Network diameter	9	5	5
Network radius	1	3	1
Network centralization	0.161	0.297	0.3
Shortest paths	7320 (85%)	210 (100%)	1718 (48%)
Characteristic path length	3.343	2.705	2.428
Avg. number of neighbors	5.462	2.4	2.867
Network density	0.059	0.171	0.049
Network heterogeneity	0.879	0.742	1.308

**Table S2.** Alignment of OTU reads unveiled during two independent surveys (2009 and 2011-2013) in the EEC. The closely related OTU reads from two independent surveys (this study and Christaki et al., 2014) were compared by performing sequences alignments using MUSCLE v3.8.31 (Edgar RC, 2004). OTU reads appeared to share enough similarities to be considered the same organism: OTU004 (this study) and EEC09\_1174 (from Christaki et al., 2014), OTU020 and EEC09\_1174, and finally OTU045 and EEC09\_1001. The percentage of identity and the differences between each couple of sequence are indicated. It is noteworthy that, when differences arise between two sequences (highlighted in yellow) they are restricted to homopolymeric region (stretch of repeated identical nucleotides). Homopolymers are known to generate sequence errors with 454-pyosequencing.

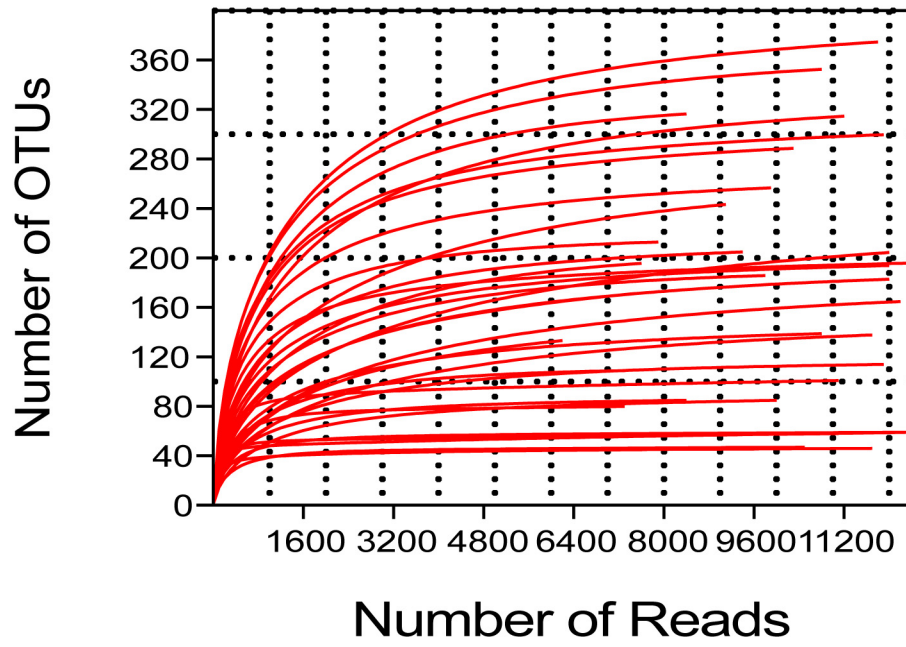
	<b>Alignment</b>		<b>Identity</b>	<b>Differences</b>
<b>Syndiniales</b>	Otu004	ATTAAACAGTTATAGTTTCCATGTCAGTTTTTTGCTACTTGGATATCCACGGTAATTCTAGAGC	98.7%	- Homopolymer 'AAAA' - Homopolymer 'TTTTTT' - Homopolymer 'TTT'
	EEC09_1168	ATTAAACAGTTATAGTTTCCATGTCAGTTTTTTGCTACTTGGATATCCACGGTAATTCTAGAGC		
	Otu004	TAATACATGCGTCCAGACCCGACTTTGCGGAAGGGTTGTGCTTATTAGACATTGAACCATCCCCG		
	EEC09_1168	TAATACATGCGTCCAGACCCGACTTTGCGGAAGGGTTGTGCTTATTAGACATTGAACCATCCCCG		
Otu004	GCTTGCCCGGTTTCGAGGTGATTGATGGTAATCGAACGAATCGCATGCTTTGCGGCGATGATTCA	98.7%	- Homopolymer 'TTT'	
EEC09_1168	GCTTGCCCGGTTTCGAGGTGATTGATGGTAATCGAACGAATCGCATGCTTTGCGGCGATGATTCA			
Otu004	TTCAAGTTTCTGACCTATCAGCTTCCGACGGTAGGGTATTGGCCTACCGT			
EEC09_1168	TTCAAGTTTCTGACCTATCAGCTTCCGACGGTAGGGTATTGGCCTACCGT			
<b>Fungi</b>	Otu020	ATTATATAAGTTATCGTTTATTTGATAGTACCTTACTACTTGGATAACCGTGGTAATTCTAGAGC	100 %	None
	EEC09_1174	ATTATATAAGTTATCGTTTATTTGATAGTACCTTACTACTTGGATAACCGTGGTAATTCTAGAGC		
	Otu020	TAATACATGCTAAAAATCCCGACTTCGGAAGGGATGTATTTATTAGATTAAAAACCAATGCCCTC		
	EEC09_1174	TAATACATGCTAAAAATCCCGACTTCGGAAGGGATGTATTTATTAGATTAAAAACCAATGCCCTC		
Otu020	TGGGCTCTTTGGTGATTCATGATAACTTCTCGAATCGCACGGCCTTGCGCCGGCGATGGTTCATT	100 %	None	
EEC09_1174	TGGGCTCTTTGGTGATTCATGATAACTTCTCGAATCGCACGGCCTTGCGCCGGCGATGGTTCATT			
Otu020	CAAATTTCTTCCCTATCAACTTTCGATGTTTGGGTATTGGCCAAACAT			
EEC09_1174	CAAATTTCTTCCCTATCAACTTTCGATGTTTGGGTATTGGCCAAACAT			



<b>Cercozoa</b>	Otu045	ATTATATCAGTAATAATTTATTTGATGGTTCCTTACTACACGGATAACTGTAGTAATTCTAGAGC	99.6%	- Homopolymer 'TTTT'
	EEC09_1001	ATTATATCAGTAATAATTTATTTGATGGTTCCTTACTACACGGATAACTGTAGTAATTCTAGAGC		
	Otu045	TAATACGTGCGTAAAATCCCGACTCTTGCGGGGAGGGATGTATTTATTAGATAAAAAGCCAGCGC		
	EEC09_1001	TAATACGTGCGTAAAATCCCGACTCTTGCGGGGAGGGATGTATTTATTAGATAAAAAGCCAGCGC		
	Otu045	GGGTTCTGCCCGGTTTTGTGGTGAATCATAGTAACTTTTCGAATCGCATGGCTTCACGCCGGCGA		
	EEC09_1001	GGGTTCTGCCCGGTTTTGTGGTGAATCATAGTAACTTTTCGAATCGCATGGCTTCACGCCGGCGA		
	Otu045	TAGTTCATTCAAATTTCTGCCCTATCAGCTTTCGACGGTAGTGTAGTGGACTACCGT		
	EEC09_1001	TAGTTCATTCAAATTTCTGCCCTATCAGCTTTCGACGGTAGTGTAGTGGACTACCGT		

Additional figure

**Figure S1** : Rarefaction curves representing the of OTUs against the number of high-quality reads





# **CHAPITRE V**

**DIVERSITE DES PROTISTES DANS UN ECOSYSTRME COTIER OLIGOTROPHE  
(MEDITERRANEE OCCIDENTALE, GOLF DU LION)**



**Diversité des protistes dans un écosystème côtier oligotrophe (Méditerranée Occidentale, Golf de Lion) évaluée par les techniques de séquençage de nouvelle génération (NGS : Next Generation Sequencing)**

*Small and (sub)mesoscale variability of protistan plankton communities in oligotrophic coastal ecosystems (NW Mediterranean, Gulf of Lions) assessed by NGS (In prep.)*

**Sara Rachik, Sébastien Monchy, Anne Patrenko, Elsa Breton, Bernard Queguiner, Urania Christaki**

**RESUME DU CONTEXTE ET OBJECTIFS**

Ce troisième chapitre a été dédié à l'étude de la diversité des communautés eucaryotes microbiennes dans un écosystème aux conditions trophiques contrastées à la Manche Orientale : La Méditerranée Occidentale (Golfe du Lion). La composante biologique de cet écosystème oligotrophe est sous l'influence de différents facteurs abiotiques (par exemple : vents, upwelling, limitation en éléments nutritifs et intrusions) et biotiques (par exemple : la concurrence vis-à-vis des éléments nutritifs et les relations inter-taxa). Plusieurs études conduites en Méditerranée Occidentale ont souligné l'impact que peut avoir les changements dans les conditions environnementales sur la communauté planctonique (par exemple : *Marty et al., 2002, Ludwig et al., 2009*). Cependant, la plupart des études menées sur la diversité biologique, majoritairement réalisées par des approches traditionnelles, concernaient principalement des genres ou groupes taxonomiques précis (par exemple : les espèces indicatrices des changements environnementaux et celles toxiques). Par ailleurs, de notre connaissance, peu d'études ont ciblé un large spectre de diversité taxonomique (*Penna et al. 2017*).

Notre étude a été réalisée dans le cadre du projet SPECiMed (Structures of Planktonic Ecosystems in the North-western Mediterranean) qui visait à établir une plate-forme d'observation de trois ans (2012-2014) des communautés planctoniques, incluant les bactéries, micro-zoo- et mésozooplancton tout en associant les cycles biogéochimiques des éléments majeurs (C, N, P, et Si). Notre étude contribuait dans ce projet à l'exploration de la variabilité de la structure et de la dynamique des communautés eucaryotes microbiennes au regard des paramètres environnementaux et de la zone géographique étudiée. Pour ceci, nous avons traité 59 échantillons provenant du Golfe du Lion et prélevés au niveau de quatre stations (deux stations situées au sud-ouest du Golfe du Lion et deux situées à son sud-est). L'étude de la diversité a été réalisée par séquençage de la région variable du gène de l'ARNr18S par la technique Illumina.

Les objectifs principaux de cette étude étaient d'évaluer en premier temps la diversité globale des protistes telle que révélée par le séquençage à haut débit et dans un second temps, de déceler les différences et les similitudes pouvant exister entre les stations côtières et celles du large et entre le sud-est et sud-ouest du Golfe du Lion et leur lien avec les variables environnementales.

## RESUME

La composition de la communauté eucaryote microbienne a été étudiée par séquençage de l'ARN18S des régions variables V2-V3 via la technique Illumina. Les données issues du séquençage ont été complétées par les comptages microscopiques des dinoflagellés et des ciliés. L'échantillonnage a été conduit au nord-ouest de la Méditerranée entre 2012-2014 dans quatre stations, se situant à l'est (SOLEMIO et JULIO) et à l'ouest (SOLA et MOLA) du Golfe du Lion (GoL). Cette région est soumise à l'influence directe du Rhône. Des mesures des variables environnementales (e.g. salinité, température, sels nutritifs) ont été effectuées à chaque prélèvement. L'analyse en composante principale sur les paramètres environnementaux a montré une saisonnalité plus claire à la station MOLA (située au large de Banyuls sur mer) comparée aux stations SOLA, SOLEMIO et JULIO. Un total de 3358 unités taxonomiques opérationnelles (OTU), réparties en sept super-groupes (Alveolata, Amoebozoa, Apusozoa, Archeplastida, Hacrobia, Opisthokonta, Rhizaria et Stramenopiles) et d'autres eucaryotes, ont été affiliées à 48 groupes taxonomiques supérieurs au sein des 59 échantillons prélevés en surface (1-5m). En prenant en considération seulement les groupes taxonomiques les plus abondants, qu'elles que soient les dates, aucune différence significative n'a été observée entre les stations de l'est ou de l'ouest du GoL en termes de diversité basée sur le nombre de séquences (Kolmogorov-test,  $p_{value} > 0.05$ ). Les alvéolaires, principalement Dinophycées et deux groupes majeurs de Syndiniales (MALV-I et MALV-II, alvéolaires marins) dominaient les assemblages en nombre de séquences et en diversité. Les ciliés n'occupaient que 1 à 2% de la richesse totale mais représentaient un nombre relativement élevé d'OTU (>60), dominés par les mixotrophes (*Laboea strobila*, *Tontonia sp.* et *Myrionecta rubra*). Quant aux Straménopiles, les MAST (Straménopiles marins) et les bacillariophycées étaient les plus représentatifs en termes de nombre de séquences et de diversité. Les Chlorophytes, Fungi et Cercozoa formaient les groupes les moins représentés dans l'ensemble des stations. Les analyses de co-inertie pour chaque station, ont montré un couplage significatif entre les conditions environnementales saisonnières et les abondances des super-groupes au sein de chaque station à l'exception de JULIO. De par cette étude, nous avons pu générer un jeu de données le plus exhaustif possible de la diversité des protistes microbiens eucaryotes dans la Méditerranée Nord-Ouest et discuter l'influence du Rhône sur leur structuration.

**Mots clefs :** Séquençage, ARN18S, Illumina, nord-ouest de la méditerranée, diversité, protistes, Golfe du Lion, Rhône

## INTRODUCTION

La Méditerranée est l'un des écosystèmes marins les plus oligotrophes au monde. Jusqu'à 90-95% pendant la période stratifiée de la production primaire est soutenue par le recyclage interne de la matière organique (production régénérée) (*Moutin et Raimbault, 2002*). ~~En plus,~~ Elle dispose d'une variabilité spatiale considérable : Par exemple, à l'ouest, le front d'Almeria-Oran est une zone caractérisée par sa production primaire élevée (*Videau et al., 1994, Van Wambeke et al., 2004*) par rapport aux eaux environnantes. Dans cette zone, les tourbillons persistants de méso-échelle en mer ouverte diffèrent non seulement en termes de biomasse et de production par rapport aux stations de référence extérieures situées dans le même bassin, mais également entre elles (par exemple, *Christaki et al., 2011*).

Le bassin de la Méditerranée du Nord-Ouest (NWM) est une mosaïque d'écosystèmes jumelés présentant un estuaire à l'embouchure d'une grande rivière, la rivière du Rhône, qui apporte des charges nutritives localement importantes sur un plateau continental, le GoL et un courant côtier, le courant nord (NC), qui sépare le continuum aquatique terre-océan d'un gyre oligotrophe (pour plus de détails : voir **Chapitre 1 partie 2**). Dans le nord-ouest de la Méditerranée, plusieurs études sur de longues séries temporelles ont détecté une dérive à long terme de la composition et de la dynamique du plancton (e.g *Gómez & Gorsky, 2003; Pereira Quiroga, 2005; Tunin Ley et al., 2007*). A la station DYFAMED, Marty et al. (2002) ont montré une tendance vers une augmentation de la biomasse phytoplanctonique due au phytoplancton de petite taille (nanoflagellés et cyanobactéries) qui a été favorisé par une période de stratification estivale allongée. Ce résultat rejoignait les observations de Goffart et al. (2002) dans la mer Ligure (NMW), où ils ont démontré un impact significatif des conditions environnementales, et en particulier de la température sur le phytoplancton. D'autres études ont souligné que la réponse des communautés planctoniques vis-à-vis des changements environnementaux, traduite par leur présence ou leur absence dans l'environnement peut être un reflet de changements dans les conditions environnementales (par exemple : *Molinero et al., 2005*). Un nombre croissant d'études soutient l'idée que les analyses sur les communautés protistes et leur rôle dans l'écosystème peut aider à mieux comprendre les processus du changements écosystémiques (*Karl et al., 2001; Mackas et al., 2001; Batchelder & Powell, 2002; Beaugrand et al., 2002; Hooff et Peterson, 2006; Chiba et al., 2008*).



Notre étude a été menée dans le Golfe de Lion (GoL). L'environnement physique du GoL jusqu'à sa limite offshore est sous le contrôle de trois processus majeurs: les intrusions du courant Nord (CN) à l'entrée Est, les entrées d'eau douce du panache du Rhône et la génération de tourbillons à mésoéchelle à la sortie Ouest. Ces processus sont fortement influencés par la fréquence des vents dominants (Mistral et Tramontane) et les régimes pluviométriques. La présente étude a été menée dans le cadre du projet SPECIMED (Structures des écosystèmes planctoniques dans le nord-ouest méditerranéen). La stratégie d'échantillonnage couvrait les stations spatialement diverses dans le GoL et s'appuyait sur les stations côtières SOMLIT (Service d'Observation en Milieu LITtoral, <http://somalit.epoc.u-bordeaux1.fr/fr/>) de Marseille (SOLEMIO) et Banyuls sur mer (SOLA). Deux autres stations situées au large ont été intégrées dans l'étude : JULIO (entrée SE du GoL) et MOLA (Sortie SW, **Fig. 1**).

L'objectif principal de la présente étude était d'explorer la variabilité de la structure de la communauté des protistes planctoniques (microbes eucaryotes) en utilisant le séquençage du gène de l'ARNr 18S (Illumina MiSeq) à la (sub)mésoéchelle dans le GoL. L'utilisation d'Illumina MiSeq a permis d'accéder à un large pan de diversité, y compris les groupes quasi impossibles à caractériser en utilisant des approches traditionnelles de microscopie et/ou de culture (par exemple, MAST, MALV, Fungi et autres). Étant donné que dans le GoL, la (sub)mésoéchelle est influencée directement par le rôle dominant du vent sur la circulation (*Millot, 1982, Hu et al., 2009*), nous nous attendons donc à ce que les quatre stations présentent des différences dans le forçage environnemental et que ces différences se reflètent dans la structure et la succession des communautés planctoniques.

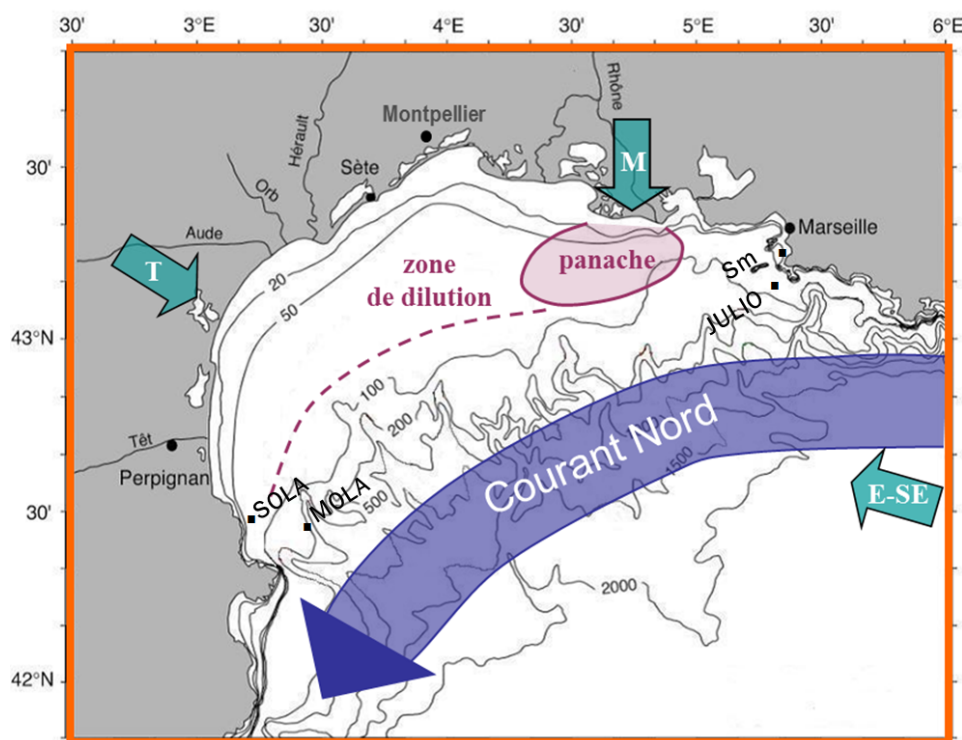
Les deux questions principales étaient les suivantes: (i) Quelle est la diversité globale des protistes telle que révélée par MiSeq ? (ii) Quelles sont les principales différences et similitudes entre les stations côtières et celles du large et entre le sud-est et sud-ouest du GoL et leur lien avec les variables environnementales ?

Dans ce chapitre, une partie est consacrée à l'évaluation de la diversité des communautés microzooplanctoniques (ciliés et dinoflagellés) par microscopie inversée. Le microzooplancton a été dénombré et identifié pour la plupart des échantillons analysés ainsi que par séquençage (**Tableau 1**).

## MATERIELS ET METHODES

### *Sites d'échantillonnage et prélèvements*

L'échantillonnage a été effectué dans le Golfe de Lion (GoL) situé dans la partie nord-ouest du bassin occidental de la mer Méditerranée. Les prélèvements ont couvert 4 stations entre 2012 et 2014 : SOLEMIO (N 43°14'30", E 05°17'30", 60 m de profondeur) et JULIO (N 43°06'00", E 05°15'00", 100 m de profondeur) localisées à 8.77 miles de distance et respectivement situées à la frontière sud de la baie de Marseille et à proximité de la bordure du plateau, tandis que SOLA (N 42°29'18 " , E 03°08'42", 27 m de profondeur) et MOLA (N 42°27'08", E 03°32'28 " , 600m de profondeur), localisées à 17.68 miles de distance et respectivement situées dans la baie de Banyuls et à l'entrée sur le flanc nord du canyon Lacaze-Duthiers (**Fig.1**). Les stations SOLA et SOLEMIO font partie du réseau SOMLIT (Service d'Observation en Milieu Littoral) (<http://somlit.epoc.u-bordeaux1.fr/fr/>).



**Figure 1.** Localisation géographique des sites d'échantillonnage (Méditerranée).

Carte réalisée avec l'aide de *Anne Patrenko*, MIO.

En ce qui concerne l'évaluation de la diversité par séquençage, les échantillons ont été prélevés en sub-surface à 1-5m de profondeur, recueillis dans des bouteilles de polyéthylène

lavées à l'acide, stockées dans l'obscurité à la température de surface *in situ* et filtrés dans les deux heures qui suivaient. La filtration sur des filtres de 0.2 µm (47mm de diamètre) a été effectuée à l'aide d'une pompe péristaltique à basse pression de filtration (15 tours/min) afin d'éviter la saturation du filtre et minimiser la destruction des organismes. Les filtres ont été immédiatement stockés à -80°C jusqu'à l'extraction de l'ADN.

De 129 échantillons prélevés à différentes profondeurs (5m, 40m, 150m et 300m), un total de 59 échantillons, correspondant à ceux prélevés en surface, a été utilisé pour la suite des analyses (15, 13, 18, 13 pour SOLA, MOLA, SOLEMIO et JULIO, respectivement, **Tableau 1**). Et comme précédemment mentionné, nous avons réalisé également le comptage des ciliés et des dinoflagellés pour les échantillons équivalents à ceux séquencés (**Tableau 1**), par contre le comptage du phytoplancton a été réalisé à Marseille

	<b>SOLA</b>	<b>MOLA</b>	<b>SOLEMIO</b>	<b>JULIO</b>
<b>2012</b>	19/04	26/01	18/04*	28/05*
	11/06	10/05	28/05*	14/06
		14/06	14/06*	06/09*
		10/10	06/09*	11/10
			11/10*	09/11
			09/11*	
			07/12*	
<b>2013</b>	21/05*	20/02	12/02	04/06*
	17/06*	21/03	30/04*	17/07*
	15/07*	17/04	04/06	04/09
	12/08*	11/06*	17/07*	01/10*
	09/09*	11/07*	01/10*	08/11
	28/11	25/09	08/11*	10/12
		24/10	10/12	
<b>2014</b>	13/01	12/06	14/01	14/01
	10/02	24/07	14/02	13/03
	24/03		13/03*	
	22/04		28/05	
	19/05			
	23/06			
	21/07			

**Tableau 1** : Tableau récapitulatif des dates d'échantillonnage, pour les 4 stations (15, 13, 18, 13 dates pour SOLA, MOLA, SOLEMIO et JULIO, respectivement). \* : Les dates des échantillons comptés au microscope inversé (5, 2, 12, 5 dates pour SOLA, MOLA, SOLEMIO et JULIO, respectivement).

Les prélèvements marins des deux régions d'échantillonnage ont été fixés au lugol (2%) puis stockés jusqu'à l'analyse à 4°C. Un volume de 250 à 520ml d'échantillon provenant de la Méditerranée a été mis à sédimenter pendant 72h, le surnageant a été ensuite aspiré (**Fig.2**). Les 100 ml restant dans le fond de l'éprouvette ont été transférés dans une chambre hydrobios pour une deuxième sédimentation pendant au moins 24h. Les comptages ont été effectués avec un Nikon EclipseTE2000-s au grossissement x400 (**Fig.2**). L'identification s'est faite selon les critères morphologiques des microorganismes.



**Figure 2:** Quelques images prises lors de l'aspiration du surnageant (image gauche) et du comptage microscopique (image droite)

Au cours de chaque échantillonnage, les paramètres physico-chimiques ont été mesurés. La température (°C), la salinité et le pH de l'eau de mer ont été mesurés par une CTD Seabird 19 SeaCat. Les éléments nutritifs ( $\text{NO}_3$ ,  $\text{NO}_2$ ,  $\text{PO}_4$  et  $\text{SiOH}_4$ ,  $\mu\text{M}$ ) ont été analysés à l'aide de l'auto-analyseur Brann Luebbe selon Aminot et Kerouel (2007). Les pigments chlorophylliens ( $\text{Chla}$ ,  $\mu\text{g L}^{-1}$ ) ont été extraits pendant 24h dans 90% d'acétone suivant le protocole fluorimétrique d'Aminot et Kerouel (2004). Quant au carbone organique particulaire (POC,  $\text{mg L}^{-1}$ ) et l'azote organique particulaire (PON,  $\text{mg L}^{-1}$ ) ont été traités à l'aide d'un analyseur élémentaire CHN Perkin Elmer 2400 Series CHNS/O. Aux stations SOLA et MOLA, des mesures supplémentaires ont été également réalisées. Les concentrations en oxygène dissous ( $\text{ml/L}$ ) ont été mesurées par microtitration, telle que proposée par Winkler (1888) (Aminot et Chaussepied, 1983). Les concentrations en matières en suspension (MES ou SPM,  $\text{mg L}^{-1}$ ) ont été déterminées par pesée avant et après filtration au travers de filtres GF/F (porosité d'environ  $0.7 \mu\text{m}$ ) (pour plus de détails : <http://somlit.epoc.u-bordeaux1.fr/fr/>).

Pour la période 2012-2014, les variables environnementales ont été enregistrées à 19, 140, 18 et 24 reprises à MOLA, SOLA, JULIO et SOLEMIO, respectivement. Cependant dans ce travail nous avons utilisé pour les analyses statistiques (e.g analyses en composante principale, ACP) que les mesures correspondantes aux dates des échantillons séquencés.

### ***Extraction d'ADN, amplification PCR et séquençage par Illumina***

L'ADN ribosomique des organismes protistes a été extrait et purifié avec le Kit d'isolement d'ADN (Qiagen Allprep DNA/RNA Mini Kit 50) en suivant le protocole du fabricant. Les échantillons traités contenaient entre 2.59 et 53.1 ng.µL<sup>-1</sup> d'ADN. Les dosages des quantités d'ADN ont été réalisés avec le Qubit 2.0 Fluorometer (Thermo Fisher Scientific Inc., Massachusetts, USA). Les échantillons d'ADN ont été amplifiés en utilisant les deux amorces universelles d'eucaryotes 18S-82F (5'-GAAACTGCGAATGGCTC-3', *Lopez-Garcia et al., 2003*) et Euk-516r (5'-ACCAGACTTGCCCTCC-3', *Amann et al., 1990*), conçues pour amplifier les régions variables des eucaryotes V2 et V3 du gène de l'ARNr18S (environ 470-480 pb). Ces amorces ont été utilisées avec succès dans des études similaires (*Monchy et al., 2012* ; *Christaki et al., 2014* ; *Georges et al., 2014* ; *Genitsaris et al., 2015* ; *2016*).

- ***Amplification PCR***

Les échantillons ont été préparés selon le protocole développé par **Genoscreen** (*Lille, France*), à partir des extraits d'ADNg normalisés à la concentration de 2.5 ng/µl. Une quantité de 5ng de matrice ADN a été utilisée pour chacune des amplifications ADNr 18S. Des contrôles qualitatifs et quantitatifs ont été réalisés tout au long du processus de préparation des bibliothèques d'amplicons. L'amplification avec les amorces spécifiques 18S montrait un polymorphisme de taille unique, entre 490 et 500 pb. Un mélange final équimolaire a été réalisé à la concentration de 4nM à partir du dosage final des bibliothèques. Ce mélange a été ensuite contrôlé sur Bioanalyzer (Agilent) à l'aide d'une puce High Sensitivity.

- ***Séquençage***

Les amplicons ont été mélangés et séquencés sur une « flow cell » (lame de séquençage) de MiSeq 2x300bp Illumina, à partir de la bibliothèque finale à 4 pM, complétée par une bibliothèque contrôle PhiX à 15% (V/V). Ce contrôle PhiX est une bibliothèque « *shotgun* » construite à partir

du génome du Phage PhiX174. Cette librairie possède un double intérêt (i) contrôler les différentes étapes du processus de séquençage (clusterisation, taux d'erreur ...) et (ii) apporter une variabilité nucléotidique suffisante à tous les cycles du séquençage dans le cadre du séquençage d'amplicons de faible diversité comme pour le gène codant pour l'ARNr18S.

### ***Analyse de la qualité et affiliations taxonomiques des séquences***

Les données générées durant le séquençage ont été traitées avec le logiciel MOTHUR v1.34.0 (Schloss et al., 2009) suivant la procédure d'analyse standard ([http://www.mothur.org/wiki/MiSeq\\_SOP](http://www.mothur.org/wiki/MiSeq_SOP)) (Schloss et al. 2011). Les séquences ont été alignées sur la base de données de référence SILVA 108 (<http://www.arb-silva.de/>). Les chimères ont été enlevées à l'aide du logiciel UCHIME (Edgar, 2010). Après normalisation, effectuée par rapport à l'échantillon ayant le plus faible nombre de séquences, le jeu de donnée contenait un total de 248.788 séquences. Les séquences ont été regroupées par la suite en OTU à 97% de similarité. Les « singletons », séquences n'apparaissant qu'une seule fois dans l'ensemble des données, ont été supprimées afin d'éviter la surestimation de la richesse spécifique des échantillons (Kunin et al., 2010 ; Behnke et al., 2011 ; Monchy et al., 2012). L'affiliation taxonomique a été réalisée en comparant les séquences de chaque OTU (total de 3358 OTU, **Fig.4**) avec celles présentes dans la base de données PR2 (Guillou et al., 2013) en utilisant le programme BLASTN (Altschul et al., 1990). Finalement, les OTUs identifiées comme métazoaires (516 OTU) ont été retirées de l'analyse.

### ***Traitements et analyses des données***

Les analyses statistiques ont été effectuées sous R (v 3.2.3). **(1)** L'analyse en Composante Principale (ACP) a été effectuée pour explorer les inter-corrélations entre les variables environnementales, en utilisant le package « ade4 » (Dray and Dufour, 2007). Pour se rapprocher d'une distribution normale des données, la transformation BoxCox (Box and Cox, 1964), a été appliquée en utilisant le package « caret » (Kuhn, 2016). La clusterisation des dates/sites a été effectuée avec la méthode hiérarchique de Ward basée sur la variance minimale. **(2)** Pour déterminer le degré de couplage entre les variations environnementales et celles des abondances des protistes une analyse de Co-Inertie (COIA, Dolédec & Chessel 1994 ; Dray et al. 2003) a été effectuée avec le package « ade4 ». Elle consiste à trouver des

axes de co-inertie qui maximisent les covariances entre les coordonnées de lignes (dates) des matrices de variables environnementales et celles d'abondances des groupes taxonomiques ou trophiques. Elle a été calculée à partir d'une ACP normalisée sur les variables environnementales et une ACP centrée sur les abondances. Le degré de couplage a été évalué avec le coefficient de corrélation multidimensionnelle (RV) et en effectuant un test de Monte-Carlo. Les données environnementales et d'abondance ont été transformées en  $\log(x+1)$  afin de minimiser l'influence des valeurs élevées et pour assurer une distribution normale (Legendre & Legendre 1998). Enfin, les données environnementales ont été normalisées afin de les rendre sans dimension (z-scores, Kenkel 2006). La significativité statistique a été fixée à  $p < 0.05$ .

## RESULTATS

### RESULTATS DES DONNEES MOLECULAIRES

#### *Les variables environnementales*

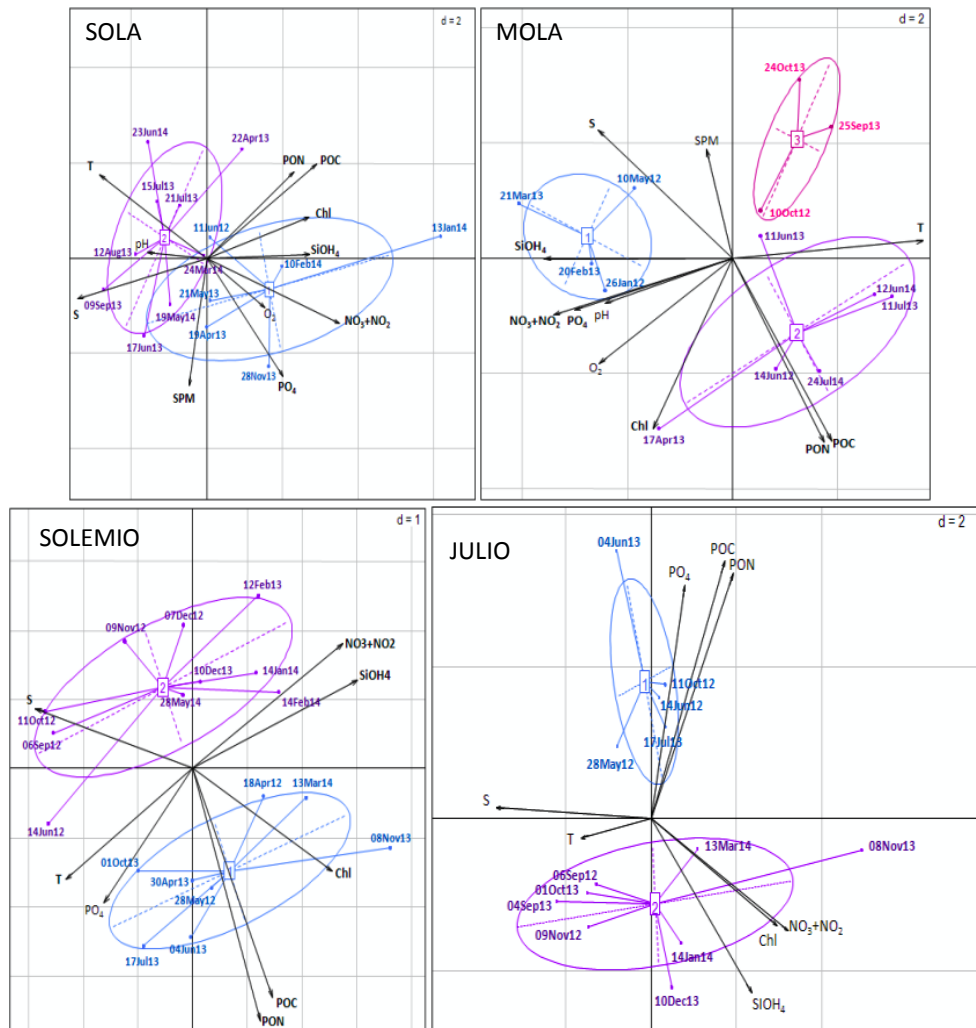
L'analyse des paramètres environnementaux a révélé que la température et la salinité les plus élevées ont été enregistrées respectivement à SOLEMIO en Juillet 2013 (24.6°C) et à SOLA en Juin 2013 (38.49 PSU), tandis que les plus basses ont été enregistrées à la station SOLA, en Novembre 2013 (12.2°C) et en Janvier 2014 (34.27 PSU). Cette salinité basse coïncidait avec les concentrations les plus élevées de  $\text{NO}_3^- + \text{NO}_2^-$ ,  $\text{SiOH}_4$  et Chl *a* (10.24  $\mu\text{M}$ , 10.99  $\mu\text{M}$  et 1.75  $\mu\text{g L}^{-1}$ , respectivement, **Tableau 2**). Pour  $\text{PO}_4^{3-}$ , la concentration la plus élevée a été enregistrée à la station JULIO atteignant 0.33  $\mu\text{M}$  en Octobre 2012. A la même station, des valeurs élevées en POC et en PON ont été observées en Juin 2013 (28.67  $\mu\text{mol L}^{-1}$  et 3.13  $\mu\text{mol L}^{-1}$ , respectivement). En général et pour l'ensemble des paramètres, la station SOLA se distinguait par sa forte salinité et sa richesse en Chl *a* (test de Kruskal-Wallis,  $p < 0.05$ ). En ce qui concerne les paramètres supplémentaires mesurés uniquement à SOLA et à MOLA, les moyennes calculées ne montraient pas de différence significative entre les deux stations. Les concentrations maximales en oxygène dissous (6.18  $\text{mL.L}^{-1}$  en Mai 2013) et en matière en suspension (16.66  $\text{mg.L}^{-1}$  en Juin 2013) ont été observées également à la station SOLA (**Tableau 2**).



Les analyses d'ACP effectuées ont montré que les  $\text{NO}_3+\text{NO}_2$ ,  $\text{SiOH}_4$ , POC, PON et la Chl *a* contribuaient à >10 % à la formation des axes à toutes les stations. La salinité contribuait à >15 % à toutes les stations sauf JULIO. (**Fig.3, Tableau 3B**). Cependant, la contribution des variables biologiques (POC, PON et Chl *a*) a été en moyenne deux fois plus faible à SOLA ( $12\pm 2$ ) qu'aux trois autres stations ( $24\pm 6$ ). De ce fait, la matière particulaire en suspension (MES ou SPM) formait ainsi le principal contributeur (26%) à la variabilité saisonnière. Une autre différence marquée entre les stations a été le découplage du  $\text{PO}_4$  des deux autres nutriments principaux ( $\text{NO}_3+\text{NO}_2$ ,  $\text{SiOH}_4$ ) aux stations situées dans la baie de Marseille (**Fig.3**). La hiérarchisation des données temporelles avec la méthode Ward (HAC), qui expliquait au total plus de 50% de l'inertie totale (**Tableau 3A**), a permis de distinguer trois périodes saisonnières à MOLA contre seulement deux pour le reste des stations (**Fig.3**). En conséquence, l'automne (Septembre et Octobre) à MOLA, caractérisée par des eaux relativement chaudes, mais riches en sels nutritifs et SPM, se distinguait nettement des deux autres saisons, i.e. hiver-printemps et été. Ces deux périodes ont été également produites à JULIO. Pour les stations côtières (SOLA et SOLEMIO), la saisonnalité s'est montrée moins marquée par rapport aux stations du large et est caractérisée par deux saisons regroupant plus ou moins hiver-printemps et été-automne (**Fig.3**). Et quoique MOLA et JULIO aient montré des périodes semblables, les variables environnementales responsables des tendances saisonnières différaient entre elles. En conséquence, la période hiver-printemps à MOLA était principalement caractérisée par des concentrations relativement élevées en nutriments mais des températures faibles, tandis que celle à JULIO était principalement caractérisée par un faible POC et PON et un  $\text{PO}_4$  relativement faible malgré une teneur élevée en  $\text{SiOH}_4$  et  $\text{NO}_3+\text{NO}_2$  (**Tableau 3B**). Des différences entre ces deux stations ont été également observées en été.

	SOLA	MOLA	SOLEMIO	JULIO
T (°C)	16.69±3.9	17.09±3.16	17.07±3.39	19.01±3.48
	12.2-23.22	12.37-21.05	12.75-24.6	14.3-24.19
S (PSU)	37.54±0.94	37.96±0.34	38.06±0.28	38.01±0.49
	34.27-38.49	37.32-38.37	37.12-38.35	36.44-38.41
O <sub>2</sub> (ml.L <sup>-1</sup> )	5.35±0.64	5.40±0.29	NA	NA
	3.88-6.18	4.97-5.89		
pH	8.17±0.12	8.25±0.12	NA	NA
	7.96-8.44	8.08-8.43		
NO <sub>3</sub> +NO <sub>2</sub> (μM)	1.13±2.6	0.54±0.95	0.62±0.53	0.5±0.95
	0.03-10.24	0.03-3.14	0.09-1.95	0.08-3.6
PO <sub>4</sub> (μM)	0.03±0.02	0.03±0.02	0.08±0.1	0.06±0.09
	0.01-0.09	0.01-0.07	0.01-0.45	0.01-0.33
SiOH <sub>4</sub> (μM)	1.93±2.65	1.38±0.71	1.34±0.76	1.28±1.16
	0.11-10.99	0.13-2.4	0.46-3.58	0.33-4.83
POC (μmol.L <sup>-1</sup> )	8.94±3.85	6.78±3.05	9.37±3.76	9.19±6.33
	4.35-18.2	3.77-14.55	4.6-18.38	4.42-28.67
PON (μmol.L <sup>-1</sup> )	1.16±0.41	0.84±0.3	0.94±0.48	0.91±0.71
	0.5-1.97	0.42-1.35	0.46-2.34	0.42-3.13
SPM (mg.L <sup>-1</sup> )	2.98±3.96	1.65±1.14	NA	NA
	0.37-16.66	0.3-3.52		
Chl (μg.L <sup>-1</sup> )	0.54±0.54	0.21±0.17	0.44±0.25	0.35±0.34
	0.05-1.75	0.03-0.67	0.04-1.05	0.02-1.35

**Tableau 2 :** Liste des moyennes (± écart type) ainsi que les valeurs minimales et maximales des différentes variables environnementales mesurées au niveau de chaque station (moyenne ± écart type, min-max). NA : Not available (non disponible)



**Figure 3 :** Analyses en composante principale (ACP) appliquées sur les variables environnementales enregistrées aux stations méditerranéennes (15, 13, 18 et 13 dates à SOLA, MOLA, SOLEMIO et JULIO, respectivement). L'ACP a été réalisé en utilisant le « package ade4 ». Les groupes ont été réalisés par utilisation du « package caret » appartenant au logiciel R (v 3.2.3). (Voir aussi tableau 2)

(A)	SOLA		MOLA		SOLEMIO		JULIO	
	Eig.	Inertia (%)	Eig.	Inertia (%)	Eig.	Inertia (%)	Eig.	Inertia (%)
Axe 1	4.33	39.42	4.66	42.44	2.69	33.74	2.98	37.32
Axe 2	1.76	16.02	2.71	24.67	2.40	30.08	2.55	31.89
Axe 3	1.56	14.18	1.40	12.77	1.03	12.97	1.19	14.94
Axe 4	1.19	10.86	0.93	8.53	0.82	10.37	0.57	7.15
Axe 5	0.94	8.59	0.50	4.59	0.42	5.30	0.39	4.92

(B)	SOLA		MOLA		SOLEMIO		JULIO	
	Axe1	Axe2	Axe1	Axe2	Axe1	Axe2	Axe1	Axe2
T	11.42	10.83	17.63	0.21	12.53	6.68	5.61	0.16
S	16.65	2.55	8.72	11.38	19.09	1.98	27.50	0.05
O <sub>2</sub>	3.4	3.89	8.64	7.75	NM	NM	NM	NM
pH	3.5	0.05	7.89	1.41	NM	NM	NM	NM
NO <sub>2</sub> +NO <sub>3</sub>	17.76	6.74	15.59	2.24	17.55	8.44	18.26	4.85
PO <sub>4</sub>	5.9	22.03	12.17	1.88	6.10	9.75	1.29	23.17
SiOH <sub>4</sub>	10.64	0.02	17.14	0	21.00	4.29	11.75	12.91
POC	12.10	13.97	4.79	23.23	4.95	28.45	6.20	28.06
PON	7.67	11.77	4.07	23.59	3.59	34.65	7.75	25.44
SPM	0.3	25.55	0.33	8.25	NM	NM	NM	NM
Chl	10.61	2.59	3.02	20.07	15.19	5.76	21.64	5.35

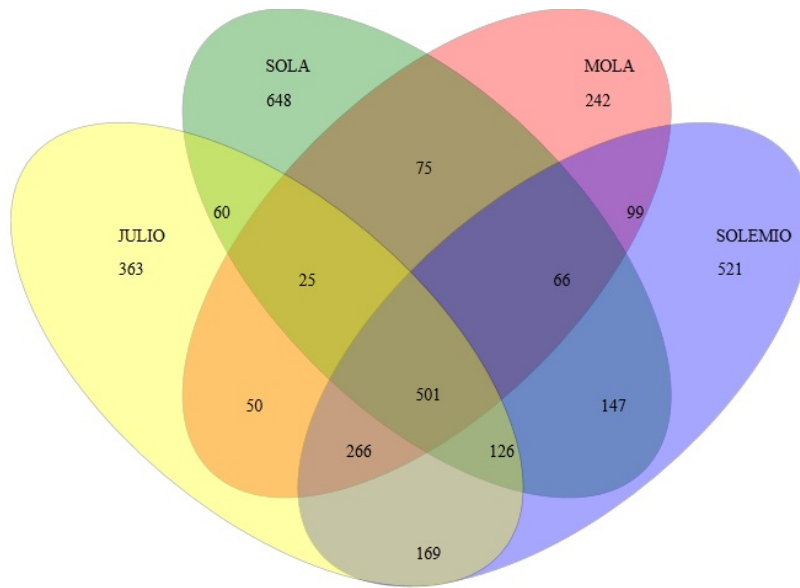
**Tableau 3** : Résultats des analyses en composante principale appliquées sur les paramètres physico-chimiques (voir fig. 3) mesurés au niveau des différentes stations (SOLA, MOLA, SOLEMIO et JULIO). (A) Valeurs propres (Eigenvalues) et inerties projetées (%) des cinq premiers axes, (B) Contributions de chaque variable environnementale dans la formation des deux premiers axes de l'ACP. En bleu : fortes contributions des variables environnementales. NM : non mesuré.

### *Diversité et distribution des protistes*

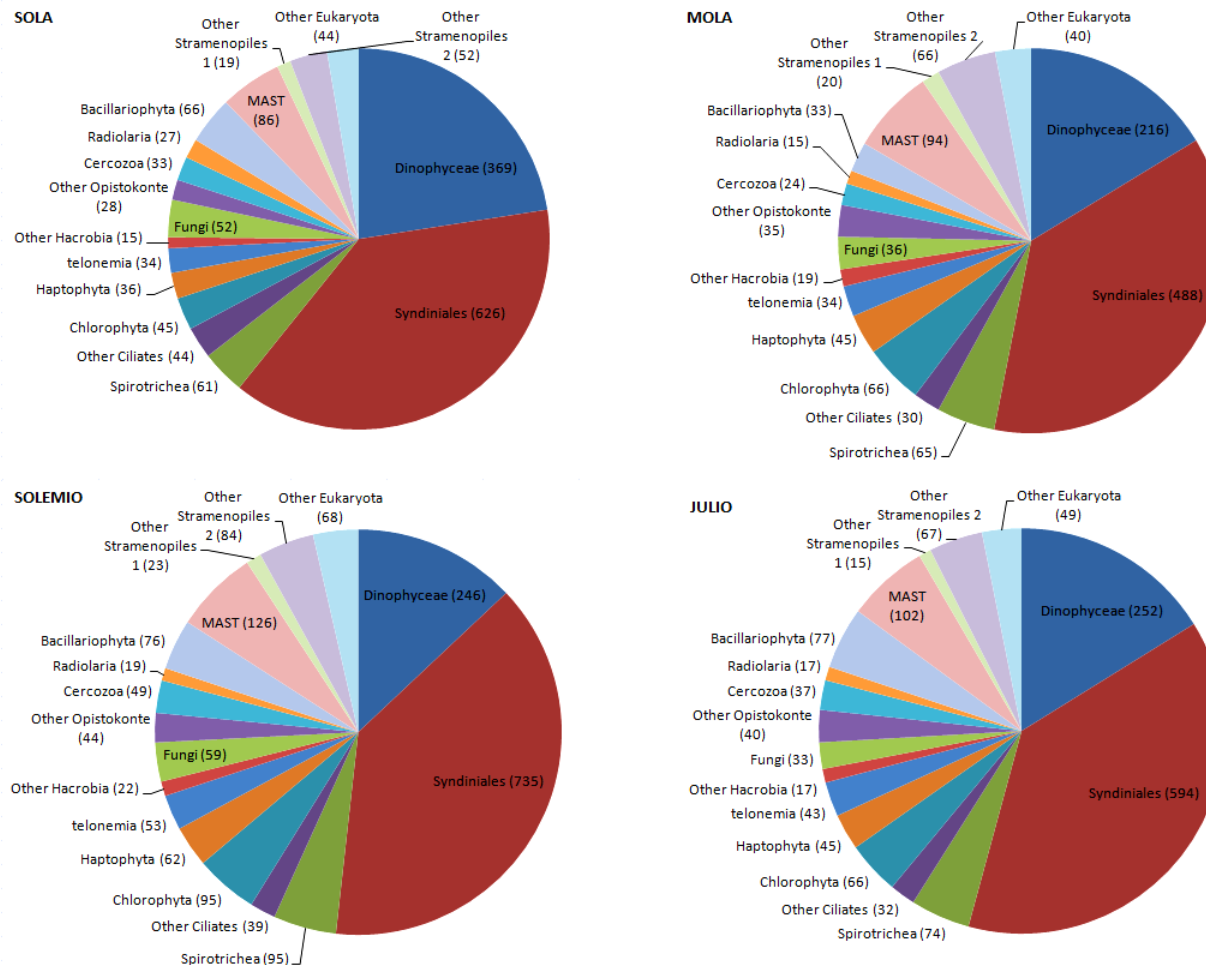
Après la normalisation, un total de 248 788 séquences a été généré à partir du séquençage de tous les échantillons. Un total de 3358 OTU a été identifié dans toutes les stations (1648, 1324, 1895 et 1560 OTU pour SOLA, MOLA, SOLEMIO et JULIO, respectivement) avec 665 OTU et 1062 OTU partagées entre SOLA & MOLA et SOLEMIO & JULIO, respectivement (**Fig.4**). Les 3358 OTU ont été réparties en sept super-groupes (Alveolata, Apusozoa, Archeplastida, Hacrobia, Opisthokonta, Rhizaria et Stramenopiles) et d'autres

eucaryotes, et affiliées à 48 groupes taxonomiques supérieurs (voir **Annexe 1**). Dans l'ensemble, si nous classons notre diversité en grands groupes taxonomiques (**Fig.5**), aucune différence significative, en termes de distribution de nombre d'OTU, n'est détectée entre les stations séparées par région (Kolmogorov-test,  $p_{value} > 0.05$  pour SOLA & MOLA et pour SOLEMIO & JULIO), et pareillement pour l'ensemble des stations (Kruskal-Wallis,  $p_{value} = 0.34$ ). Le test Kruskal-Wallis appliqué sur le nombre de séquences des grands groupes (**Fig.6**) n'a également montré aucune différence dans la distribution des nombres de séquences dans l'ensemble des stations.

En général, les Syndiniales dominaient la composition taxinomique des 4 stations (626 OTU, 488 OTU, 735 OTU, 488 OTU pour SOLA, MOLA, SOLEMIO et JULIO, respectivement) (**Fig.5**) représentant entre 30% et 50% du total des séquences générées (32%, 50%, 40% et 46% pour SOLA, MOLA, SOLEMIO et JULIO, respectivement) (**Fig.6**), suivis par les Dinophyceae aux stations SOLA et JULIO (29% et 13%, respectivement) et par les chlorophycées à MOLA et SOLEMIO (12% et 14%, respectivement) (**Fig.6**). Dans les Dinophyceae, les données moléculaires ont enregistré une présence conséquente des *Dinophyceae sp. Clade 87* et de *Gymnodinium* en Mai 2013 et 2013 à SOLA, SOLEMIO et JULIO. Pour les ciliés, ils ne représentaient que 1-2% du total des séquences dans l'ensemble des stations (**Fig.6**) marqués par une apparition considérable de *Laboea strobila* en Mai-Juin 2013 à JULIO. Les groupes Stramenopiles les plus représentatifs, en termes de richesse, étaient les MAST (Marine Stramenopiles) suivis par les bacillariophycées pour SOLA et MOLA et inversement pour SOLEMIO et JULIO. Nos données moléculaires ont révélé également une présence considérable de *Pseudonitzschia sp.* en Juin-Juillet dans les stations SOLA et JULIO et de *Chaetoceros sp.* en Juin 2013 à SOLEMIO. Les groupes les moins importants en termes de nombre de séquences et d'OTU étaient Cercozoa, Chlorophytes et Fungi (**Fig.5 et 6**).

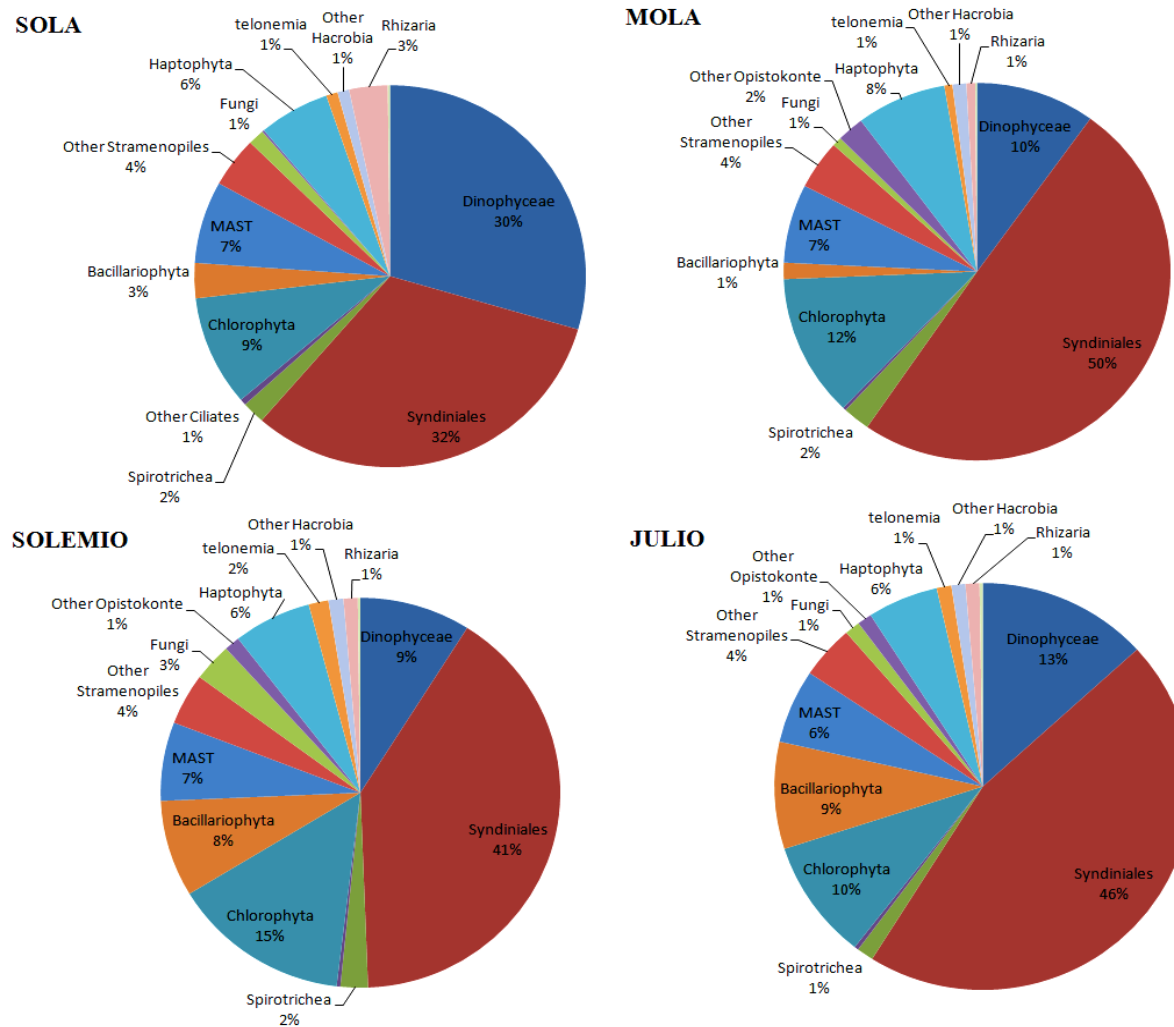


**Figure 4** : Le schéma représente un digramme de Venn de la diversité partagée entre les différentes stations. La richesse totale de l'ensemble des stations est de 3358 OTU. Le nombre d'OTU identifiées est de 1648, 1324, 1895 et 1560 à SOLA, MOLA, SOLEMIO et JULIO, respectivement. Le nombre d'OTU partagées entre SOLA & MOLA et SOLEMIO & JULIO est de 667 et 1062 OTU, respectivement.



**Figure 5 :** Le nombre d'OTU des grands groupes taxonomiques à chaque station.

Other Stramenopiles 1: Bicoecia, Pirsonia, Hyphochytriomyceta, Labyrinthulea, Labyrinthulomycetes; Other Stramenopiles 2: Chrysomerophyceae, Chrysophyceae-Synurophyceae, Dictyochophyceae, Eustigmatophyceae, Bolidophyceae and Relatives, Ochrophyta, Pelagophyceae, Phaeophyceae, Pinguiophyceae, MOCH, Raphidophyceae



**Figure 6 :** L'abondance relative des grands groupes taxonomiques à chaque station. Pour faciliter la lecture, le pourcentage des séquences attribué à un groupe spécifique est indiqué dans les graphiques que lorsque celui-ci était supérieur à 1%.



### ***La variabilité spatiale et structure de la communauté***

L'analyse COIA a été réalisée pour tester la corrélation entre 8 à 11 variables environnementales (température, salinité, NO<sub>3</sub>NO<sub>2</sub>, PO<sub>4</sub>, SiOH<sub>4</sub>, PON, POC et Chl *a* mesurées pour les quatre stations, en plus des SPM, O<sub>2</sub> et pH mesurés uniquement à SOLA et MOLA) et les dix super-groupes taxonomiques (Dinophyceae, Syndiniales, Ciliés, Chlorophyta, Haptophyta, Champignons, Cercozoa, Radiolaria, Bacillariophyta et MAST). Les sous-groupes, utilisés dans la description de la diversité des protistes, ont été assemblés en super-groupes pour une meilleure lisibilité des diagrammes COIA.

Les analyses de co-inertie effectuées sur les dix super-groupes taxonomiques (précédemment cités) ont montré un coulage significatif entre les conditions environnementales saisonnières et les abondances des super-groupes au sein de chaque station à l'exception de JULIO ( $p_{value} = 0,008$ ,  $p_{value} = 0,026$ ,  $p_{value} = 0,001$ , pour SOLA, MOLA et SOLEMIO, respectivement (**Tableau 4A**). Les deux premiers axes ont expliqué 86%, 64% et 62% de la variance totale à SOLA, MOLA et SOLEMIO, respectivement (**Tableau 4A**). Au niveau de ces stations, dont le couplage s'est montré significatif, la température, NO<sub>3</sub>NO<sub>2</sub>, PO<sub>4</sub>, SiOH<sub>4</sub> et PON ont été les variables les plus importantes dans la structuration de la dynamique des groupes taxonomiques (**Tableau 4B**). D'autres contribuaient également à cette structuration taxonomique: salinité, pH et Chl *a* chez SOLA et MOLA, SPM chez SOLA, O<sub>2</sub> chez MOLA et POC chez SOLEMIO (**Tableau 4B**).

Du côté des groupes taxonomiques, les ciliés, les chlorophycées et les radiolaires ont montré une forte contribution à la formation des axes COIA des trois stations, suivis par Syndiniales, Haptophyta et Cercozoa dans les stations côtières (SOLA et SOLEMIO), Dinophyceae dans les stations situées à Banyuls (SOLA et MOLA), Bacillariophyta et MAST chez MOLA et SOLEMIO, et finalement Fungi chez SOLA (**Tableau 4C**). Les radiolaires dominaient dans les eaux relativement froides qu'elle que soit la station (**Annexe 2**). De plus, ce super-groupe était associé à des teneurs élevées en SPM et en Chl *a* mais faibles en PON avec un pH bas aux stations situées dans la partie Ouest du GoL (SOLA, MOLA, **Annexe 2a et b**) alors qu'à SOLEMIO, les radiolaires étaient plutôt associées à des teneurs élevées en SiOH<sub>4</sub> et NO<sub>3</sub>NO<sub>2</sub> mais faibles en PO<sub>4</sub> (**Annexe 2c**). Pour les Syndiniales, ils présentaient les mêmes associations avec les variables environnementales sauf à nouveau à SOLEMIO (**Annexe 2c**). Ce super-groupe était plutôt associé à des eaux pauvres en POC et en PON. De la même

manière, les super-groupes restants présentaient des associations différentes avec les variables environnementales à SOLEMIO contrairement aux stations SOLA et MOLA. Par conséquent, un couplage significatif a été observé entre les stations SOLA et MOLA contrairement à SOLEMIO et JULIO, où les stations ont évolué indépendamment l'une de l'autre et également de celles de la partie Ouest du GoL. A SOLA et MOLA, les Dinophyceae ont évolué de manière saisonnière à l'opposé de Radiolaria et Syndiniales (**Annexe 2a et b**). De plus, des dynamiques saisonnières de Bacillariophyta et de Chlorophyta ont été associées, mais opposées à celles de Cercozoa, des ciliés, d'Haptophyta et de MAST (**Annexe 2a et b**). Par conséquent, contrairement à ces trois derniers super-groupes, Bacillariophyta et Chlorophyta dominaient lorsque les eaux étaient riches en nutriments et relativement peu salées (**Annexe 2a et b**). À la station SOLEMIO, Chlorophyta était également associée à des concentrations élevées de nutriments, sauf pour  $PO_4$  (**Annexe 2c**). Comme à l'Ouest du GoL, les ciliés étaient opposés aux diatomées (**Annexe 2c**).

(A)	p value	RV	% Axis 1	% Axis 2
SOLA	<b>0.008</b>	0.51	76.06	10.11
MOLA	<b>0.029</b>	0.69	69.62	17.39
SOLEMIO	<b>0.001</b>	0.55	48.21	37.80
JULIO	0.062	0.53	58.44	26.77

(B)	SOLA		MOLA		SOLEMIO		JULIO	
	Axe1	Axe2	Axe1	Axe2	Axe1	Axe2	Axe1	Axe2
T	9.28	16.74	18.79	2.88	0.35	20.84	7.13	26.7
S	13.61	2.97	13.42	13.24	0.01	5.86	1.89	11.16
O <sub>2</sub>	0.3	0.12	7.07	11.32	NM	NM	NM	NM
pH	0.24	27.22	6.85	2.21	NM	NM	NM	NM
NO <sub>2</sub> +NO <sub>3</sub>	15.93	1.43	11.49	2.57	1.22	24.73	1.89	16.43
PO <sub>4</sub>	16.17	0.23	11.28	0	1.98	24.55	4.98	2.65
SiOH <sub>4</sub>	15.58	0.32	17.98	0.94	0.32	22.21	4.96	14.2
POC	10.38	6.79	5.04	5.5	45.38	0.51	40.21	1.5
PON	6.94	17.38	5.9	18.86	45.53	0.26	35.43	2.77
SPM	1.38	15.89	0.94	0.03	NM	NM	NM	NM
Chl	10.10	10.91	1.14	42.45	5.21	1.03	3.51	24.59

(C)	SOLA		MOLA		SOLEMIO		JULIO	
	Axe1	Axe2	Axe1	Axe2	Axe1	Axe2	Axe1	Axe2
Dinophyceae	0.63	39.48	0	26.86	0.17	0.03	13.94	7.88
Syndiniales	0.24	13.79	0	4.07	12.72	0.84	19.66	3.11
Ciliates	11.58	1.30	34.04	0.9	21.7	2.13	0.07	3.07
Chlorophyta	19.82	18.83	24.71	2.07	0.35	21.5	7.94	2.71
Haptophyta	11.58	0.12	8.89	2.45	11.02	26.33	2.14	9.25
Fungi	10.04	1.96	0.06	8.28	3.1	6.84	3.79	0.05
Cercozoa	27.11	3.70	0.92	0.15	10.11	10.69	0.3	10.61
Radiolaria	7.44	20.63	22.39	3.93	2.67	23.14	5.93	58.63
Bacillariophyta	9.87	0.10	1.13	28.71	22.81	0.16	19.03	4.6
MAST	1.69	0	7.86	22.49	15.35	8.34	27.2	0.09

**Tableau 4:** Analyses de Co-inertie (PCA-PCA COIA) entre les variables environnementales et les groupes taxonomiques aux 4 stations. (A) La  $p_{value}$  (test de randomisation avec 1000 permutations) renseigne sur la significativité de la relation entre les 2 tableaux dans chaque cas. La corrélation entre les tables est donnée par le coefficient RV. "Pourcentage de co-

variation" s'explique par les deux premiers axes de co-inertie (notés % Axe 1 et % Axe 2 respectivement) (B) Contributions de chaque variable environnementale sur les deux premiers axes des analyses COIA en %. En bleu, les fortes contributions des variables environnementales (>10%). NM: non mesuré. (C) Contributions de chaque groupe taxonomique sur les deux premiers axes des analyses COIA en %. En bleu, les fortes contributions des groupes taxonomiques (>10%).

## RESULTATS DES COMPTAGES MICROSCOPIQUE

Comme précédemment mentionné, la composition du microzooplancton (dinoflagellés et ciliés) a été également étudiée, en complément des données obtenues à partir du gène 18S ARN-Illumina, grâce à des observations morphologiques. Le choix des échantillons à étudier s'est basé sur les dates des échantillons séquencés (**Tableau 1**).

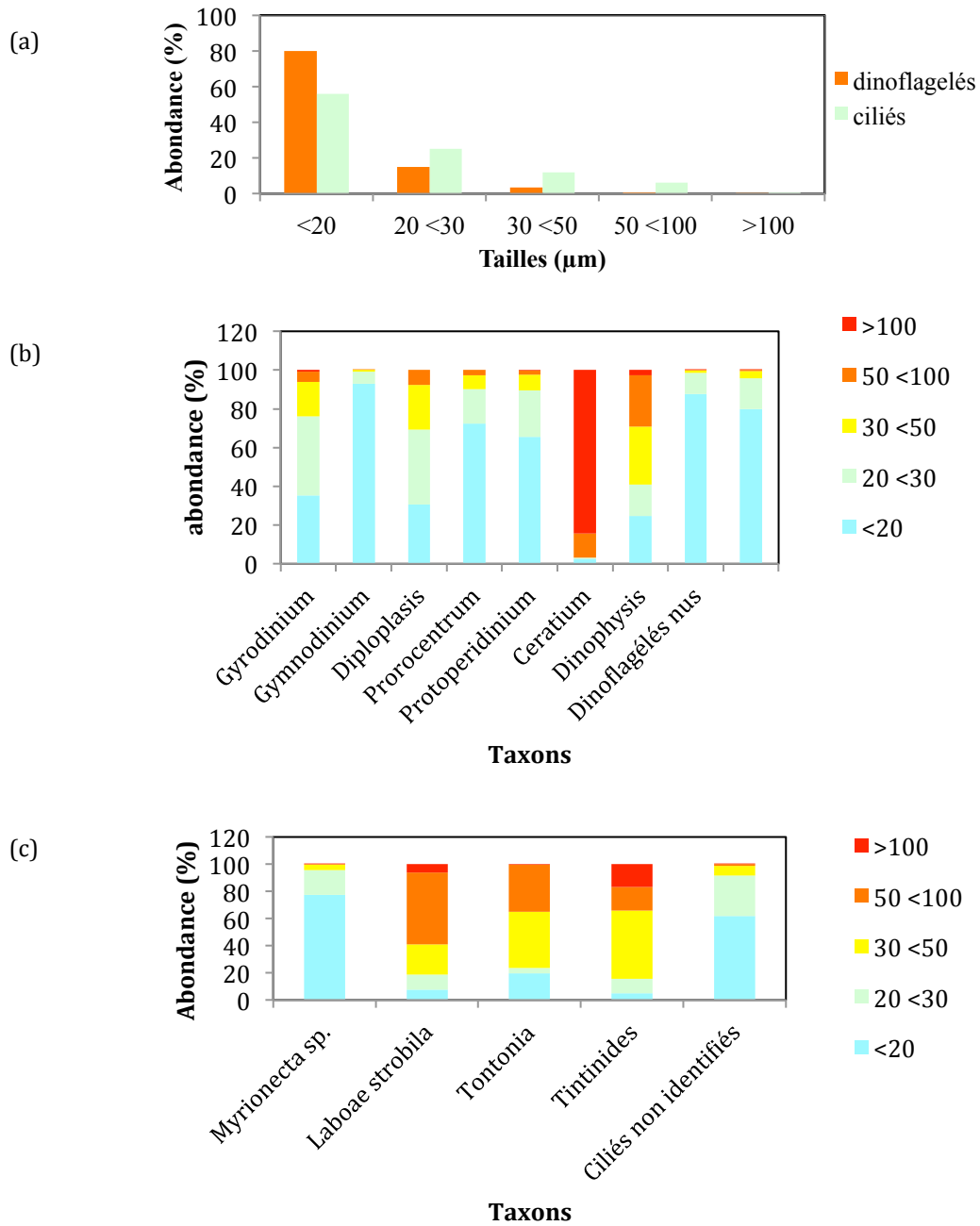
En microscopie, 14 morphotypes principaux souvent présents dans les quatre stations ont été dénombrés (7 taxons pour les dinoflagellés et 4 pour les ciliés, **Annexe 3**). Les autres morphotypes inidentifiables ou présentant une faible abondance et/ou rares (très peu de cellules) ont été classés dans deux grands groupes pour les dinoflagellés (dinoflagellés nus et enveloppés) contre un groupe unique pour les ciliés (oligotriches) (**Annexe 3**). Un « guide » des taxons observés dans l'ensemble des échantillons traités a également été effectué et présenté à la fin de ce chapitre (**Annexe 4**).

Dans cette partie, les dinoflagellés et les ciliés ont été classés par classe de taille (**Annexe 3**). Dans l'ensemble des stations, les observations microscopiques ont montré une dominance des dinoflagellés par rapport aux ciliés (**Fig.7 & 8**). Pour les dinoflagellés comme pour les ciliés, les cellules de petite taille (<20µm) dominaient la composition microzooplanctonique et occupaient 80% de l'abondance totale du microzooplancton (**Fig.7a**). Pour les dinoflagellés, seul le genre *Ceratium* présentait une taille supérieure dépassant les 100µm (**Fig.7b**) quant aux ciliés, la taille dépassait les 50µm pour *Laboea strobila* et *Tontonia sp.* et les 100µm pour *Tintinnide sp.* (**Fig.7c**).

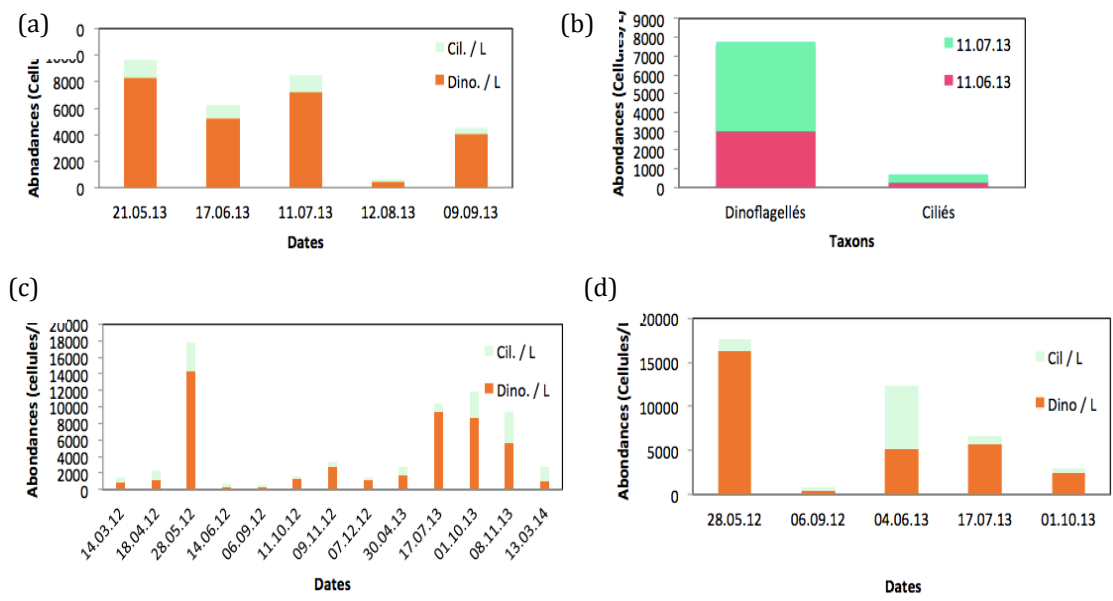
D'après les données obtenues par microscopie, l'abondance des dinoflagellés était très importante pendant la période estivale (Mai - Juillet) dans les stations SOLA, SOLEMIO et JULIO et diminuait pendant la période Automne-Hiver (**Fig.8**). Le

nombre limité d'échantillons observés à MOLA, ne permet pas de confirmer cette tendance pour cette station même si l'on remarque une augmentation d'abondance des dinoflagellés allant de Juin à Juillet (**Fig.8**). A la station SOLA, en période estivale (Mai - Juillet), les micro-brouteurs (*Gyrodinium sp.*) et les mixotrophes (*Protoberidinium sp.*, *Myrionecta sp.*, *Laboea strobila* et *Tontonia sp.*) dominaient en termes d'abondance (**Fig.9**). Pour les échantillons MOLA, *Gyrodinium sp.*, *Prorocentrum sp.* et *Protoberidium sp.*, étaient abondamment présents pendant l'été (Juin-Juillet 2013) en plus de *Myrionecta sp.*, et de *Tontonia sp.*. En Juillet 2013, les Tintinides dominaient la population des ciliés (**Fig.9**). Pour SOLEMIO, en Mai 2012, *Gyrodinium sp.* est devenu 18 fois plus abondant par rapport à la période printanière (Mars-Avril 2012) accompagné de *Prorocentrum sp.*, *Protoberidinium sp.*, *Myrionecta sp.* et *Tontonia sp.*. Pendant la période automnale (Octobre-Novembre 2012), *Gyrodinium sp.* et *Gymnodinium sp.* dominaient la composition des dinoflagellés. Alors qu'en Octobre 2013, *Protoberidinium sp.*, *Prorocentrum sp.* et *Myrionecta sp.* ont été les plus fréquemment observés (**Fig.9**). A la station JULIO, pendant Mai 2012 & Juin 2013, *Tontonia sp.* était très abondante en plus de *Prorocentrum sp.* et *Protoberidinium sp.*. Au moment des observations, un bloom de *Pseudonitzschia sp.* a été observé en Mai 2012 (**Fig.10**).

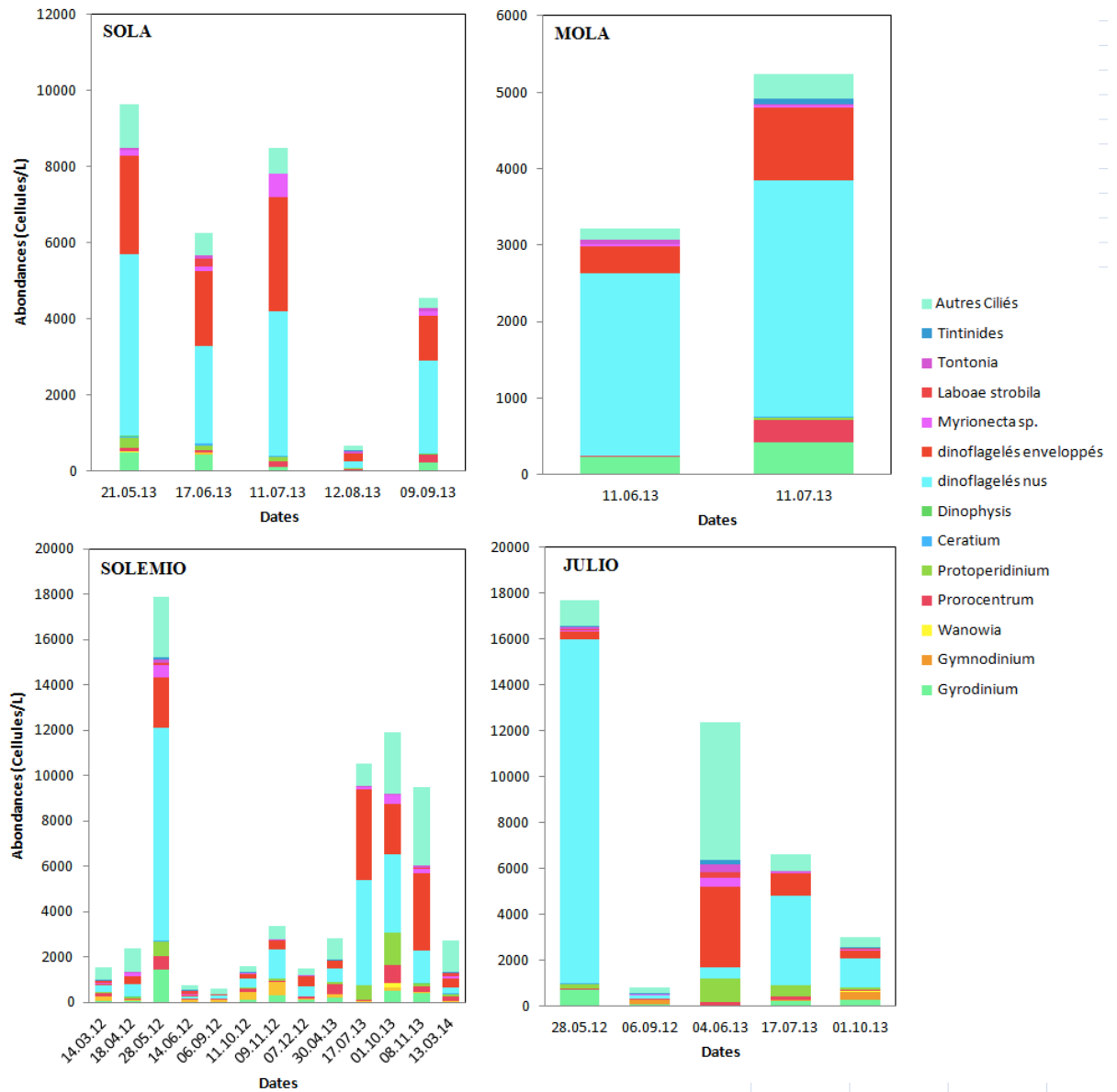
En conclusion, les dinoflagellés dominaient la composition du microzooplancton dont la taille prépondérante est celle <20µm. Les organismes mixotrophes et micro-brouteurs ont été les plus fréquemment observés. Les abondances importantes des différents taxons ont été enregistrées pendant la période estivale (Mai - Juillet).



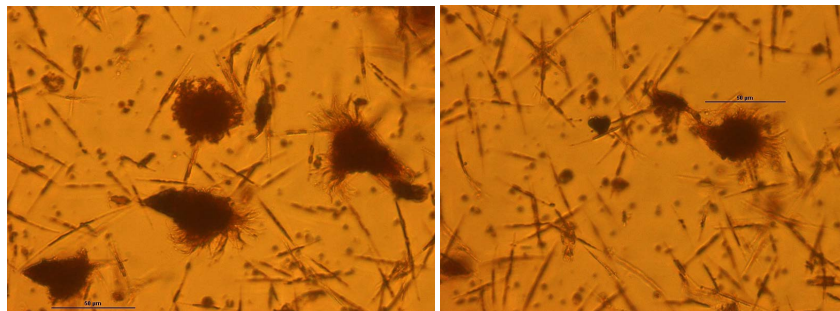
**Figure 7 :** Pourcentage d'abondance des taxons présents dans l'ensemble des stations (SOLA, MOLA, SOLEMIO et JULIO) classés par facteur de taille. (a) composition générale des dinoflagellés et ciliés classés par facteur taille. (b) abondance par taille des différents taxons appartenant aux dinoflagellés. (c) abondance par taille des différents taxons appartenant aux ciliés.



**Figure 8:** Abondances (exprimées en cellules/Litre) des dinoflagellés et des ciliés dans les quatre stations méditerranéennes (a) SOLA, (b) MOLA, (c) SOLEMIO et (d) JULIO). Les dates mentionnées dans chaque graphique représentent les échantillons qui ont fait l'objet des observations microscopiques et qui correspondaient à ceux séquencés.



**Figure 9:** Abondances (exprimées en cellules/Litre) des différents morphotypes appartenant aux dinoflagellés et aux ciliés dans les quatre stations méditerranéennes (SOLA, MOLA, SOLEMIO, JULIO). Les dates mentionnées dans chaque graphique représentaient les échantillons qui ont fait l'objet des observations microscopiques et qui correspondaient à ceux séquencés.



**Figure 10:** Aperçu du bloom de *Pseudonitzschia sp.* observé à la station JULIO en surface durant Mai/Juin 2012 accompagné d'une haute présence de *Tontonia sp.* .



## DISCUSSION ET CONCLUSIONS

### *Aspect général*

Etant considérée comme une mer semi-fermée, en grande partie oligotrophique, avec des eaux chaudes et relativement salées, pauvres en éléments nutritifs (Duarte *et al.*, 1999), la Méditerranée a un aspect intéressant concernant la coexistence de processus océaniques particuliers, comme la formation d'eau dense observée en hiver dans le bassin occidental (Béthoux & Gentili, 1999), et des caractéristiques côtières typiques, comme la pression anthropique marquée favorisant l'eutrophisation (Duarte *et al.*, 1999). Dans cette mer, les tendances à l'augmentation de la température et de la salinité de l'eau sont observées depuis quelques décennies (Béthoux *et al.*, 1990 ; Rohling & Bryden, 1992 ; Béthoux *et Gentili*, 1996 ; Johnson, 1997 ; Marbà & Duarte, 1997). Les concentrations de nutriments, en particulier le phosphate et les nitrates, ont connu une augmentation depuis le début des années 1960, probablement en raison des apports atmosphériques et terrestres, tandis que la concentration en silicate est restée relativement constante (Béthoux *et al.*, 2002). De par ses particularités hydrologiques, la Méditerranée forme un bon lieu d'études des interactions biotiques et abiotiques.

Notre étude, qui s'est conduite en Méditerranée occidentale, s'est intéressée particulièrement à l'étude des communautés protistes et à l'impact des variables environnementales sur leur variabilité saisonnière. Pour ceci, le Séquençage de Nouvelle Génération (NGS) a été utilisé pour déterminer la composition de la communauté planctonique par séquençage de l'ARN 18S de quatre stations soumises aux effets du Rhône. Ce dernier représente une source importante d'éléments nutritifs pour le Golfe du Lion dans le nord-ouest de la Méditerranée, influençant grandement sa productivité (Lefèvre *et al.*, 1997). Dans le nord-ouest de la Méditerranée, un certain nombre d'études ont abordé la composition et la dynamique du microplancton à long terme en utilisant des méthodes classiques (e.g. Gómez *et Gorsky*, 2003; Pereira Quiroga, 2005; Tunin Ley *et al.*, 2007).

Les études basées sur l'ADN ont grandement amélioré notre compréhension de la diversité en fournissant une évaluation complète comprenant également les taxons « invisibles » (petite taille, symbiotes etc.) et rares (e.g. Lopez-Garcia *et al.*, 2001 ;

*Medinger et al., 2010 ; Behnke et al 2010 ; Caron et al., 2012 ; de Vargas et al., 2015 ; Massana et al. 2015 ; Genitsaris et al., 2015, 2016*). Dans cette étude, l'analyse génomique nous a permis de réaliser la liste la plus exhaustive possible de la diversité sur une période de deux ans (2012-2014) et d'évaluer sa variabilité au travers les saisons.

### ***L'environnement nutritif***

Selon les données SOMLIT (Service d'Observation en Milieu Littoral, <http://somlit.epoc.u-bordeaux1.fr/fr/>), les stations situées à Marseille (SOLEMIO & JULIO) et à Banyuls/mer (SOLA & MOLA) ont montré un changement progressif qui s'est traduit par une période de limitation accrue en phosphate en hiver de 1996 à 2007. Ceci s'est prolongé sur les périodes 2007-2011 (*Fraysse et al., 2013*) et 2012-2014 (**Tableau 2**). Pour les stations SOMLIT situées dans la baie de Marseille, les valeurs de température, de nutriments, de POC, de PON et de chlorophylle *a* des stations se situaient dans la même fourchette que celles trouvées dans la même zone durant la période 2007-2011 (*Fraysse et al., 2013*). Les concentrations les plus élevées en chlorophylle *a* ont été enregistrées à la station SOLA pendant la période hiver-printemps, plus particulièrement en 2014 où les concentrations ont doublé par rapport à l'année 2013. Ceci peut être lié au volume de la dérive des panaches fluviaux dont SOLA est soumise (*Hu et al., 2009*). Pour les Stations situées à Marseille, les concentrations de chlorophylle *a* atteignaient des niveaux comparables ou supérieures à ceux de Banyuls sur mer. Cette augmentation a été effectivement observé depuis 2007 (*Fraysse et al. 2013*). D'après les données SOMLIT, jusqu'en 2003 la chlorophylle *a* montrait des valeurs typiquement oligotrophes (généralement  $< 2 \mu\text{g L}^{-1}$ , pendant le bloom) pour augmenter progressivement en 2004 et atteindre en 2007 des niveaux similaires à ceux de SOLA et MOLA.

L'analyse en composante principale (ACP, **Fig.3**) a montré une saisonnalité un peu plus marquée aux stations situées au large, plus particulièrement MOLA contrairement aux stations côtières, SOLEMIO & SOLA où les dates étaient mélangées. Ceci peut être expliqué par les phénomènes d'intrusions induites par le panache du Rhône et son extension distale (*Hu et al. 2009*). JULIO est également influencé par la présence à proximité du Courant Nord assurant le transport des

nutriments (Hu et al., 2009). Les intrusions du Courant Nord induites peuvent influencer les taux en éléments nutritifs et ainsi masquer l'observation d'une saisonnalité nette. Fraysse et al., (2013) ont conclu également que le fonctionnement de l'écosystème de la Méditerranée du Nord (ex : la baie de Marseille) est complexe et fortement affecté par l'hydrodynamique et les apports terrestres.

### ***Diversité, succession saisonnière et structuration des communautés protistes***

Notre étude est une des premières à générer un aperçu général de la diversité des protistes microbiens eucaryotes dans la Méditerranée Nord-Ouest. La majorité des travaux jusqu'ici, généralement basés sur l'usage des techniques classiques, se sont plutôt intéressés à l'étude des genres ou groupes précis; en particulier des espèces indicatrices des changements environnementaux (Tunin-Ley et al., 2007,2009) et celles potentiellement toxiques (Harmful Algal Blooms) (ex, *Ostreopsis sp.* (Cohu et al., 2011; Blaufune et al., 2012)), ainsi qu'*Alexandrium sp.* (Bolch et al., 2007). Une étude récente (Penna et al., 2017) dans trois grandes régions méditerranéennes (Occidentale, Adriatique et Orientale) s'est intéressée à l'étude des assemblages phytoplanctoniques via le séquençage à haut débit. Penna et al., (2017) ont conclu que les assemblages phytoplanctoniques sont plus homogènes au niveau de la méditerranée orientale contrairement aux deux autres régions et que les trois écosystèmes sont significativement différents. Cette étude a montré également qu'au niveau des trois régions, les activités anthropiques n'ont pas de liens avec la diversité du phytoplancton.

Etant donnée les dates d'échantillonnage distantes et différentes d'une station à une autre, il n'était pas possible de comparer la diversité entre les stations pour chaque date. Nous avons donc regardé la composition des protistes dans l'ensemble et au niveau de chaque station. Le séquençage a révélé une diversité globale dominée par les alvéolaires majoritairement représentées par les Syndiniales (e.g. *Dino-group1-Clade1*, *Dino-group1-Clade7*) et les dinoflagellés (e.g. *Gyrodinium sp.*), respectivement (**Fig. 5 et 6**). Ces groupes présents dans l'ensemble des échantillons montraient une augmentation considérable en printemps (Avril-Mai-Juin) en nombre de séquences par rapport au restant de l'année. Pour les dinoflagellés, leur grande

abondance a été également confirmée dans les échantillons dénombrés par microscopie où par exemple *Gyrodinium sp.* était 18 fois plus abondant en Mai 2012 à SOLEMIO par rapport au mois précédent. D'après Fraysse et al. (2013), la plupart des intrusions se produisent entre Mai et Octobre, une période de l'année où la colonne d'eau est soit appauvrie, soit limitée en nutriments. Pendant cette période, les nutriments apportés par les intrusions du Rhône sont rapidement consommés, alimentant la production biologique locale. En revanche, les intrusions en hiver ne devraient pas avoir d'impact significatif sur la production primaire car les nutriments ne sont généralement pas limitants en raison du fort mélange vertical hivernal qui conduit à une limitation légère mais pas en nutriments (Fraysse et al., 2013). Les Syndiniales (e.g. Marines Alveolates, MALVI et MALVII) apparaissent dans tous les suivis de diversité marine (López-García et al., 2001 ; Massana et Pedrós-Alió, 2008 ; Georges et al., 2014 ; Genitsaris et al., 2015, 2016 ; Christaki et al. 2017). Il est admis que l'ensemble des Syndiniales (MALV) soit composé de parasites marins alors que leur abondance et leur diversité considérables suggèrent des interactions avec divers hôtes (Skovgaard et al., 2005, Massana et Pedrós-Alió, 2008). Quelques études ont montré que les Syndiniales interagissent avec les dinoflagellés ainsi que d'autres hôtes tels que les Radiolaires, les ciliés, les copépodes et également les poissons (Skovgaard et al., 2005 ; Harada et al., 2007 ; Guillou et al., 2008 ; Brate et al., 2012). Une étude récente sur les corrélations entre les symbiotes/décomposeurs/parasites et protistes eucaryotes a montré une corrélation positive avec les Dinophycées, plus particulièrement les *Gyrodinium sp.* (Christaki et al., 2017).

Un autre groupe considéré comme des hôtes de symbiotes et de parasites de dinoflagellés était présent dans les échantillons : les Radiolaires, représentés par > 15 OTU (à l'exception de SOLA qui comportait que 4 OTU, **Annexe 1**). Dans notre étude, les Radiolaires sont associés aux MALV (COIA, Annexe 2 a et c) ce qui peut soutenir l'hypothèse que les Radiolaires peuvent être des réservoirs importants de MALV (Brate et al., 2012).

Les ciliés, qui sont des brouteurs importants du phytoplancton de petite taille, n'avaient pas une grande abondance relative, comparés aux autres super-groupes, puisqu'ils ne représentaient que 1 à 2% d du nombre totale des séquences dans l'ensemble des stations (**Fig.6**). Par contre, ils représentaient un nombre relativement

élevé d'OTU (>60 OTU, **Fig.5, Annexe 1**) et ils étaient dominés par les mixotrophes. Ainsi, la composition taxinomique de la communauté des ciliés était dominée par les oligotriches nus (*Strombidium sp.*, *Laboea strobila*, *Tontonia sp.*) et l'haptoride *Myrionecta rubra* (**Fig.7**). Ceci a été observé aussi bien dans les comptages microscopiques que dans le jeu de données moléculaires. La présence élevée de mixotrophes au sein de la communauté des ciliés sont courants et ont déjà été signalés dans divers écosystèmes y compris la méditerranée NW (par exemple *Stoecker et al.*, 1987, 1999; *Sanders*, 1995; *Crawford & Lindholm*, 1997; *Modigh*, 2001 ; *Christaki et al.*, 2008, 2009 ; *Georges et al.*, 2014). Les mixotrophes, *Laboea strobila* et *Myrionecta rubra* sont des organismes connus pour être présents dans les eaux côtières les plus diverses (*McManus et Fuhrman*, 1986 ; *Stoecker et al.*, 1987 ; *Sanders* 1995 ; *Modigh* 2001). Ces ciliés disposent d'un double rôle de consommateur et de producteur, et des modèles de réseaux trophiques marins suggèrent que la mixotrophie entraîne une augmentation de l'efficacité trophique dans la mer (*Stickney et al.*, 2000).

En ce qui concerne les straménopiles, les MAST (Straménopiles marins) et les bacillariophycées ils étaient les plus représentés en termes de richesse surtout aux deux stations le plus proches de la côte (SOLEMIO & SOLA, **Fig.5, Fig.6, Annexe 1**). Les MAST sont largement distribués dans les océans du monde et ont été identifiés comme étant des flagellés hétérotrophes bactérivores (*Massana et al.*, 2006, *Jürgens et Massana*, 2008). Dans nos échantillons, d'autres protistes hétérotrophes ont été identifiés tels que les Telonemia (>20 OTUs) et Choanoflagellida (>30 OTUs). Pour les bacillariophycées, les données moléculaires ont révélé un bloom de *Pseudonitzschia sp.* en Juin-Juillet dans les stations SOLA et JULIO et de *Chaetoceros sp.* en Juin 2013 à SOLEMIO. Le bloom de *Pseudonitzschia sp.* a également été observé par microscopie pour la station JULIO (**Fig.10**). Les Chlorophytes et les groupes parasites Fungi et Cercozoa formaient les groupes les moins représentés, en termes de séquences et d'OTU, dans l'ensemble des stations avec une diversité moyennement plus élevée dans les stations le plus côtières (SOLA et SOLEMIO, **Fig.5 et 6**).

Les analyses de co-inertie appliquées sur les dix super-groupes taxonomiques ont montré un coulage significatif entre les conditions environnementales saisonnières et les abondances des super-groupes au sein de chaque station à l'exception de JULIO

(**Tableau 3A**). Elles ont révélé également un couplage significatif entre les stations SOLA et MOLA, par contre SOLEMIO et JULIO évoluaient indépendamment l'une de l'autre et également de celles de la partie Ouest du GoL. La structuration des groupes étaient différentes entre les stations et plus particulièrement entre SOLA & MOLA et SOLEMIO (**Annexe 2**). La structure hydrographique près du Rhône est hétérogène. Les résultats biologiques sont difficiles à expliquer sur la base des paramètres hydrographiques. Plusieurs facteurs interviennent dans cette structuration, principalement l'impact des intrusions du Rhône qui dépendait fortement du moment de l'intrusion et du type d'intrusion (petite, éphémère ou grande). SOLEMIO est sous influence directe des intrusions provenant du Rhône quant à SOLA est influencé par la dérive des panaches fluviaux et son volume (*Hu et al., 2009*). Par contre Julio est soumise aux apports transportés par le Courant du Nord. Il existe également d'autres facteurs pouvant jouer un rôle important sur la biologie et sa structuration comme l'upwelling, les vents et les courants secondaires.

Notre étude a accentué le point sur l'exploration de la diversité dans cet écosystème complexe et de déceler les différences et les similarités entre les stations étudiées. Pour une meilleure compréhension de la dynamique biologique, Il se peut qu'un suivi des intrusions et des autres paramètres (vent, courants) ainsi que les interactions biotiques permettra d'apporter plus d'éléments de réponse. Un échantillonnage plus au moins synchronisé entre les 4 stations sera également recommandé afin de pouvoir comparer minutieusement les quatre stations.

## REFERENCES BIBLIOGRAPHIQUES

- Altschul SF, Gish W, Miller W et al**, (1990). Basic local alignment search tool *J. Mol. Biol.*, 215
- Amann RI, Binder BJ, Olson RJ, Chisholm SW, Devereux R and Stahl DA**, (1990). Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations, *Appl. Environ. Microbiol.*, 56, 1919–192.
- Aminot A, Chaussepied M**, (1983). Manuel des analyses chimiques en milieu marin. CNEXO, Brest: 395 pp.
- Aminot A, Kerouel R**, (2004). Dissolved organic carbon, nitrogen, and phosphorus in the N-E Atlantic and the N-W Mediterranean with particular reference to non-refractory fractions and degradation *Deep-Sea Res, I* pp. 51–99
- Aminot A, K erouel R**, (2004). Hydrologie des  cosyst mes marins. Param tres et analyses. Ifremer: 336 pp.
- Aminot A, K erouel R**, (2007). Dosage automatique des nutriments dans les eaux marines. Ifremer: 188 pp.
- Batchelder HP, Powell TM**, (2002). Physical and biological conditions and processes in the northeast Pacific Ocean. *Progress in Oceanography*, 53, 105–
- Beaugrand G, Reid PC, Ibanez F, Lindley JA, Edwards M** (2002). Reorganization of North Atlantic Marine copepod biodiversity and climate. *Science*, 296, 1692– 1694.
- Behnke A, Engel M, Christen R et al**, (2011). Depicting more accurate pictures of protistan community complexity using pyrosequencing of hypervariable SSU rRNA gene regions *Environ. Microbiol.*, 13 : 340–349.
- Bethoux JP & Gentili B**, (1999). Functioning of the Mediterranean Sea: past and present changes related to freshwater input and climate changes. *Journal of Marine Systems*, 20: 33-47.
- Bethoux JP et Gentili B**, (1996). The Mediterranean Sea, coastal and deep-sea signatures of climatic and environmental changes. *Journ Mar. Syst.*, 7: 383-394.
- Bethoux JP, Gentili B, Raunet J et Tailliez D**, (1990). Warming trend in the western Mediterranean deep water. *Nature*, 347: 660-662.
- B thoux JP, Morin P, Ruiz-Pino DP**, (2002). Temporal trends in nutrient ratios: chemical evidence of Mediterranean ecosystem changes driven by human activity. *Deep-Sea Res II* 49:2007–2016.
- Blanfune A, Cohu S, Mangialajo L, Lem e R, Thibaut T**, (2012). Preliminary assessments of the impact of *Ostreopsis cf. ovata* (Dinophyceae) development on macroinvertebrates in the North Western Mediterranean Sea. *CRYPTOGAMIE ALGOL.* 33, 129-136.
- Bolch CJS, de Salas MF**, (2007). A review of the molecular evidence for ballast water introduction of the toxic dinoflagellates *Gymnodinium catenatum* and the *Alexandrium*. *Harmful Algae* 6 (4) 465–485.
- Box GEP, and Cox D**, (1964). R. An Analysis of Transformations *Journal of the Royal Statistical Society, Series B (Methodological)* Vol. 26, No. 2, pp. 211-252
- Br te J, Krabberot AK, Dolven JK, Ose RF, Kristensen T, Bjorklund K R et al** (2012). Radiolaria associated with large diversity of marine alveolates. *Protist* 163, 767–777. doi: 10.1016/j.protis.2012.04.004.
- Caron DA, Countway PD, Jones AC, Kim DY & Schnetzer A**, (2012). Marine protistan diversity. *Annual review of marine science*, 4, 467-493.
- Chiba S, Aita MN, Tadokoro K, Saino T, Sugisaki H, Nakata K**, (2008) . From climate regime shifts to lower- trophic level phenology: Synthesis of recent progress in retrospective studies of the western North Pacific. *Progress In Oceanography*, 77:112-126.
- Christaki U, Courties C, Massana R, Catala P, Lebaron P, Gasol P and Zubkov MV**, (2011). Optimized routine flow cytometric enumeration of heterotrophic flagellates using SYBR Green I. *Limnol. Oceanogr. Methods*, 9, 329–339.
- Christaki U, Genitsaris S, Monchy S, Li L, Rachik S, Breton E, Sime-Ngando T** (2017). Parasitic eukaryotes in a meso-eutrophic coastal system with marked phytoplankton blooms *Front. Mar. Sci.* 4:416. doi: 10.3389/fmars.2017.00416 .
- Christaki, U, Kormas KA, Genitsaris S, Georges C, Sime-Ngando T, Viscogliosi E et al**, (2014). Winter-summer succession of unicellular eukaryotes in a meso-eutrophic coastal system. *Microb. Ecol.* 67, 13–23. doi: 10.1007/s00248-013-0290-4
- Cohu S, Thibaut T, Mangialajo L, Labat JP, Passafiume A, Blanfun  A, Simon N, Cottalorda JM, Lem e R**, (2011). Occurrence of the toxic dinoflagellate *Ostreopsis cf. ovata* in relation with environmental factors in Monaco (NW Mediterranean). *Mar. Pol. Bull.* 62, 2681-2691.
- Crawford DW, Lindholm T**, (1997). Some observations on vertical distribution and migration of the phototrophic ciliate *Mesodinium rubrum* (= *Myrionecta rubra*) in a stratified brackish inlet. *Aquat Microb Ecol* 13:267–274.
- De Vargas C, Audic S, Henry N, Decelle J, Mah  F, Logares R et al** (2015). Eukaryotic plankton diversity in the sunlit ocean. *Science* 348: 1261605.
- Dol dec S, Chessel D** (1994). Co-inertia analysis: an alternative method for studying species-environment relationships. *Freshw Biol* 31: 277–294
- Dray S and Dufour AB**, (2007). The ade4 package: implementing the duality diagram for ecologists. *Journal of Statistical Software* 22(4): 1-20.
- Dray S, Chessel D, Thioulouse J**, (2003). Co-inertia analysis and the linking of ecological data tables. *Ecology* 84: 3078–3089
- Duarte CM, Agusti S, Cole JJ and del Giorgio PA**, (1999). Is the oligotrophic ocean heterotrophic?, *Science*, 284, 1735.

- Edgar RC**, (2010). Search and clustering orders of magnitude faster than BLAST Bioinformatics, 26 pp. 2460–2461
- Fraysse M, Pinazo C, Faure VM, Fuchs R, Lazzari P, et al** (2013) Development of a 3D Coupled Physical-Biogeochemical Model for the Marseille Coastal Area (NW Mediterranean Sea): What Complexity Is Required in the Coastal Zone? *PLoS ONE* 8(12): e80012. doi:10.1371/journal.pone.0080012.
- Genitsaris S, Monchy S, Breton E, Lecuyer E, Christaki U**, (2016). Small-scale variability of protistan planktonic communities relative to environmental pressures and biotic interactions at two adjacent coastal stations *Mar. Ecol. Prog. Ser.*, 548 pp. 61-75, doi: 10.3354/meps11647
- Genitsaris S, Monchy S, Viscogliosi E, Sime- Ngando T, Ferreira S, Christaki U** (2015). Seasonal variations of marine protist community structure based on taxon-specific traits using the eastern English Channel as a model coastal system *FEMS Microbiol. Ecol.*, 91
- Georges C, Monchy S, Genitsaris S and Christaki U**, (2014). Protist community composition during early phytoplankton blooms in the naturally iron-fertilized Kerguelen area (Southern Ocean). *Biogeosciences* 11, 5847–5863. doi: 10.5194/bg-11-5847
- Goffart A, Hecq JH, Legendre L**, (2002). Changes in the development of the winter-spring phytoplankton bloom in the Bay of Calvi (NW Mediterranean) over the last two decades: A response to changing climate? *Mar Ecol Prog Ser* 236:45–60
- Gómez F, Gorsky G**, (2003). Annual microplankton cycles in Villefranche Bay, Ligurian Sea, NW Mediterranean, *Journal of Plankton Research*, 25:323-339.
- Guillou L, Bachar D, Audic S et al**, (2013). The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy *Nucleic Acid Res.* 41, pp. 597–604
- Guillou L, Viprey M, Chambouvet A, Welsh RM, Kirkham AR, Massana R et al**, (2008). Widespread occurrence and genetic diversity of marine parasitoids belonging to Syndiniales (Alveolata). *Environ. Microbiol.* 10, 3349–3365. doi: 10.1111/j.1462-2920.2008.01731.x.
- Harada A, Ohtsuka S and Horiguchi T, (2007). Species of the parasitic genus *Duboscquella* are members of the enigmatic marine alveolate group, I. *Protist* 158, 337–347. doi: 10.1016/j.protis.2007.03.005.
- Hooff RC and Peterson WT**, (2006). Copepod biodiversity as an indicator of changes in ocean and climate conditions of the northern California current ecosystem. *Limnol. Oceanogr.* 51: 2607–2620. doi:10.4319/lo.2006.51.6.2607
- Hu ZY, Doglioli AM, Petrenko AA, Marsaleix P et Dekeyser I** (2009). Dekeyser. Numerical simulations of eddies in the Gulf of Lion. *Ocean Modelling*, 28(4): 203–208.
- Johnson RG** (1997). Climate control requires a dam at the strait of Gibraltar. *EOS Trans Am Geophys Union* 78:277–281.
- Jürgens K and Massana R**, (2008). Protistan grazing on marine bacterioplankton, in: *Microbial Ecology of the Oceans*, edited by: Kirchman, D. L., 2nd Edn., 383–441.
- Karl DM, Bidigare RR, Letelier RM**, (2001). Long-term changes in plankton community structure and productivity in the North Pacific Subtropical Gyre: The domain shift hypothesis. *Deep Sea Research Part II: Topical Studies in Oceanography*, 48:1449-1470.
- Kenkel NC**, (2006). On selecting an appropriate multivariate analysis. *Can J Plant Sci* 86: 663–676
- Kuhn M, Contributions from Wing J, Weston S, Williams A, Keefer C, Engelhardt A, Cooper T, Mayer Z, Kenkel B, the R Core Team, Benesty M., Lescarbeau R, Ziem A, Scrucca I, Tang Y and Candan C**, (2016). Caret: Classification and Regression Training R package version 6.0-68.
- Kunin V, Engelbrektson A, Ochman H et al**, (2010). Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates *Environ. Microbiol.*, 12
- Lefevre D, Minas HJ, Minas M, Robinson C, Williams PJ** (1997). Le B. Woodward E.M.S.. Review of gross community production, primary production, net community production and dark respiration in the gulf of Lions. *Deep-Sea Res.* 44 801-832.
- Legendre P, Legendre L**, (1998). Numerical ecology, 2nd edn. Elsevier, Amsterdam
- López-García P, Philippe H, Gail F and Moreira D**, (2003). Autochthonous eukaryotic diversity in hydrothermal sediment and experimental microcolonizers at the Mid-Atlantic Ridge, P. *Natl.Acad. Sci. USA*, 100, 697–702.
- Lopez-Garcia P, Rodriguez-Valera F, Pedros-Alio C, Moreira D** (2001). Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. *Nature* 409:603–7.
- Mackas L, Thomson RE, Galbraith M**, (2001). Hanges in the zooplankton community of the British Columbia continental margin, 1985-1999, and their covariation with oceanographic conditions. *Canadian Journal of Fisheries and Aquatic Sciences*, 58:685-702.
- Marbà N, Duarte CM**, (1997). Interannual changes in seagrass (*Posidonia oceanica*) growth and environmental change in the Spanish Mediterranean littoral. *Limnol Oceanogr* 42: 800–810.
- Marty JC, Chiavérini J, Pizay MD, Avril B**, (2002). Seasonal and interannual dynamics of nutrients and phytoplankton pigments in the western Mediterranean Sea at the DYFAMED time-series station (1991–1999). *Deep-Sea Res II* 49:1965–1985
- Massana R, Gobet A, Audic S, Bass D, Bittner L, Boute C et al**, (2015). Marine protist diversity in European coastal water and sediments as revealed by high-throughput sequencing. *Environ. Microbiol.* 10, 4035–4049. doi: 10.1111/1462-2920.12955.
- Massana R and Pedrós-Alió C**, (2008). Unveiling new microbial eukaryotes in the surface ocean, *Curr. Opin. Microbiol.*, 11, 213–218.
- Massana R, Terrado R, Forn I, Lovejoy C and Pedrós-Alió C**, (2006). Distribution and abundance



- of uncultured heterotrophic flagellates in the world oceans, *Environ. Microbiol.*, 8, 1515–1522.
- McManus GB, Fuhrman JA**, (1986) Photosynthetic pigments in the ciliate *Laboea strobila* from Long Island Sound, USA. *J Plankton Res* 8:317–327.
- Medinger R, Nolte V, Pandey RV, Jost S, Ottenwalder B, Schlotterer C and Boenigk J**, (2010). Diversity in a hidden world: potential and limitation of next-generation sequencing for surveys of molecular diversity of eukaryotic microorganisms, *Mol. Ecol.*, 19, 32–40.
- Millot C**, (1982). Analysis of upwelling in the Gulf of Lions. In: Nihoul J.C.J. (Ed.), *Hydrodynamics of Semi-Enclosed Seas: Proceedings of the 13th International Liège Colloquium on Ocean Hydrodynamics*. Elsevier Oceanographical Series, vol. 34, Elsevier Scientific Publications, Amsterdam, The Netherlands, pp. 143–153.
- Modigh M**, (2001). Seasonal variations of photosynthetic ciliates at a Mediterranean coastal site. *Aquat Microb Ecol* 23:163–175.
- Monchy S, Grattepanche J-D, Breton E, Meloni D, Sanciu G, Chabé M, Delhaes L, Viscogliosi E, Sime-Ngando T, Christaki U**, (2012). Microplanktonic community structure in a coastal system relative to a *Phaeocystis* bloom inferred from morphological and tag pyrosequencing methods. *PLoS One* 7:e39924.
- Moutin T, Raimbault P**, (2002). Primary production, carbon export and nutrients availability in western and eastern Mediterranean Sea in Early summer 1996 (MINOS cruise). *Journal of marine systems*. 33-34, 273-288.
- Penna A, Casabianca S, Guerra AF, Vernesi C, Scardi M**, (2017). Analysis of phytoplankton assemblage structure in the Mediterranean Sea based on high-throughput sequencing of partial 18S rRNA sequences. *Marine Genomics* 36 49–55.
- Pereira Quiroga AI**, (2005). Evolution des populations phytoplanctoniques de la Baie de Banyuls-sur-Mer, avec une attention particulière pour les genres possédant des espèces potentiellement toxiques. *Thèse de doctorat en Océanologie biologique*, Université Pierre et Marie Curie, Paris, 186 pp.
- Rohling EG, Bryden HL**, (1992). Man-induced salinity and temperature increases in western Mediterranean deep water. *J Geophys Res* 97:11191–11198.
- Rohling EJ, Bryden H**, (1992). Man-induced salinity and temperature increases in Mediterranean deep water. *J. Geophys. Res.*, 97 (C7), pp. 11191–11981
- Sanders RW**, (1995). Seasonal distributions of the photosynthetic ciliates *Laboea strobila* and *Myrionecta rubra* (= *Mesodinium rubrum*) in an estuary of the Gulf of Maine. *Aquat Microb Ecol* 9:237–242.
- Schloss PD, Gevers D, Westcott SL**, (2011). Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies *PLoS One*, 6 (e27310)
- Schloss PD, Westcott SL, Ryabin T et al**, (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities *Appl. Environ. Microb.*, 75 (7537–41)
- Skovgaard A, Massana R, Balague V and Saiz E**, (2005). Phylogenetic position of the copepod-infesting parasite *Syndinium turbo* (Dinoflagellata, Syndinea). *Protist* 156, 413–423. doi: 10.1016/j.protis.2005.08.002.
- Stickney HL, Hood RR, Stoecker DK**, (2000). The impact of mixotrophy on planktonic marine ecosystems. *Ecol Model* 125:203–230.
- Stoecker DK**, (1987). Photosynthesis found in some single-cell marine animals. *Oceanus* 30, 49–53.
- Stoecker DK**, (1999). Mixotrophy among dinoflagellates. *J. Eukaryot. Microbiol.* 46:397–401
- Tunin-Ley A, Ibañez F, Labat JP, Zingone A, Lemée R**, (2009). Phytoplankton biodiversity and NW Mediterranean Sea warming: Changes in the dinoflagellate genus *Ceratium* in the 20th century. *Mar. Ecol. Prog. Ser.* 375:85–99. doi: 10.3354/meps07730.
- Tunin-Ley A, Labat JP, Gasparini S, Mousseau L, Lemée R**, (2007). Annual cycle and diversity of species and infraspecific taxa of *Ceratium* (Dinophyceae) in the Ligurian Sea, northwest Mediterranean. *J. Phycol.* 43:1149–1163. doi: 10.1111/j.1529-8817.00417.x.
- Van Wambeke F, Lefevre D, Prieur L, Sempéré R, Bianchi M, Oubelkheir K and Bruyant F**, (2004). Distribution of microbial biomass, production, respiration dissolved organic carbon and factors controlling bacterial production across a geostrophic front (Almeria-Oran, SW Mediterranean Sea), *Mar. Ecol. Prog. Ser.*, 269, 1–15.
- Videau C, Sournia A, Prieur L, Fiala M**, (1994). Phytoplankton and primary production characteristics at selected sites in the geostrophic Almeria–Oran front system (SW Mediterranean Sea). *J. Mar. Sys.* 5, 235–250.
- Winkler LW**, (1888). Die Bestimmung des in Wasser gelösten Sauerstoffes. *Berichte der Deutschen Chemischen Gesellschaft*, 21: 2843–2855.

**ANNEXE 1 :** Nombre d'OTUs identifiées dans les différentes stations : SOLA, MOLA, SOLEMIO et JULIO (SOLA : 15 dates, MOLA : 13 dates, SOLEMIO : 18 dates et JULIO : 13 dates).

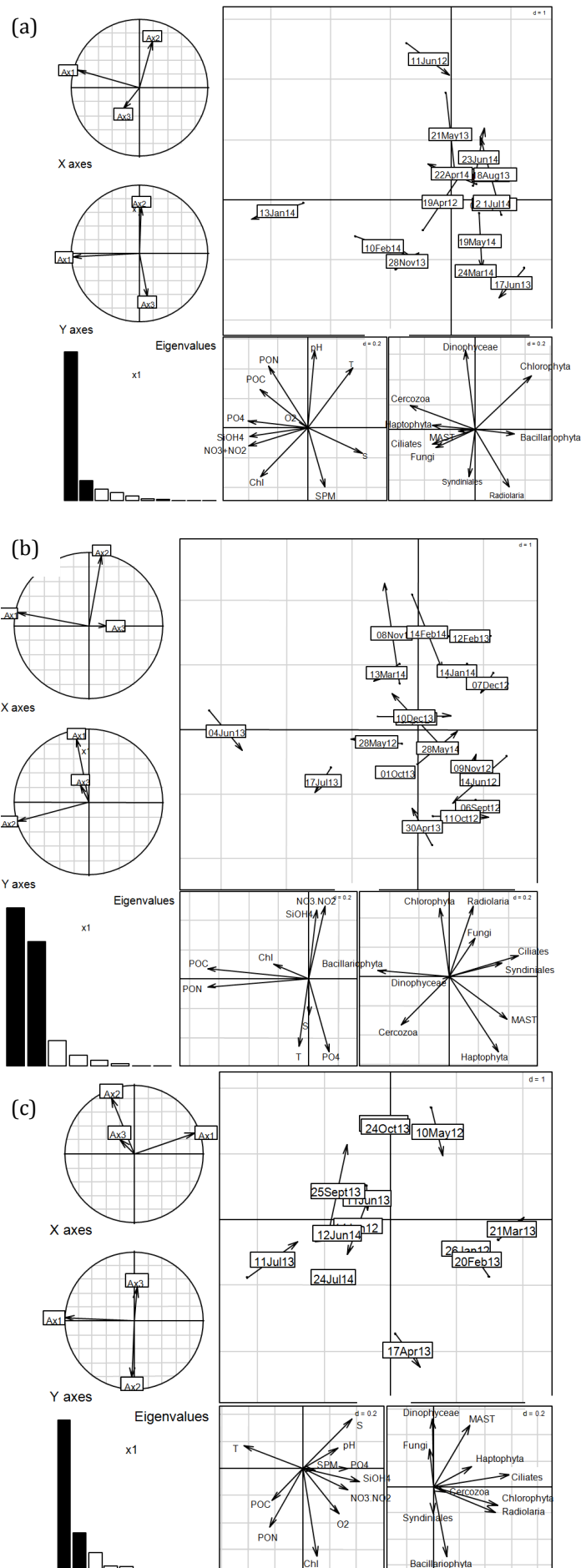
Super groups	groups	OTUs Number			
		MOLA	SOLA	SOLEMIO	JULIO
	Apicomplexa	2	6	3	1
	Perkinsea	0	2	1	0
<b>Alveolata</b>	Colpodea	4	5	7	7
	Ciliates Heterotrichea	0	0	1	1
	Listomatea	9	15	15	7
	Oligohymenophore	6	10	12	7
	Phyllopharyngea	5	5	3	3
	Prostomatea	1	3	1	1
	Spirotrichea	65	102	95	73
	Other Ciliates	2	2	0	2
Dinophyta	Dinophyceae	216	359	246	252
	Syndiniales	488	450	735	594
	Others	3	3	4	2
<b>Apusozoa</b>	Apusomonadidae	1	0	0	0
<b>Archaeplastida</b>	Chlorophyta	66	84	95	66
	Rhodophyta	0	4	1	0
	Streptophyta	1	7	5	2
<b>Hacrobia</b>	Centrohelioczoa	4	4	6	3
	Cryptophyta	2	1	3	2
	Haptophyta	45	44	62	45
	Katablepharidophyta	2	3	0	2
	Picobiliphita	11	9	13	10
	Telonemia	34	30	53	43
	Choanoflagellida	23	31	30	31

	Fungi	36	67	59	33
<b>Opisthokonte</b>	Mesomycetozoa	12	14	14	8
	Others	0	1	0	0
	Cercozoa	24	53	49	37
<b>Rhizaria</b>	Radiolaria	15	4	19	17
	Bacillariophyta	33	88	76	77
<b>Stramenopiles</b>	Bicoecia	1	3	5	0
	Bolidophyceae and Relatives	11	10	9	12
	Chrysomerophyceae	0	1	0	0
	Chrysophyceae-Synurophyceae	21	23	25	21
	Dictyochophyceae	20	20	23	16
	Eustigmatophyceae	1	2	1	0
	Hyphochytriomyceta	0	1	1	0
	Labyrinthulea	19	38	15	15
	Labyrinthulomycetes	0	2	1	0
	MAST	94	97	126	102
	MOCH	9	11	12	8
	Ochrophyta	0	1	0	0
	Oomycota	1	13	12	6
	Opalinata	0	0	0	0
	Pelagophyceae	4	5	12	8
	Phaeophyceae	0	2	0	0
	Pinguiphyceae	0	0	2	2
	Pirsonia	0	0	1	2
	Raphidophyceae	0	1	0	0
	Others	13	9	22	17
Unidentified	17	4	20	19	

**ANNEXE 2 :** Analyses de Co-inertie (PCA-PCA COIA) entre les variables environnementales et l'abondance des OTU triés en groupes taxonomiques. Les axes X montrent les projections des trois premiers composants de la PCA à partir des variables explicatives (variables environnementales). Les axes Y montrent les projections des trois premiers axes PCA à partir des variables de réponse (groupes taxonomiques). Les cercles représentent une vue de la rotation nécessaire pour associer les deux ensembles de données. Les valeurs de  $P$  ont été calculées en utilisant le test de permutation de Monte-Carlo (1000 permutations). Le diagramme de dispersion de l'échantillon montre à quelle distance les échantillons étaient relativement à leurs variables explicatives et de réponses. Le début de la flèche montre la position de l'échantillon décrite par les variables explicatives et la fin par les variables de réponse. Les variables ont été transformées en  $\log(x + 1)$  et les variables environnementales ont été normalisées. RV: coefficient de corrélation entre les 2 tables ('R' pour la corrélation et 'V' pour le vectériel).

(a) SOLA, (b) MOLA, (c) SOLEMIO

Pour les valeurs de  $P_{value}$ , RV et les pourcentages des axes, voir **Tableau 6A**



**ANNEXE 3 :** Liste des différents morphotypes, classes de taille, biomasse, biomasse moyenne basée sur 100-300 mesures cellulaires par morphotype (facteur de conversion du biovolume en biomasse Menden pour les dinoflagellés et Stoeker pour les ciliés) et le pourcentage d'abondance calculé par rapport à chaque classe de taille.

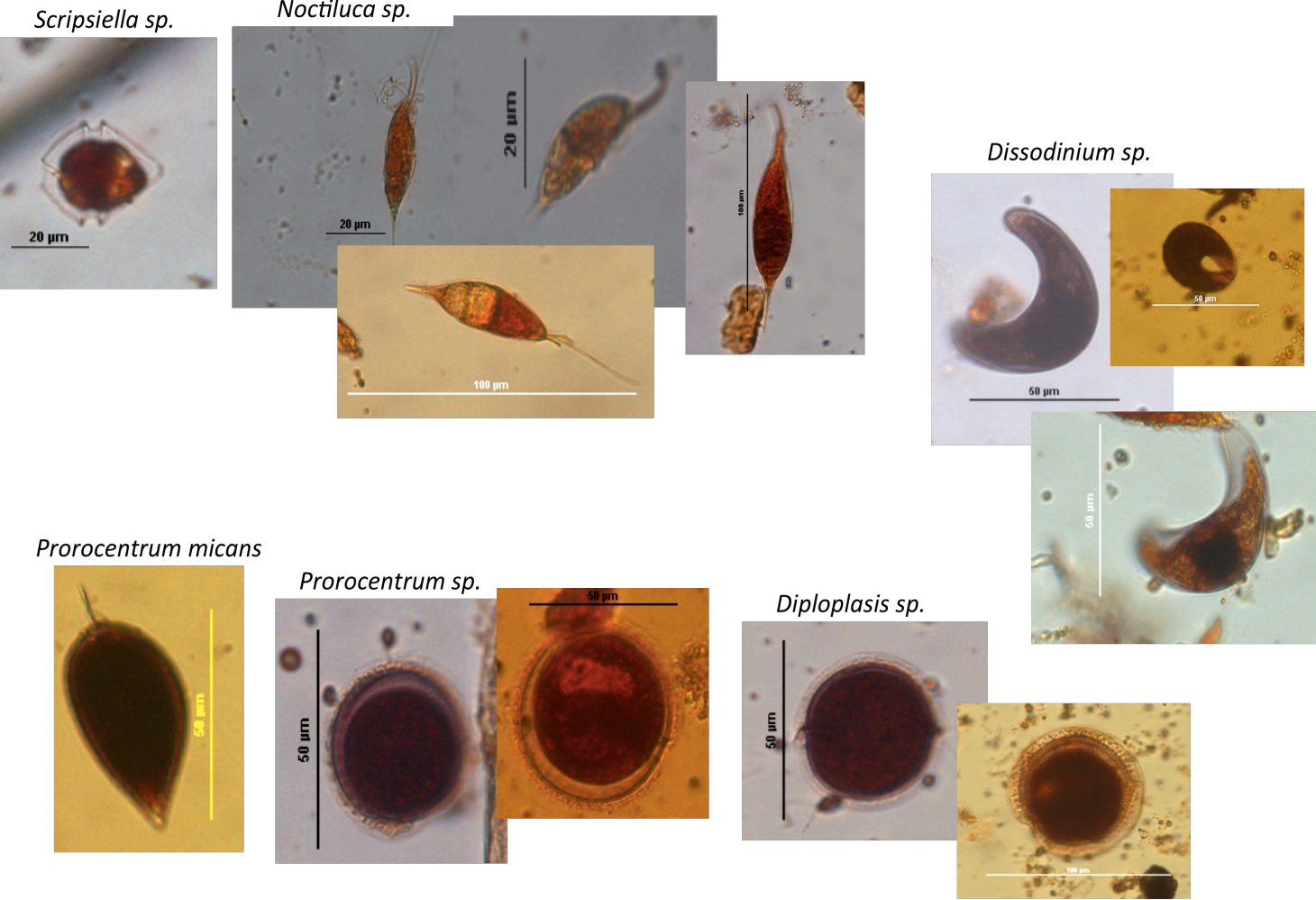
<b>Morphotype</b>	<b>Size (<math>\mu\text{m}</math>)</b>	<b>ESD (<math>\mu\text{m}</math>) <math>\pm</math> SD</b> <b>min-max</b>	<b>Mean biomass</b> <b>ng C Cell<sup>-1</sup></b>
<b>Dinoflagellates</b>			
<b>Gyrodinium</b>	<20	<b>13.94<math>\pm</math>3.2</b> 6.6-19.4	0.28 $\pm$ 0.2
	20-30	<b>25.38<math>\pm</math>3.2</b> 20.5-29.9	1.09 $\pm$ 0.5
	30-50	<b>42.77<math>\pm</math>5.49</b> 34.5-47.6	4.57 $\pm$ 2
	50-100	<b>52.44<math>\pm</math>0.1</b> 52.3-52.5	8.22 $\pm$ 0
		mean biomass (ng C Cell <sup>-1</sup> ) $\pm$ SD	<b>1.23<math>\pm</math>1.9</b>
<b>Gymnodinium</b>	<20	<b>12.87<math>\pm</math>3</b> 7.8-19.9	
	mean biomass (ng C Cell <sup>-1</sup> ) $\pm$ SD	<b>0.18<math>\pm</math>0.1</b>	
<b>Diploplasis</b>	<20	<b>18.15<math>\pm</math>1.3</b> 16-18.6	0.41 $\pm$ 0
	20-30	<b>22.62<math>\pm</math>1.3</b> 21.1-25.2	0.77 $\pm$ 0.1
	30-50	<b>38.21<math>\pm</math>5.6</b> 31.4-47	3.53 $\pm$ 1.4
	mean biomass (ng C Cell <sup>-1</sup> ) $\pm$ SD	<b>2.88<math>\pm</math>4.8</b>	
	<b>Prorocentrum</b>	<20	<b>15.3<math>\pm</math>2.3</b> 11.9-17
20-30		<b>24.16<math>\pm</math>3.2</b> 20-29.7	0.96 $\pm$ 0.3
30-50		<b>35.31<math>\pm</math>5.4</b> 30.2-48.1	2.86 $\pm$ 1.4
50-100		<b>74.87<math>\pm</math>32.6</b> 51.7-97.9	27.87 $\pm$ 28.2
mean biomass (ng C Cell <sup>-1</sup> ) $\pm$ SD		<b>2.4<math>\pm</math>6.3</b>	
<b>Protopteridinium</b>	<20	<b>19.27<math>\pm</math>0.02</b> 19.26-19.27	0.49 $\pm$ 0
	20-30	<b>25.37<math>\pm</math>3</b> 20.8-29.3	1.1 $\pm$ 0.3
	30-50	<b>37.52<math>\pm</math>5.6</b> 30.6-48.5	3.37 $\pm$ 1.4
	50-100	<b>58.01<math>\pm</math>7.6</b> 51.7-75.2	11.37 $\pm$ 4.7
	mean biomass (ng C Cell <sup>-1</sup> ) $\pm$ SD	<b>4.06<math>\pm</math>4</b>	
<b>Ceratium</b>	<20		

	20-30		
	30-50	<b>46.25±2.6</b>	5.81±0.9
	50-100	<b>56.7±4.3</b>	
		41.6-49.6	
		50.4-66.2	10.39±2.3
mean biomass (ng C Cell <sup>-1</sup> )±SD		<b>8.83±2.9</b>	
<b>Dinophysis</b>	<20	<b>16.44±2.2</b>	0.32±0.1
		13.7-19	
	20-30	<b>27.07±2.6</b>	1.3±0.3
		23.2-29.8	
	30-50	<b>38.53±5.4</b>	3.6±1.5
		31.4-49.6	
	50-100	<b>61.79±10.7</b>	13.86±6.7
		51.1-76.6	
mean biomass (ng C Cell <sup>-1</sup> )±SD		<b>4.14±5</b>	
<b>Dino. Naked</b>	<20	<b>11.91±3.2</b>	0.15±0.1
		6.3-19.9	
	20-30	<b>25.62±2.6</b>	1.11±0.3
		22-28.7	
	30-50	<b>38.23±3.4</b>	3.42±0.8
		33.4-40.9	
mean biomass (ng C Cell <sup>-1</sup> )±SD		<b>0.49±2.1</b>	
<b>Dino. Armored</b>	<20	<b>14.36±2.7</b>	0.23±0.1
		8.1-19.2	
	20-30	<b>24.84±2.7</b>	1.03±0.3
		20.2-29.7	
	30-50	<b>38.68±6.9</b>	3.76±1.8
		30.3-48.4	
	50-100	<b>63.49±10.8</b>	15.04±7.3
		51.5-82.1	
mean biomass (ng C Cell <sup>-1</sup> )±SD		<b>2.48±7.4</b>	
<b>Ciliates</b>			
<b>Myrionecta</b>	<20	<b>14.28±3.1</b>	0.33±0.2
		9.3-19.9	
	20-30	<b>21.4±1.5</b>	0.98±0.2
		20-23	
	30-50	<b>36.39±5.3</b>	5.02±2.1
		30-42.7	
mean biomass (ng C Cell <sup>-1</sup> )±SD		<b>0.97±1.7</b>	
<b>Laboea strobila</b>	<20	<b>16.356±2.2</b>	0.44±0.1
		13.1-19	
	20-30	<b>22.35±3.1</b>	0.95±0.3
		20.1-27.8	
	30-50	<b>40.28±5.7</b>	4.06±1.3
		31-48.4	
	50-100	<b>62.22±8.5</b>	11.79±4
		52.2-76.9	
mean biomass (ng C Cell <sup>-1</sup> )±SD		<b>5.36±5</b>	

<b>Tontonia</b>	<20	<b>13.58±4.3</b> 7-19.4	0.31±0.2
	20-30	<b>25±2.6</b> 20.1-29.9	1.24±0.3
	30-50	<b>32.81±2.4</b> 30.2-40.1	2.39±0.4
	mean biomass (ng C Cell <sup>-1</sup> )±SD	<b>1.3±0.6</b>	
<b>Tintinnids</b>	<20	<b>18.95±1.6</b> 17-19.9	0.68±0.1
	20-30	<b>24.69±2</b> 20.1-27	1.52±0.3
	30-50	<b>41.77±3.9</b> 33.1-49.2	7.43±1.9
	50-100	<b>55.67±2.4</b> 51.7-58.6	17.25±2.2
	mean biomass (ng C Cell <sup>-1</sup> )±SD	<b>6.65±5.3</b>	
<b>Ciliates</b>	<20	<b>15.17±02.9</b> 6.5-19.8	0.38±0.1
	20-30	<b>23.93±2.8</b> 20-29.5	1.42±0.5
	30-50	<b>37.29±5.2</b> 31.3-47.5	5.45±2.4
	50-100	<b>68.15±19.7</b> 54.2-82	35.4±27.6
	mean biomass (ng C Cell <sup>-1</sup> )±SD	<b>1.51±4.7</b>	

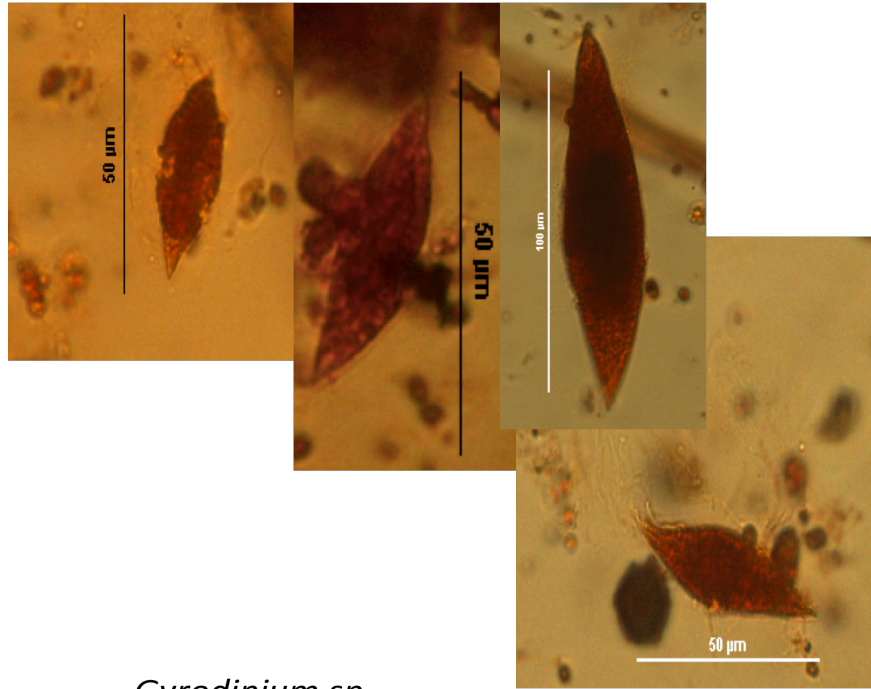
ANNEXE 4 : Trombinoscope des taxons observés au microscope inversé en Méditerranée.

**DINOFLAGELLES**

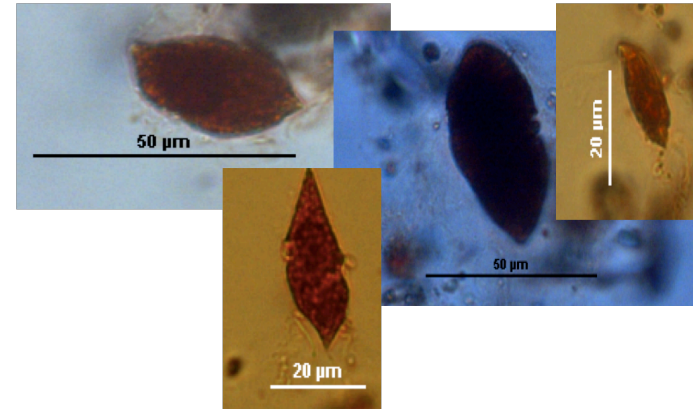




*Gyrodinium fusiforme*



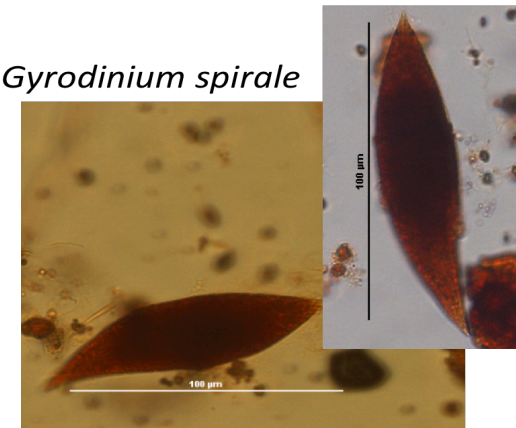
*Gyrodinium rhadomante*



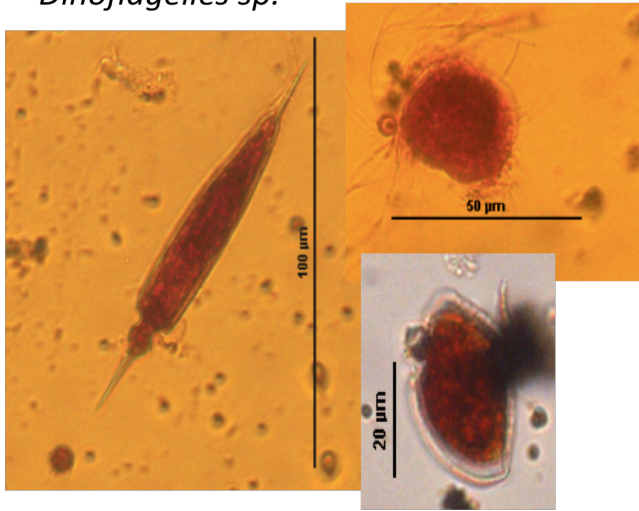
*Gyrodinium sp.*



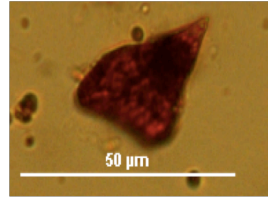
*Gyrodinium spirale*



*Dinoflagellés sp.*



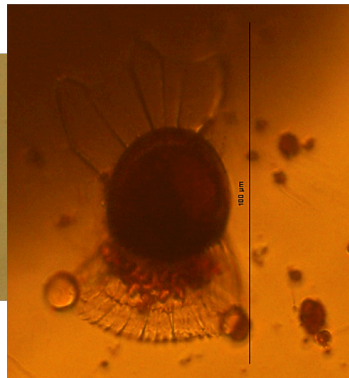
*Dinoflagellés sp.*



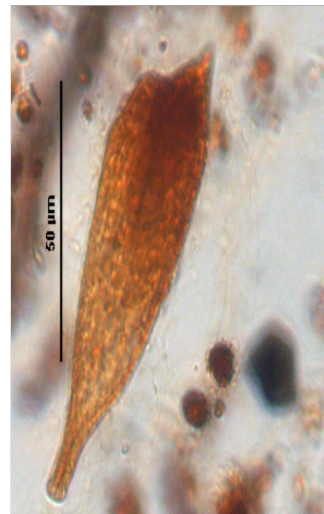
*Dinoflagellés sp.*



*Dinoflagellés sp.*



*Katodinium sp.*



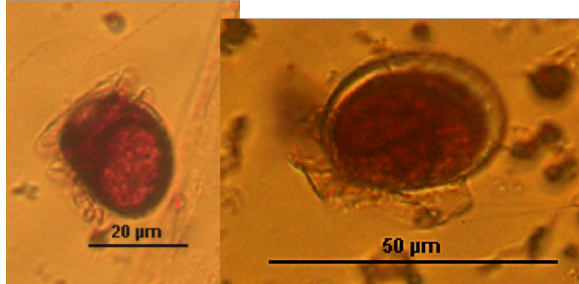
*Katodinium glaucum*



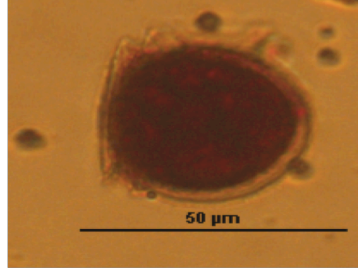
*Amphidinium sp.*



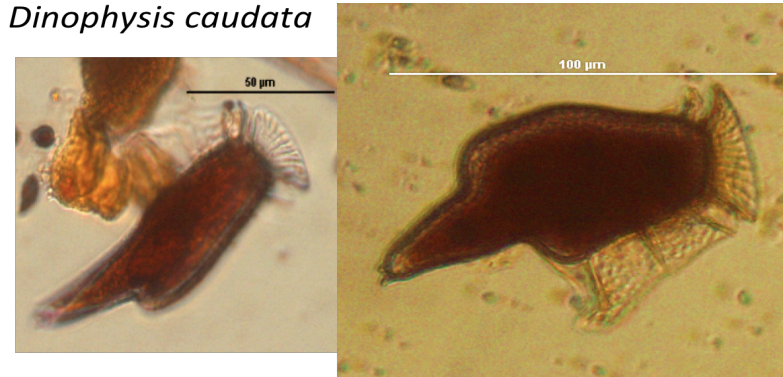
*Dinophysis ovum*



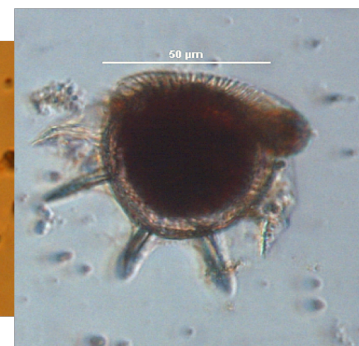
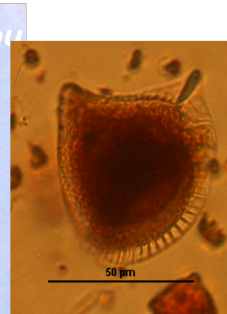
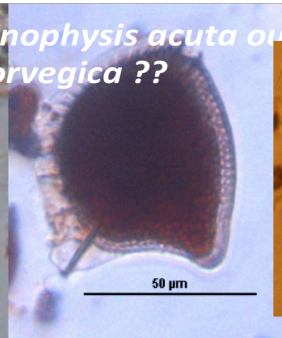
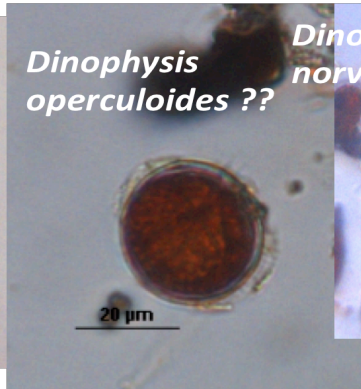
*Dinophysis tenuivelata*



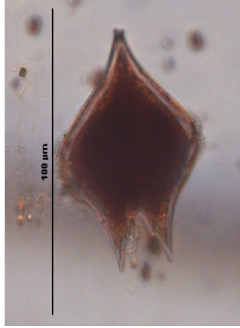
*Dinophysis caudata*



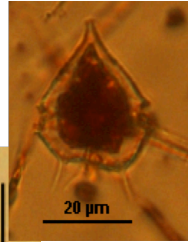
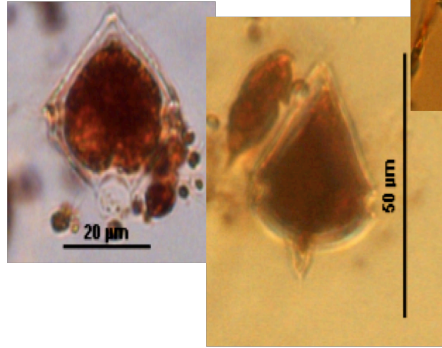
*Dinophysis sp.*



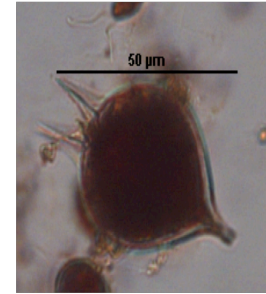
*Protoperidinium pallidum*



*Protoperidinium sp.*



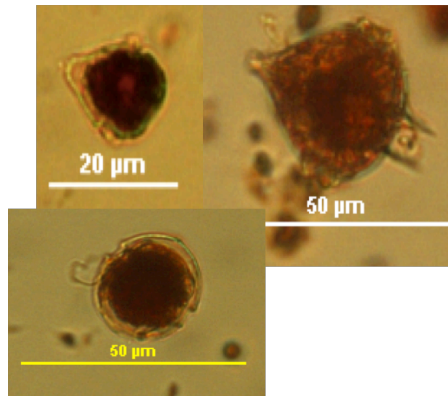
*Protoperidinium bellulum* ??



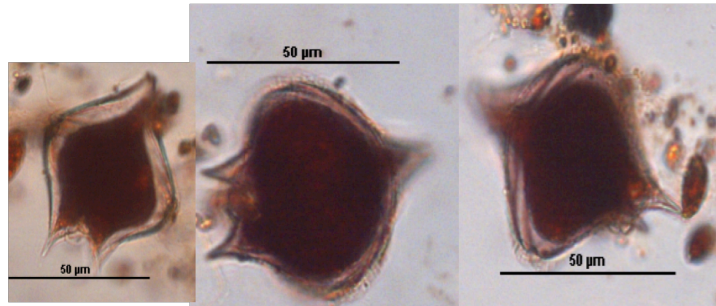
*Protoperidinium sp.*



*Protoperidinium sp.*



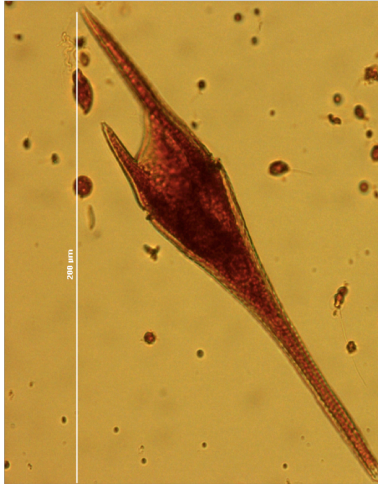
*Protoperidinium depressum*



*Protoperidinium pyriforme*



*Ceratium Lineatum*



*Ceratium sp.*



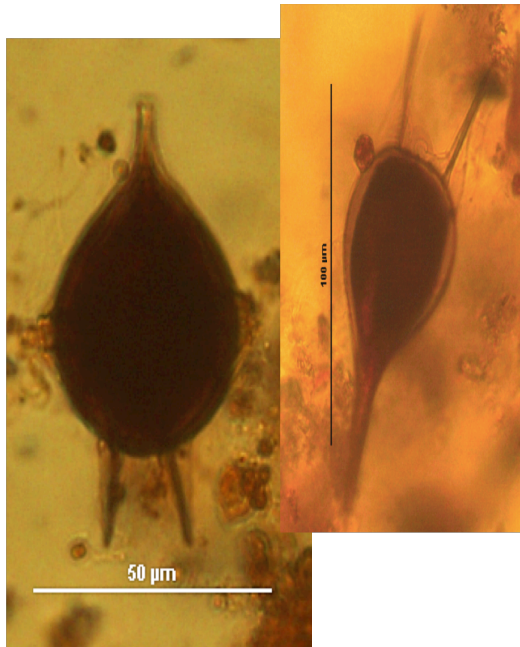
*Ceratium fusus*



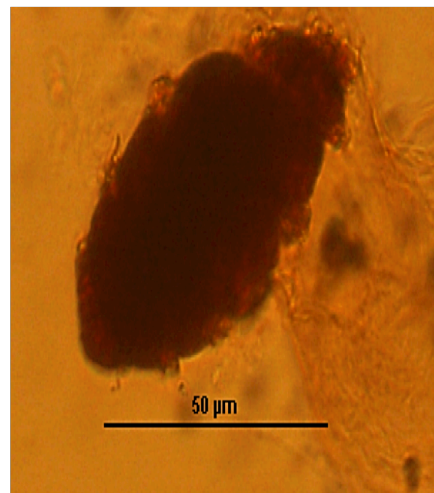
*Ceratium sp.*



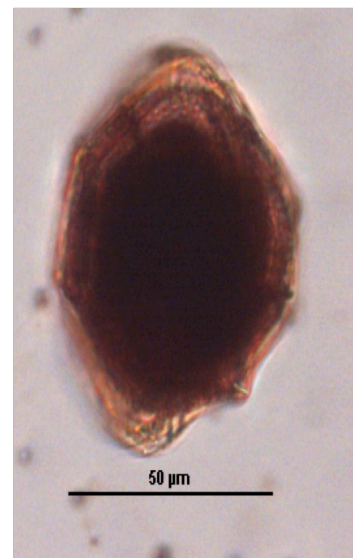
*Podolampas sp.*



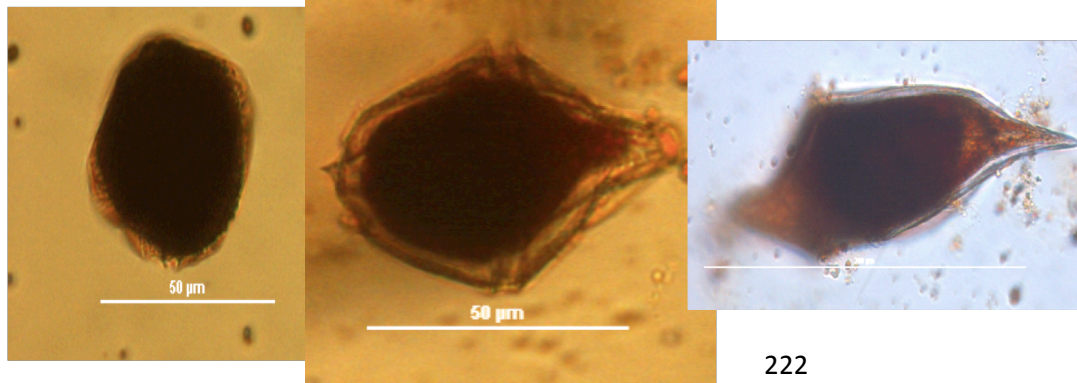
*Warnovia sp.*



*Gonyaulax digitale*

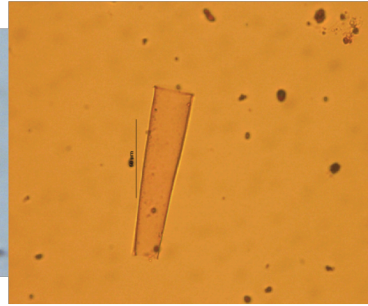
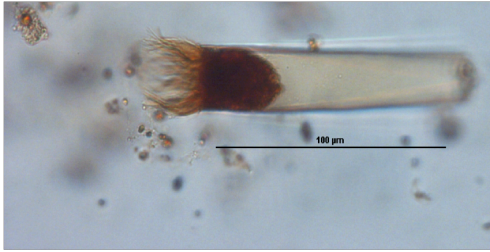


*Dinoflagellés sp.*

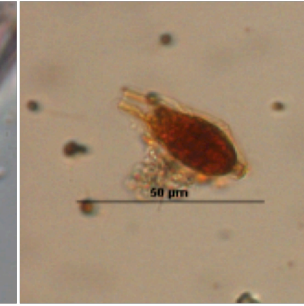
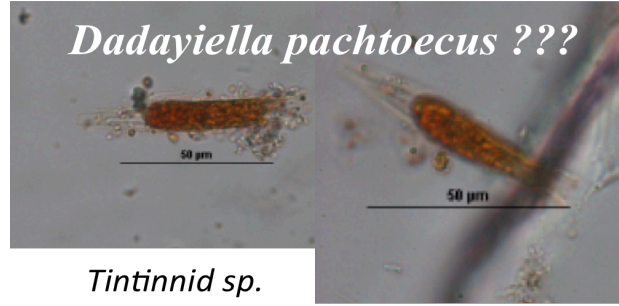


**CILIES**

*Eutintinnus*

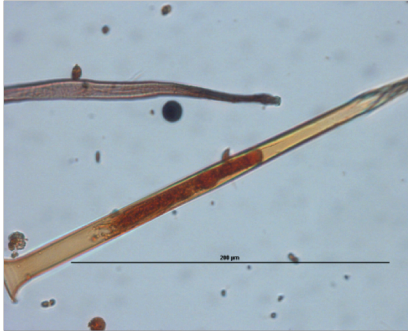


*Dadayiella pachtoecus* ???

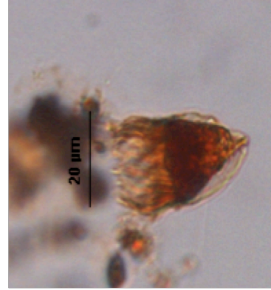


*Tintinnid* sp.

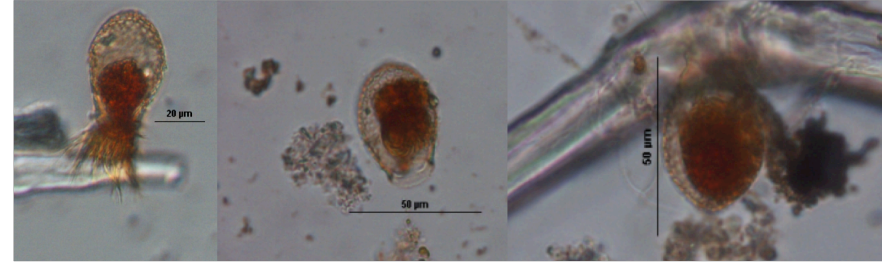
*Salpingella acuminatoides*



*Cymatocylis* Cf. *affinis*

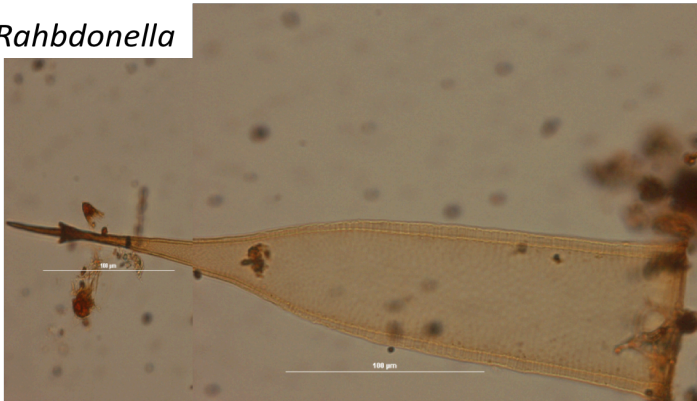


*Cadonellopsis soyaii*

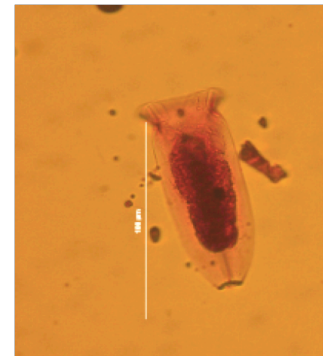
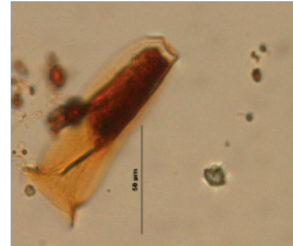


*Amphorides quadrilineata*

*Rahbdonella*



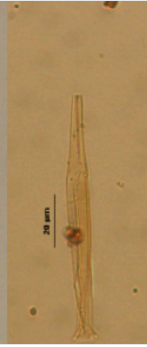
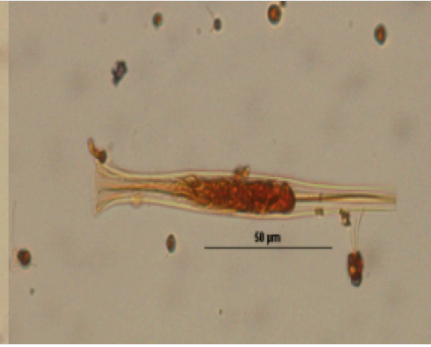
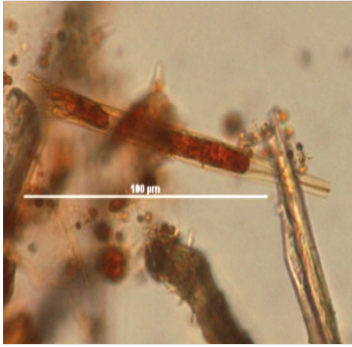
*Amphorella* sp.



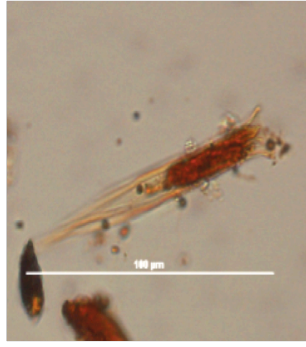
*Dictyocysta*



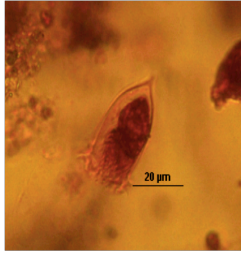
*Salpingella* sp.



*Salpingella decurtata*



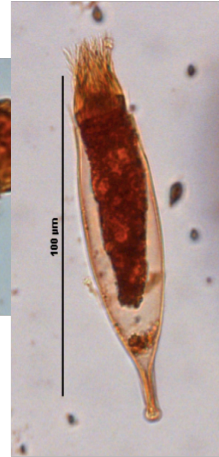
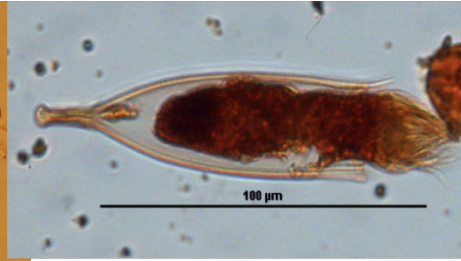
*Proplectella parva* ??



*Proplectella*

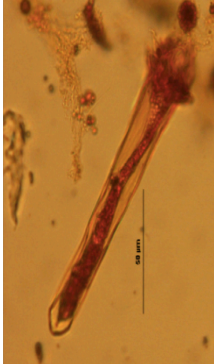


*Favella*

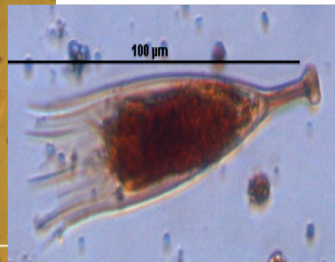
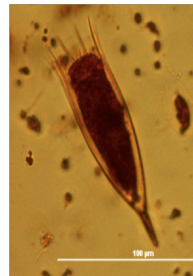


*Tintinids* sp.

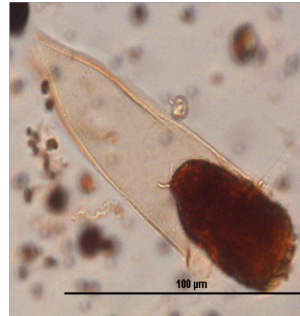
*Steenstrupiella* ??



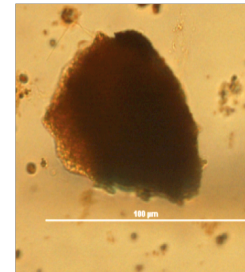
*Dadayella*



*Rahbdonella* ??

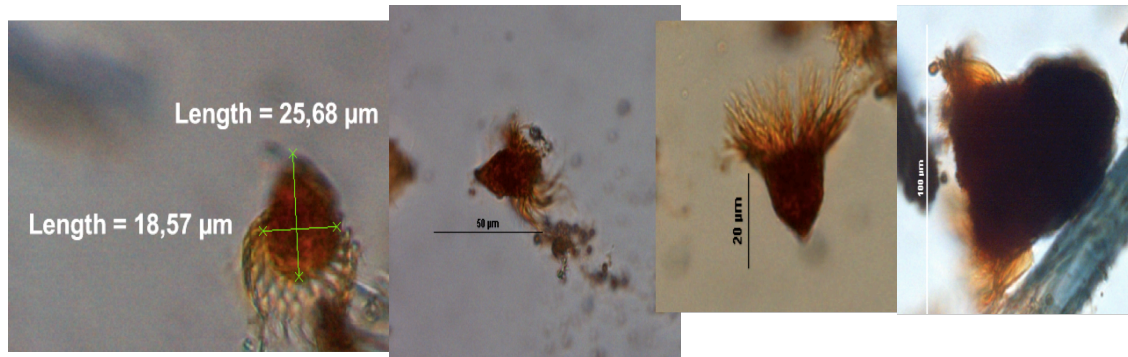


*Tintinnopsis mayeri* ??

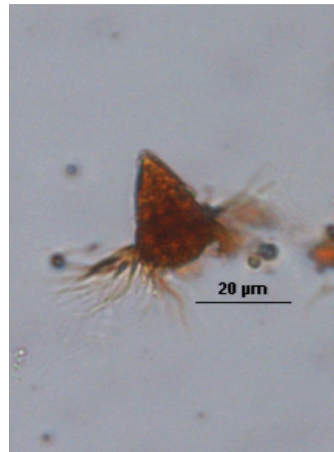




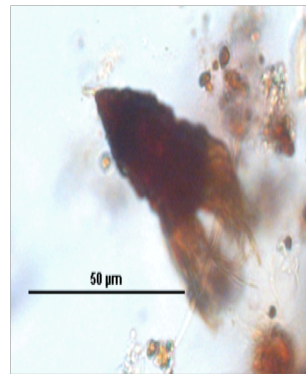
*Strombidium* sp.



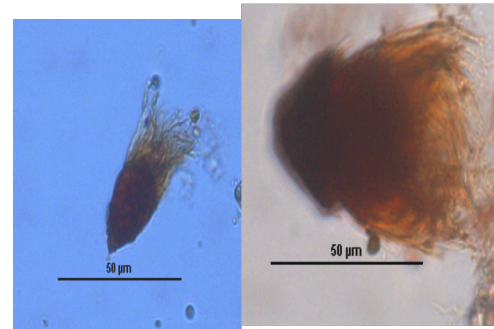
*Strombidium conicum*



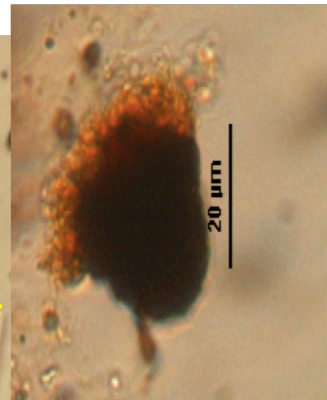
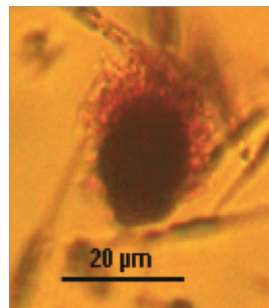
*Strombidium emergens*



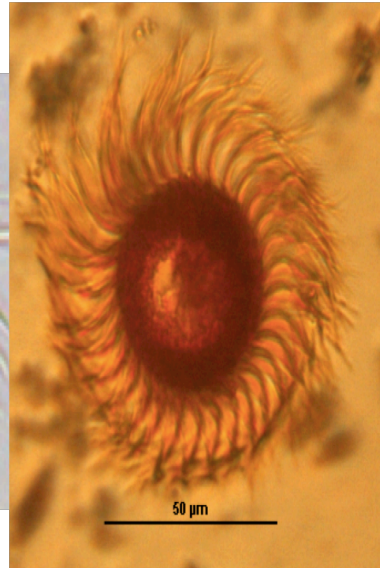
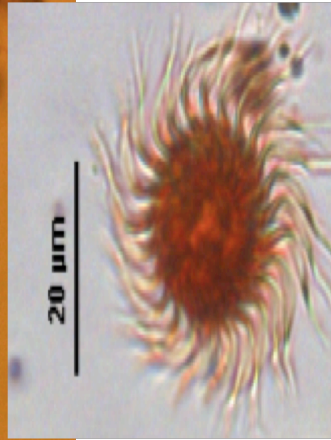
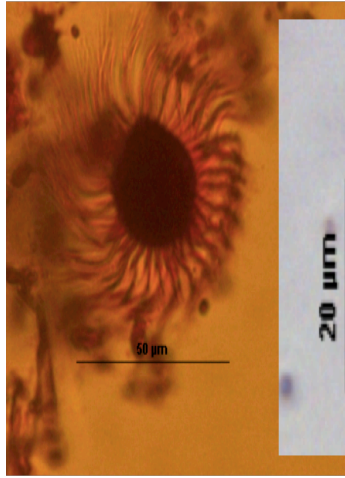
*Strombidium epidemum*



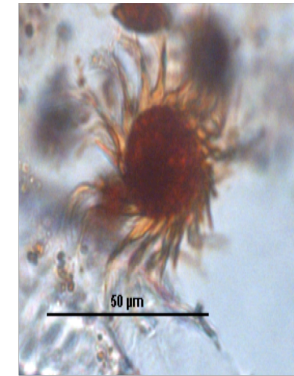
*Strombidium lynni*



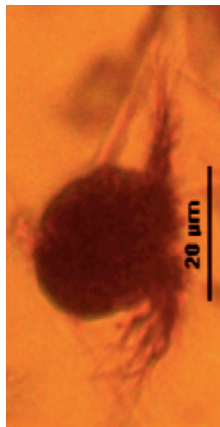
*Strombilidium spiralis*



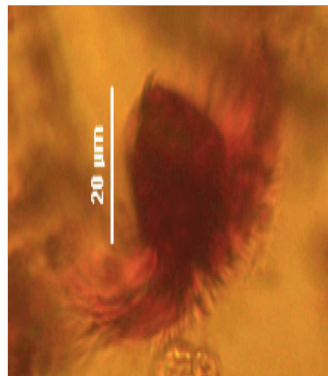
*Strombilidium sphearicum*



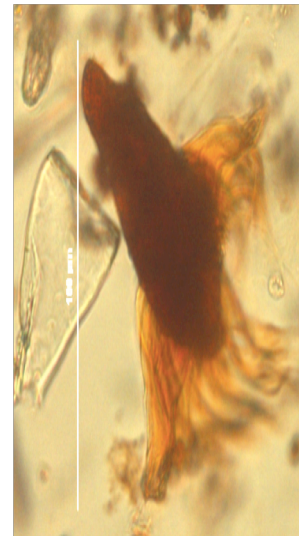
*Strombilidium elegans??*



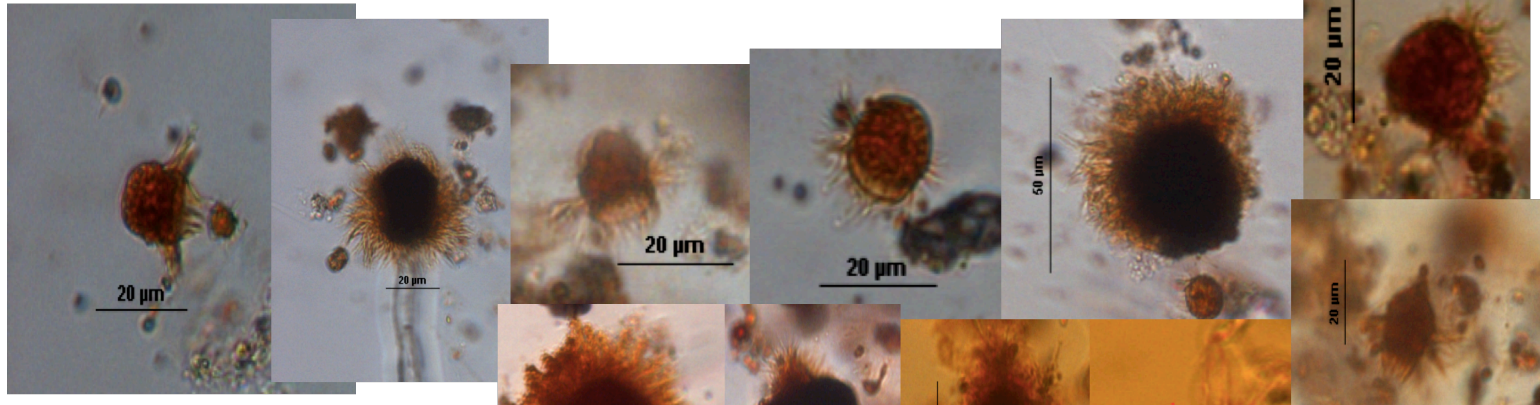
*Strombilidium neptuni*



*Strombilidium sp.*



*Myrionecta rubra*



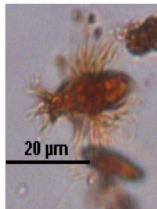
*Myrionecta pulex*



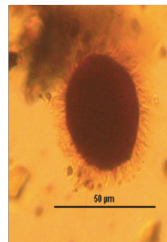
*Unknown*



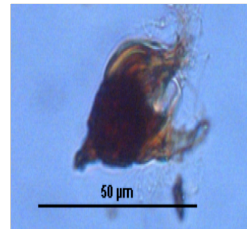
*Unknown*



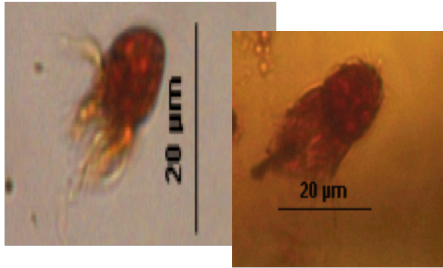
*Ciliés sp*



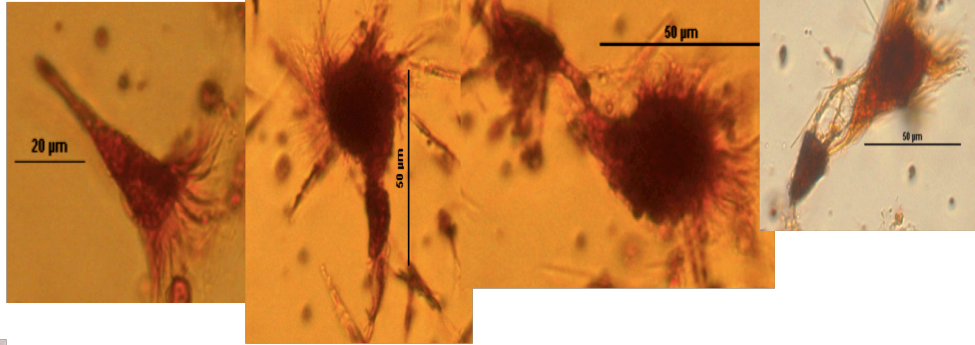
*Ciliés sp*



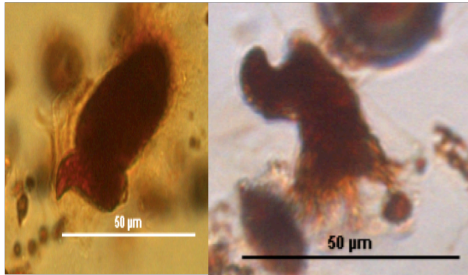
*Lohmanniella oviformis*



*Tontonia* sp.

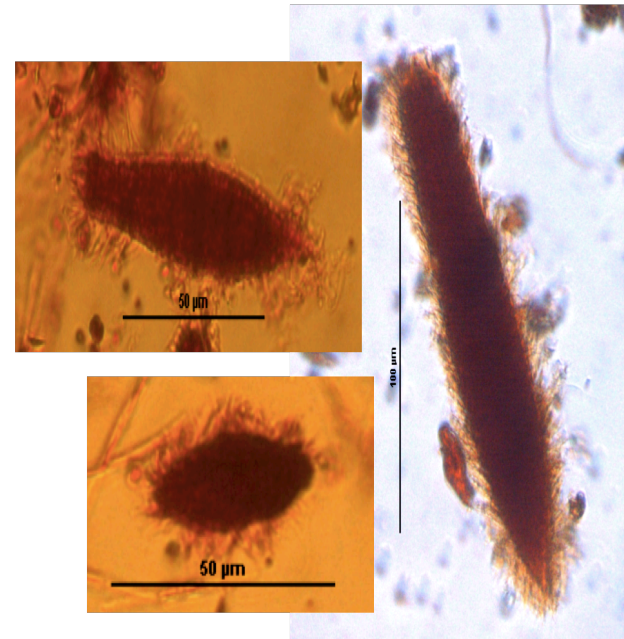
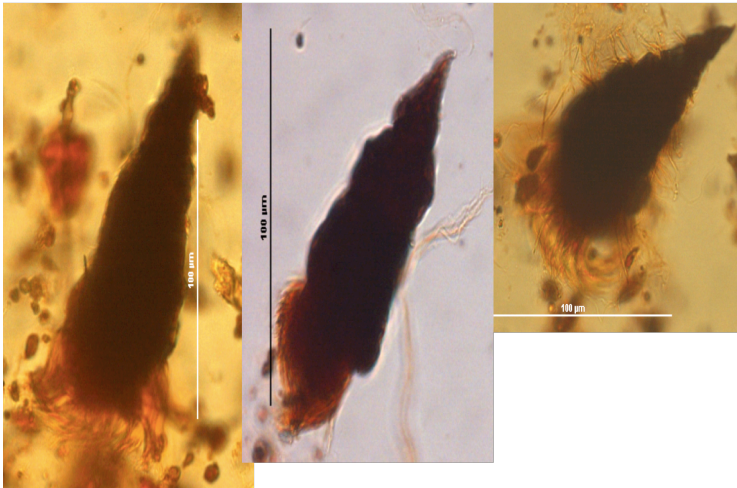


*Tontonia gracillima*

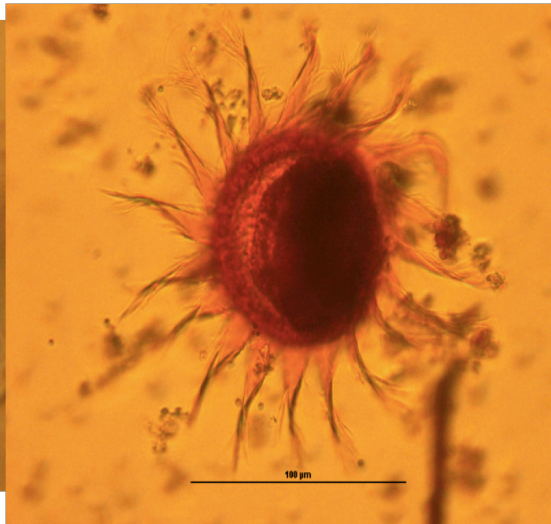
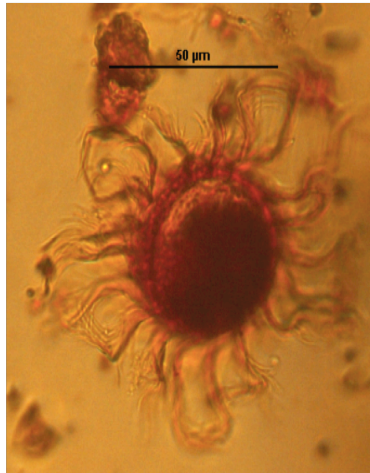


*Frontonia* sp.

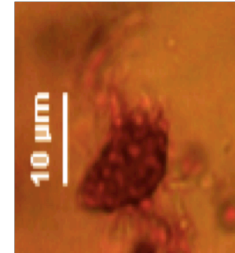
*Laboea strobila*



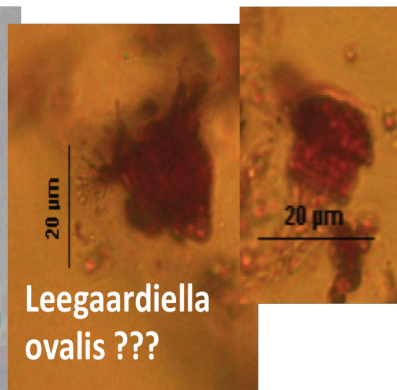
*Leegaardiella sol*



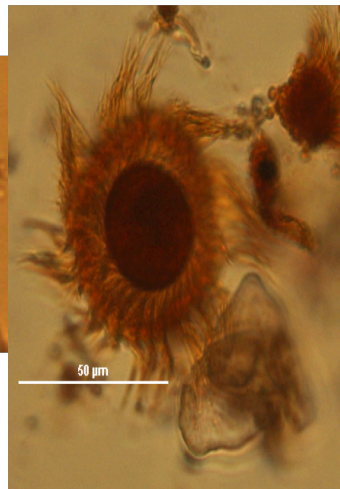
*Leegaardiella ovalis*



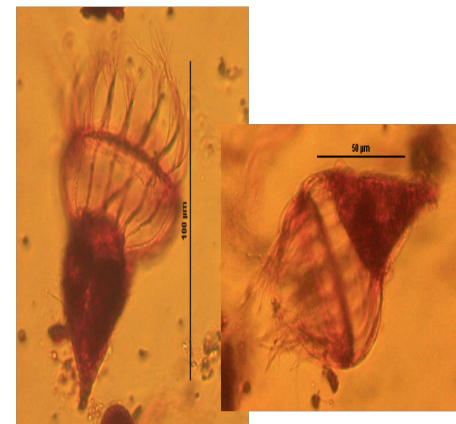
*Leegaardiella sp.*



*Leegaardiella sol ???*



Unknown, *Leegaardiella sp. ??*





# **CHAPITRE VI – DISCUSSION ET PERSPECTIVES**





## CHAPITRE VI – DISCUSSION ET PERSPECTIVES

L'intérêt de comprendre la composition et la structure des populations protistes en réponse aux facteurs biotiques et abiotiques est un des enjeux majeurs de l'écologie. La diversité indénombrable que contient l'écosystème marin et dont un grand nombre est toujours méconnue ainsi que les nombreuses interactions environnementales, rendent ceci difficile. Les premières études de diversité se sont basées sur la microscopie. Cette technique a permis de dénombrer et de caractériser morphologiquement les protistes eucaryotes et d'illustrer certaines fonctions métaboliques aux seins de différents écosystèmes (e.g. marin, dulçaquicole) (voir revue *Caron et al., 2012*). Avec la volonté croissante d'explorer à fine échelle et en profondeur la diversité des micro-organismes, d'autres techniques modernes ont été mises en place telles que les approches moléculaires (e.g. Pyroséquençage, Illumina), permettant une caractérisation globale de tous les groupes phylogénétiques des protistes présents et permettant également d'accéder à la biosphère rare, difficilement identifiable par les approches classiques en raison de leur faible abondance (e.g. microscope). L'utilisation des techniques moléculaires dans ce projet doctoral, dédiées essentiellement à l'exploration de la diversité eucaryote, permet d'obtenir de nouvelles avancées dans les recherches en écologie.

Les objectifs principaux de cette thèse étaient d'examiner la diversité des protistes eucaryotes dans deux écosystèmes contrastés (Manche Orientale et la Méditerranée), de décrire leur dynamique spatiotemporelle et d'identifier les facteurs environnementaux potentiels pouvant influencer leur distribution. Chacun de ses écosystèmes étudiés présente des particularités biologiques et hydrologiques qui lui sont propres, impactant fortement la diversité et la distribution des populations protistes qui est y habitent.

De par nos études et au travers des techniques moléculaires à haut débit, notre but était (i) de renforcer nos connaissances sur la diversité et l'activité des protistes dans un écosystème méso-eutrophe (cas de la Manche Orientale) et mettre également en évidence les successions saisonnières communautaires des protistes eucaryotes pendant le bloom de *Phaeocystis globosa* (*Rachik et al., en presse*) (ii) d'étudier la diversité et la variabilité spatiale des protistes dans un écosystème oligotrophe

contrasté à la Manche Orientale (cas du Golfe du Lion) (*Rachik et al., in prep*). (iii) et enfin, de mettre évidence les relations existantes entre les parasites eucaryotes et les autres taxons ainsi que les paramètres environnementaux (*Christaki et al., 2017*).

## **I. Les outils d'évaluation de la diversité des protistes eucaryotes**

### **1. Difficultés rencontrées pour l'analyse simultanée des données de diversité issues de deux approches NGS différentes (454-pyroséquençage et Illumina)**

Alimentés par la volonté d'évaluer en profondeur la diversité des micro-organismes dans différents écosystèmes, et de pouvoir séquencer de plus en plus de génomes complets, les chercheurs ont développé une nouvelle génération de séquenceurs dits à haut débit opérant en parallèle sur un très grand nombre de séquences courtes et reposant sur de nouvelles technologies physico-chimiques avec des débits jusqu'à 1000 fois supérieurs. Les capacités de ces techniques sont de plus en plus grandes pour un prix de revient de plus en plus faible. La commercialisation de ces techniques n'exigeant plus un passage par une étape de clonage, voit le jour durant les années 2000. Le 454-pyroséquençage (454 GS-FLX Titanium, Roche) est la première de ces technologies de séquençage nouvelle génération (NGS) à avoir été mise sur le marché (*Margulies et al., 2005 ; Shendure & Ji, 2008*). Jusqu'alors, cette technique était la plus utilisée pour explorer la diversité microbienne (*Zinger et al., 2011*), en raison de sa capacité à générer des séquences relativement plus longues (entre 300 et 500 pb environ, *Gilles et al., 2013*) que les autres approches NGS de l'époque. En 2013, nous avons assisté à l'obsolescence de la méthode 454-pyroséquençage populaire à l'époque et à la montée en puissance du séquençage Illumina. Cette technique permet de produire environ 1 milliard de bases séquencées pour un run de quelques heures, à un jour, et surtout pour un coût nettement moindre qu'avec la méthode 454-pyroséquençage. Couplé à du multiplexage, ces approches NGS (pyroséquençage et Illumina) permettent de traiter plusieurs échantillons simultanément afin d'accéder et de comparer la diversité des espèces dans différents échantillons, à différentes dates, et pour différents sites.

Durant notre étude de la diversité et l'activité des communautés microbiennes en Manche Orientale (Chapitre 3), le 454-pyroséquençage utilisé lors du premier suivi

(2011-2013) a été progressivement abandonné au profit d'Illumina utilisée lors du second suivi (2013-2015). Ainsi, afin de réaliser un bilan de la diversité microbienne sur quatre ans (de Mars 2011 à Juillet 2015), nous devions combiner les données de séquences Illumina issues de l'étude (réalisée entre 2013 et 2015) avec celles précédemment obtenues par pyroséquençage entre Mars 2011 à Juillet 2013 (GenBlanc-SRA accession SRX768577, *Genitsaris et al., 2015*). Ce travail, ayant pour but d'explorer l'ensemble des données au sein d'une analyse commune, a été fastidieux et a demandé beaucoup de temps, surtout en terme de curation manuel pour l'affiliation des espèces.

- *Optimisation de l'analyse simultanée des deux jeux de données*

Dans un premier temps, nous avons souhaité réaliser une vue d'ensemble de la diversité, issues du séquençage à haut débit, en Manche Orientale sur une période allant de 2008 à 2015 (*Monchy et al., 2012 ; Christaki et al., 2014 ; Genitsaris et al., 2015, 2016 et Rachik et al., en presse*). Tout d'abord, nous avons réalisé des tests de regroupement des différents jeux de données à plusieurs niveaux de la procédure MOTHUR<sup>®</sup> de traitement des données de séquençage à haut débit. Le regroupement des jeux de données s'est fait après *denoising* (commande *shhh.flows*) pour le pyroséquençage et après assemblage des *paired end* pour Illumina. A cette étape 5.326.486 séquences ont été compilées, dont 2.663.244 séquences de moins de 267 pb (paires de bases) ont été conservées et 2.663.242 séquences de plus de 267 pb (**Tableau 3**). Après compilation des différents jeux de données, les séquences conservées ont été alignées sur un alignement de référence via la commande *align.seqs* de MOTHUR<sup>®</sup> (**Tableau 4**). Une fois l'alignement réalisé, un problème dans le choix des bornes de l'alignement allait survenir du fait des différences dans la taille des amorces et dans la taille des séquences générées entre les approches Illumina et 454-pyroséquençage. La position de départ +19 pb a été choisie afin de conserver un maximum de séquences issues à la fois de l'approche pyroséquençage et Illumina (un choix < à +19 pb excluait d'office les séquences issues de l'approche Illumina), Pour la position 3' de fin d'alignement, plusieurs bornes ont été testées : +451 pb (valeur pour laquelle 97.5% des séquences alignées semblent être conservées), + 636 pb (75%), +1552 pb (50%) et +2578 pb (25%) (**Tableau 5**). A cette étape, la borne d'alignement end +1552 a été choisie en complément de la borne

start +19. Ce bornage avait l'avantage de conserver environ la moitié des séquences (2 336 902 reads sur un total de 5 326 486 reads) en obtenant des séquences d'une taille d'environ 267 pb. Par comparaison avec les bornes +451 et +636, la taille des séquences n'était que de respectivement 92 pb et 109 pb. A l'opposé, un bornage en 3' > à 1552 pb excluait la plupart des séquences issues du pyroséquençage du fait de leurs tailles plus courtes (**Tableau 5**). Afin de valider définitivement le choix des bornes de l'alignement, un tableau du pourcentage de séquences dans un échantillon en fonction du nombre total de séquences a été réalisé pour chaque jeu de données afin de s'assurer de l'homogénéité dans l'origine des séquences. Une fois le bornage optimisé, la procédure s'est poursuivie normalement jusqu'à l'obtention des OTUs et des séquences correspondantes. Pour finir, le bornage choisi (+19pb, +1552 pb) a permis d'obtenir 3430 OTU.

- *Validation des jeux de données consolidés*

Une fois l'ensemble des séquences de chaque OTU récupéré, une affiliation préliminaire a été faite par comparaison la base de données PR2 (the Protist Ribosomal Reference Database). Seules les affiliations des 3430 OTUs issues de l'alignement borné à +1552 ont été utilisées, même si nous avons quand même réalisé des affiliations des tests d'alignement aux bornes de fin à +451 et à +636. Une fois les affiliations connues, nous avons créé un fichier complet avec l'affiliation des 3430 OTUs, le nombre de séquences par échantillons et par OTU et quelques statistiques (abondances relatives, statistiques d'affiliation). L'étape qui s'en suivait concernait la validation du fichier final non normalisé afin de s'assurer que le résultat est concluant. Pour les données Illumina, nous avons comparé la table de contingence des données Illumina analysées seules, avec celles du sous jeu de données Illumina (2013-2015) obtenues après regroupement avec les données de pyroséquençage. Pour cela, les abondances relatives des 20 OTU les plus abondantes ont été calculées et des courbes de successions temporelles ont été réalisées. Lorsque les deux modèles de courbes (données originales et celles issues de la fusion des données Illumina et 454-pyroséquençage) sont identiques, la comparaison est classée comme validée ; lorsqu'un ou deux pics ne sont retrouvés que dans une seule courbe, elle est classée en passable ; lorsque les deux modèles sont incompatibles, elle est classée comme

invalidée. Suite à ces comparaisons, 16 OTUs sur les 20 OTUs les plus abondantes possédaient des courbes de successions temporelles similaires.

Après validation des différents sous-jeux de données (Monchy *et al.*, 2012 ; Christaki *et al.*, 2014 ; Genitsaris *et al.*, 2015, 2016 et Rachik *et al.*, en presse), nous avons procédé à la normalisation des données afin de tenir compte des inévitables différences dans le nombre de séquences par échantillon issues de différentes méthodes de séquençage à haut débit (pyroséquençage et Illumina). Etant donné que la quantité d'OTU douteuse est corrélée au nombre de séquence, il est possible de standardiser le nombre de séquences par échantillon (pour qu'ils aient tous le même nombre de séquences) afin de lisser de manière homogène la probabilité d'accepter de fausses OTUs au sein de chaque échantillon. La normalisation a été réalisée via la commande *sub.sample* sur MOTHUR<sup>®</sup>. A ce niveau différents tests ont été réalisés pour pouvoir choisir la valeur de normalisation permettant de conserver le maximum d'échantillons (i.e. 1500 reads, 2000 reads, 5000 reads, 6115 reads et 12000 reads). Après comparaison des nombres de séquences avant et après chaque normalisation ainsi que le nombre d'échantillons perdus, nous avons décidé de normaliser les données à 5399 séquences. Cette normalisation nous a permis de conserver 100 échantillons pour 2423 OTU.

- *Optimisation de l'affiliation des séquences*

Une fois la table de contingence finale obtenue, il reste à affilier les séquences à des espèces. Pour cela, nous avons affilié les OTUs avec deux bases de données de référence : SILVA (<https://www.arb-silva.de/>) et PR2 (Guillou *et al.*, 2013). Pour cela, chaque OTU a été « Blastée » sur chaque base de références, et les 10 meilleurs résultats BLAST ont été conservés. Nous avons procédé manuellement au choix du meilleur résultat, parmi les 10, en se basant sur le pourcentage de similarité (100%, 99%, ...), l'Evaluate (la plus proche possible de 0) et sur notre connaissance de l'écosystème (données d'observation du microscope). Ensuite, les affiliations retenues pour chaque base de donnée (PR2 et SILVA) ont été comparées entre elles. Pour la majorité des cas, les affiliations de PR2 ont été conservées au détriment de SILVA, lorsque les deux bases de données n'étaient pas en concordance ni dans l'affiliation ni dans l'identifiant de Blast. En effet, PR2 est une base de donnée spécifiquement

conçue pour l'étude de la diversité des microorganismes marins, et les affiliations obtenues semblaient davantage cohérentes au regard de notre connaissance de l'écosystème (i.e. observation au microscope). Pour les autres OTU douteuses ou non affiliées par PR2 ou SILVA, une analyse BLAST dédiée a été réalisée sur la base de données complète du NCBI (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>).

- *Cohérence des jeux de données de diversité consolidés*

Au final, après avoir combiné les données de 454-pyrosequencing et d'Illumina et les avoir analysé *in silico*, il est apparu que les deux jeux de données ne peuvent pas être directement regroupés et analysés ensemble. En effet, bien que techniquement possible, et malgré les optimisations réalisées, la mise en commun des données issues de deux approches de séquençage à haut débit a conduit à des reconstructions communautaires en fonction des méthodes de séquençage et non en fonction de la structuration des communautés par date. Bien que décevant, ce résultat n'est pas surprenant du fait que les deux méthodes présentent des différences en termes de préparations de bibliothèques, de longueur et de quantités de séquences générées. Etant donné, que la comparaison entre les deux techniques ne faisait pas l'objet de cette étude, nous avons décidé d'analyser indépendamment les données, puis de comparer les résultats obtenus. Pour cela, nous avons traité bioinformatiquement les deux jeux de données (454-pyrosequencing et Illumina) indépendamment (voir section I.1.3, Chapitre II), et avons procédé aux mêmes optimisations (en particulier en ce qui concerne l'affiliation des OTUs). Une fois l'affiliation optimisée, nous avons procédé à la comparaison des résultats obtenus pour la diversité et la succession des espèces (*Rachik et al.*, soumis). En conclusion, même si nous n'avons pas réussi à fusionner les deux jeux de données issus d'approches NGS différentes, l'analyse indépendante de données de diversité a produit des résultats très comparables. Cette estimation similaire de la structure de la communauté microbienne, par approche pyrosequencing et Illumina, a déjà été observée en Méditerranée lors de l'étude de la diversité des bactéries et du picoplancton eucaryote marins (Ferrera et al., 2016).

Quartiles	Base de début	Base de fin	Longueur (en pb)	Polymères	Sommes des séquences
0	1	35	35	2	1
2,50%	1	82	82	3	133163
25%	1	125	125	3	1331622
50%	1	267	267	4	2663244
75%	1	474	474	4	3994865
97,50%	1	482	482	6	5193324
100%	1	540	540	287	5326486
Moyenne:	1	271.416	271.416	4.11583	
Séquences uniques	639336				
Nombre total de séquences	5326486				

**Tableau 3 :** Statistiques relatives aux séquences présentes après compilation des 4 jeux de données issus du séquençage d'eucaryotes microbiens de Manche Orientale par NGS (pyroséquençage et Illumina).

Quartiles	Base de début	Base de fin	Longueur (en pb)	Polymères	Sommes des séquences
0	-1	-1	0	1	1
2,50%	1	451	11	2	133163
25%	1	636	125	3	1331622
50%	1	1552	266	4	2663244
75%	19	2578	473	4	3994865
97,50%	2568	2578	482	6	5193324
100%	2578	2578	484	134	5326486
Moyenne:	96.7241	1530.21	266.102	4.01596	
Séquences uniques	639336				
Nombre total de séquences	5326486				

**Tableau 4 :** Statistiques relatives aux séquences présentes après alignement des 4 jeux de données compilés issus du séquençage d'eucaryotes microbiens de Manche Orientale par NGS (pyroséquençage et Illumina).

Borne end	Groupes enlevés	Nombre de Séquences	Nombre de Séquences uniques	Taille moyenne séquences	Projet(s) Gardé(s)	Nombre de Single singletons	Nombre d'OTU
+451	0	4689575	11320	92,0796	Tous	NA	NA
+636	0	4309600	11709	108,837	Tous	53	5851
+1552	0	2336902	11781	266,996	Tous	21	3430
+2578	62	1433083	11424	461,326	Projet 2013-2015	17	1996

**Tableau 5** : Statistiques relatives aux séquences retenues lors de différents tests de bornes d'alignement. Paires de bases utilisées comme borne côté 5' (borne end), nombre de groupes de séquences enlevées par MOTHUR (groupes enlevés), nombre de séquences totales et de séquences uniques, la taille moyenne des séquences, les projets conservées lors des différents tests de bornes d'alignements, le nombre de séquences uniques retrouvées ainsi que d'OTU identifiées lors de chaque test. Les séquences ont été obtenues par compilation de 4 jeux de données issus du séquençage NGS (pyroséquençage et Illumina) de l'ADNr de protistes de Manche Orientale entre 2008 et 2015. Les valeurs de borne end test utilisées correspondent à différents quartiles de conservation des séquences : 97.5% (+451), 75% (+636), 50% (+1552) et 25% (+2578). Tous ces tests ont été réalisés en utilisant une borne start de +19.

## 2. Les limites des approches moléculaires à haut débit : 454-pyroséquençage et Illumina, et nécessité des méthodes classiques

Dès leur apparition, les approches moléculaires ont permis de s'affranchir de l'étape de la mise en culture et d'améliorer notre compréhension de la diversité des protistes tout en intégrant la composante rare (*Diez et al., 2001; Diez et al., 2004; Lovejoy et al., 2006; Stoeck et al., 2006; Lovejoy et al., 2007; Stoeck et al., 2010; Lovejoy & Potvin, 2011; Logares et al., 2014*), ainsi que les organismes pico- et nanoplancton (*Dawson et al., 2009*), permettant ainsi de découvrir un large pan d'organismes sans représentants cultivés (*Moreira et López-García, 2002*). Néanmoins, ces techniques présentent certaines limites. En termes de qualité, certaines OTU considérées comme des lignées divergentes peuvent être éventuellement artefactuelles (e.g. chimères, erreurs de séquençage (*Kunin et al., 2010*)). En termes de quantité, pour certains organismes, l'extraction d'ADN est plus difficile que chez d'autres, les amorces universelles utilisées n'amplifient pas aussi bien les marqueurs moléculaires de tous les organismes (*Zinger et al., 2011*) et également, le nombre de copies d'ADN18S dans les génomes varie d'une espèce à une autre.

Dans les deux zones d'études (Manche Orientale et Méditerranée), l'amplification des régions V2 et V3 du gène codant pour l'ARNr 18S a été réalisée par l'utilisation des amorces universelles d'eucaryotes : 18S-82F (5'-GAAACTGCGAATGGCTC-3',



*Lopez-Garcia et al., 2003*) et Euk-516r (5'-ACCAGACTTGCCCTCC-3', *Amann et al., 1990*), nous permettant ainsi d'identifier une large diversité de protistes. Néanmoins, ces amorces peuvent surestimer certains groupes en dépit d'autres. Par exemple, dans les deux écosystèmes, nous avons remarqué une surreprésentation des Alvéolaires, particulièrement les dinoflagellés. Cette surestimation peut être expliquée par le nombre important de copies du gène 18S (*Dyal et al., 1995; Zhu et al., 2005*), contrastant avec le nombre supposé faible de copies dans la plupart des organismes flagellés (*Zhu et al., 2005*). A l'opposé, certains groupes comme les diatomées sont mal détectés (*Breton et al., in prep*). Ceci peut être dû à la difficulté de la destruction de leurs frustules en silice afin d'extraire leur ADN (*Medinger et al., 2010*). Il est également possible que l'utilisation des amorces non spécifiques puisse affaiblir la détection de certains organismes. D'avoir recours à plusieurs couples d'amorces permettra d'améliorer l'amplification des séquences et de recouvrer des séquences dans les bases de données (*Stoeck et al., 2006 ; Debroas et al., 2017*).

Finalement, ces approches moléculaires ne peuvent pas remplacer les observations classiques des espèces. En effet, l'affiliation des séquences environnementales d'un taxon ne peut pas avoir lieu sans une description préalable des espèces. De plus, elles ne fournissent aucun renseignement sur la biologie, le métabolisme, ou encore l'écologie des organismes détectés. La microscopie peut donc constituer une approche complémentaire de celle moléculaire. Dans la 3<sup>ème</sup> étude conduite en Méditerranée, les dinoflagellés et les ciliés ont fait l'objet des comptages au microscope inversé. Mis à part que le comptage des diatomées était confié aux collègues de l'Institut Méditerranéen d'Océanologie (MIO), l'importance de se concentrer sur une partie du microplancton permet la spécialisation de l'observateur et ainsi d'éviter certains biais comme la subjectivité dans la classification des genres ambigus.

Dans cette étude, la microscopie a permis d'obtenir des données quantitatives (abondance de chaque organisme), de classer les organismes par taille convenant de façon approximative à des niveaux trophiques (*Sheldon et al., 1972*) et d'estimer les biovolumes des organismes permettant par la suite le calcul de biomasses par l'utilisation de facteurs de conversion en carbone.

Toutefois cette approche classique présente certains biais (Chapitre 2, section II) et ne peut permettre d'avoir une vue d'ensemble de la diversité des protistes dans un

écosystème. En général, les approches classiques couplées à celles moléculaires peuvent permettre de limiter les inconvénients liés à chaque technique et d'obtenir à la fois des données qualitatives et quantitatives.

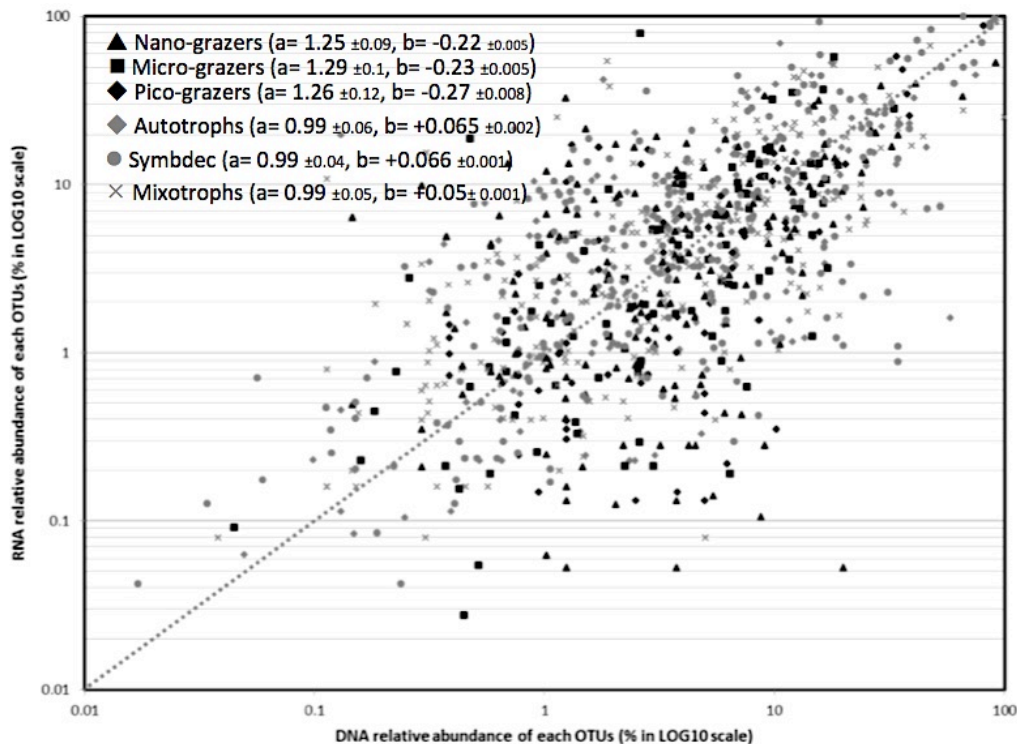
## **II. L'étude de la diversité et l'activité en Manche Orientale (*Rachik et al.*, en presse)**

Cette étude nous a permis d'établir la liste la plus exhaustive possible des espèces microbiennes présentes en Manche Orientale, et de leur dynamique saisonnière durant une période de quatre ans (de 2011 à 2015). En plus, la partie la plus novatrice de ce travail consistait à estimer le rapport rRNA:rDNA. Le nombre des séquences correspondant au ciblage de l'ADNr 18S permet d'estimer l'abondance relative de chaque OTU. De la même manière, le nombre de séquence obtenu en ciblant l'ARNr 18S permet d'estimer l'activité relative d'une cellule en se basant sur le fait qu'une cellule métaboliquement active produit davantage de ribosomes (et donc d'ARNr). Ainsi, le rapport rRNA:rDNA permet une approximation de l'activité métabolique globale de chaque OTU. Plusieurs études récentes ont proposées cette méthode chez les bactéries marines (*Mills et al.*, 2005; *Moeseneder et al.*, 2005; *Gentile et al.*, 2006; *Campbell et al.*, 2011), les archées (*Hugoni et al.*, 2013), et les eucaryotes (*Stoeck et al.*, 2007; *Not et al.*, 2009; *Orsi et al.*, 2013; *Logares et al.*, 2014; *De Vargas et al.*, 2015; *Charvet et Al.*, 2014; *Massana et al.*, 2015; *Hu et al.*, 2016). Toutefois, certaines limitations demeurent quant à l'utilisation de ce ratio, liées essentiellement aux différentes stratégies de vie microbienne ainsi qu'aux activités non liées à la croissance (*Blazewicz et al.*, 2013). Ce rapport ARNr : ADNr pourrait éventuellement ne pas prendre en compte certains événements associés au cycle de vie des espèces (*Blazewicz et al.*, 2013). Par exemple, il a été démontré que certains procaryotes développent une stratégie de vie anticipée (*Tagkopoulos et al.*, 2008; *Mitchell et al.*, 2009; *Barnard et al.*, 2013) consistant à accumuler de l'ARNr dans des conditions environnementales défavorables, en période de faible activité, juste avant que les conditions redeviennent adéquates (*Zinser et al.*, 2009; *Lepp et al.*, 1998). Ceci n'a pas été signalé pour les protistes eucaryotes dans les milieux marins.

Le calcul de ce rapport a montré que la plupart des groupes taxonomiques présentent un rapport 1 :1 suggérant une relation linéaire entre l'abondance (ADNr) et l'activité

(ARNr) (**Chapitre 3, Fig. 6a**) à l'exception de *Ciliophora* et *Chaonoflagelida* qui ont générés davantage de séquences d'ARNr que d'ADNr et inversement pour les *Fungi*. La même analyse réalisée également sur les 100 OTU les plus abondantes, classées cette fois-ci par rôle trophique : autotrophes (e.g. *Bacillariophyta*) ; micro-brouteurs (e.g. *Dinophyceae*) ; mixotrophes (e.g. *Chrysophyceae*, *Dinophyceae*) ; nano-brouteurs (e.g. *Spirotricheae*, *Litostomateae*) ; pico-brouteurs (e.g. *Choanoflagellates*) et parasites (SymbDec, e.g. *Syndiniales*, *Oomycota*, *Chytridiomycota*), avait montré les mêmes prévalences que celles des groupes taxonomiques (**Fig. 23**). Les nano-micro et pico-brouteurs possédaient un ratio moyen d'ARNr : ADNr  $> 1$  ( $a > 1.25$ , **Fig. 23**) suggérant une production d'ARNr supérieure à celle d'ADNr. Au contraire des autotrophes, des SymbDec et des mixotrophes qui semblaient plutôt pencher vers la catégorie prédominante de l'ADNr (plus de séquences d'ADNr que d'ARNr). Dans l'ensemble, les groupes trophiques présentaient un rapport ARNr : ADNr moyen d'environ 1 :1 (entre 0.9 et 1.25), suggérant une relation linéaire entre l'abondance des cellules (ADNr) et les activités ARNr. Ce ratio montre plus de variabilité à l'échelle de l'OTU et s'avère ainsi plus informatif si on traitait chaque OTU individuellement (par exemple cas de *Pheocystis globosa*) que quand les données sont rassemblées par groupes trophiques ou taxonomiques.

En conclusion, dans cette étude, le ciblage simultané d'ARNr et l'ADNr a, en outre, permis de démontrer que le rapport rRNA: rDNA peut être utilisé comme indicateur des transitions écologiques d'espèces microbiennes 'clés' (par exemple ici *Pheocystis globosa*). Ainsi, l'utilisation de ce rapport pour chaque OTU devrait permettre d'apporter une information supplémentaire indispensable pour une meilleure compréhension du fonctionnement de l'écosystème et de l'influence des paramètres biotiques et abiotiques sur la structuration des communautés microbiennes. Il devrait aboutir à mieux comprendre la structuration future des communautés face aux changements globaux.



**Figure 23 :** Relation entre l'abondance (séquences de l'ADNr) et l'activité cellulaire relative (séquences de l'ARNr) pour les 100 OTU les plus abondantes classées par fonction trophique. Chaque OTU a été tracé en fonction de son nombre relatif d'ADNr (axe des abscisses) et d'ARNr (axe des y) sur le nombre total de séquences. Chaque point représente des pourcentages appariés (en échelles logarithmiques) de chaque groupe trophique de chaque échantillon.

### III. L'étude de la diversité des protistes eucaryotes dans le Golfe du Lion (*Rachik et al., in prep*)

Dans cette étude, l'utilisation d'Illumina nous a permis d'identifier une diversité de 3358 OTU dans toutes les stations d'où les échantillons ont été prélevés, réparties en huit super-groupes (Alveolata, Amoebozoa, Apusozoa, Archeplastida, Hacrobia, Opisthokonta, Rhizaria et Stramenopiles) et d'autres eucaryotes, et affiliées à 48 groupes taxonomiques supérieurs (**Chapitre 5, Annexe 1**), pendant différentes saisons de l'année ainsi que d'évaluer par la suite la variabilité spatiale des protistes au dépend des variables environnementales. De par cette étude, nous avons pu établir la liste la plus exhaustive possible de la diversité dans le Golfe du Lion, de l'ensemble de la communauté planctonique alors que la plupart des études conduites dans la zone d'étude ou dans les environs, ne concernaient sur des groupes spécifiques (e.g toxiques, invasives, indicatrices) étudiés via des méthodes classiques (microscopie,

culture). Le couplage entre le séquençage à haut débit et la microscopie a permis d'étudier la diversité en ralliant les aspects quantitatifs et qualitatifs et de valider entre autre les résultats issus du séquençage du microzooplancton (au moins pour les espèces identifiées).

La difficulté rencontrée dans cette étude est en premier temps les dates d'échantillonnage distantes et différentes d'une station à une autre, ne permettant pas de comparer la diversité entre les stations pour chaque date. Un échantillonnage synchronisé aurait permis de mieux déceler les différences, les similarités ainsi que le mouvement des communautés protistes sous contraintes hydrodynamiques (e.g. vent, courants secondaires). Dans cette zone fortement soumise aux intrusions du Rhône, l'explication de la variabilité spatiale des groupes taxonomiques s'avérait difficile en tenant compte seulement des variables environnementales. Suite à la réalisation des analyses de co-inertie sur les variables environnementales et les principaux groupes taxonomiques, nous avons constaté les différents groupes taxonomiques sont associés à différentes conditions environnementales dans les différentes stations. Ceci suggère qu'il existe d'autres facteurs influençant cette répartition. Il se peut qu'une intégration des données physiques du lieu d'études (e.g. les jours d'intrusions, type d'intrusion) ainsi que les interactions biotiques puisse améliorer notre compréhension de la dynamique des populations protistes.

Toutefois, les données issues joueront certainement un rôle dans les prochaines études conduites dans le Golfe du Lion.

#### **IV. Quelques aspects de similitude de dissimilitude entre la Manche Orientale et la Méditerranée Occidentale**

- *Caractéristiques physiques et diversité taxonomique*

A l'instar des résultats obtenus dans les deux écosystèmes, certains points de ressemblances et de différences ressortaient (**Tableau 6**).

A l'échelle des caractéristiques physiques, les deux écosystèmes présentaient des particularités qui leur sont propres. Le bassin nord-occidental de la Méditerranée est reconnu en tant que relativement productif quoique le système reste oligotrophique stratifié (Ludwig et al., 2009). Son fonctionnement biogéochimique est influencé par les apports du Rhône, constituant la principale source d'eau douce et de nutriments. Ces entrées influencent directement sur les variables environnementales et ainsi sur la production primaire. Il a été estimé que l'apport en nutriments du Rhône pourrait supporter entre 23% et 69% de la production primaire moyenne dans le Golfe du Lion (Ludwig et al., 2009). Cet écosystème est également sous l'influence des forçages exercés par les vents (vent nord-nord-ouest, qui favorisent les remontées d'eau, (upwelling) et les vents du sud-est, qui par contre favorisent le downwelling) et le Courant Nord (CN) (**Tableau 6**). En raison des latitudes relativement basses, la Méditerranée montre naturellement une faible saisonnalité, cependant les forçages hydrodynamiques et les intrusions du Rhône entraînent une difficulté supplémentaire quant à l'observation d'une nette saisonnalité. Quant à la Manche Orientale, est un écosystème méso-eutrophe connu par son importante productivité (Brylinski et al., 1996). Cet écosystème est sous l'influence des marées, le vent et les vagues et alimenté par un courant résiduel de marée dirigé vers la Mer du Nord, riche en éléments transportés (matières en suspension : MES (Guéguéniat et al., 1995; Lafite et al., 2000 ; cellules algales, Hault et al., 1994 ; sels nutritifs, Bentley et al., 1993). Cette région est annuellement sujette à une prolifération de *Phaeocystis globosa* au printemps (par exemple : Brylinski et al., 1984 ; Brunet et al., 1992, Brunet et al., 1996 ; Hoch et Garreau, 1998 ; Breton et al., 2000 ; Monchy et al., 2012) liée à un déséquilibre dans les sels nutritifs, caractérisé par un excès d'azote et une limitation en silicate (par exemple : Peperzak et al., 1998 ; Schapira et al., 2008). En plus des conditions environnementales, cette espèce régit la structuration des communautés protistes eucaryotes, plus particulièrement, les dinoflagellés, les ciliés et les symbiotes/décomposeurs (parasites) (Grattepanche et al., 2011 a, b ; Christaki et al., 2017).

De point de vue biologique, les études moléculaires dans les deux écosystèmes nous ont permis d'identifier un large pan de diversité des protistes eucaryotes et leur structuration en relation avec les conditions environnementales. Dans cette partie, le but a été de regarder l'assemblage des groupes taxonomiques dans sa globalité et de

ressortir ainsi les ressemblances et/ou les différences dans la structuration des protistes dans les deux écosystèmes contrastes. Dans cet inventaire de diversité (**Tableau 6**) seules les données d'Illumina (2013-2015) du côté de la Manche Orientale ont été utilisées afin que ça soit en cohérence avec les données d'Illumina de la Méditerranée Occidentale.

Les mêmes super-groupes identifiés ont été détectés dans les deux études (Alveolata, Stramenopiles, Opisthokonta, Archaeplastida, Hacrobia et Apusozoa). Un total de 38 et de 25 sous-groupes taxonomiques a été identifié, pour la Méditerranée Occidentale et la Manche Orientale, respectivement (**Tableau 6**). Il est à noter que le nombre supérieur des groupes taxonomiques en Méditerranée Occidentale par rapport à la Manche Orientale peut être lié au nombre d'échantillons traités (59 et 30 échantillons pour la Méditerranée Occidentale et la Manche Orientale, respectivement), les périodes différentes d'échantillonnage (2012-2014 et 2013-2015, **Tableau 6**) ainsi que la variété des stations étudiées (SOLA, MOLA, SOLEMIO et JULIO dans le Golfe du Lion).

La diversité dans les deux zones d'études a été dominée, en termes de nombre des séquences et d'OTU, par les Alvéolaires, représentés majoritairement par les Syndiniales (MALVI et MALV II) et les dinoflagellés du genre *Gyrodinium*. La surreprésentation de ce groupe a été observée également dans d'autres écosystèmes (par exemple : Océan Austral ; *Georges et al., 2014*). Les Syndiniales sont qualifiés d'ubiquistes, leur présence a été signalée dans plusieurs études dans différents écosystèmes marins (*López-García et al., 2001, Massana et Pedrós-Alió, 2008 ; Georges et al., 2014 ; Genitsaris et al., 2015, 2016 ; Christaki et al. 2017*). Dans les deux analyses, les MALV I et le genre *Gyrodinium* ont été les deux groupes ayant la richesse la plus élevée en OTU. Christaki et al. (2017) ont suggéré l'existence d'un couplage entre les Syndiniales (MALV) et les Dinophycées, plus particulièrement le genre *Gymnodinium*. D'autres études ont également souligné l'interaction des Syndiniales avec les dinoflagellés ainsi qu'avec d'autres hôtes (e.g. radiolaires (observés uniquement dans les échantillons du Golfe du Lion), ciliés) (*Skovgaard et al., 2005; Harada et al., 2007 ; Guillou et al., 2008 ; Brate et al., 2012*).

En termes de séquences, les ciliés étaient peu représentés dans les échantillons du Golfe du Lion contrairement à la Manche Orientale (1-2% et 13%, respectivement). Toutefois, ce groupe présentait dans l'ensemble une grande diversité en termes d'OTU (> 80 OTU). Dans les deux jeux de données, les oligotriches nues du genre *Strombidium* dominaient la composition taxonomique des ciliés. Ce groupe joue un rôle clef dans le transfert de la biomasse bactérienne et phytoplanctonique aux niveaux trophiques supérieurs par le broutage. En dehors des ciliés herbivores, plusieurs OTU appartenant à des genres ou espèces mixotrophes (comme par exemple : *Laboea strobila*, *Strombidium conicum*, *Strombidium basimorphum*) ont été rencontrées dans les deux jeux de données. Les ciliés mixotrophes sont présents aussi bien dans les eaux plutôt oligotrophes, comme par exemple dans la mer de Ligurie (Perez et al., 2001) ou encore en mer Méditerranée le long d'un transect ouest-est (Pitta et al., 2001) que dans les eaux plus riches de la Méditerranée Occidentale, comme par exemple les lentilles relativement dessalées du Rhône (Christaki et al., 2009).

Dans les deux jeux de données, les Bacillariophycées (diatomées) montraient une diversité considérable, en termes d'OTU, dans la Méditerranée Occidentale par rapport à la Manche Orientale (66 OTU, 33 OTU, 76 OTU, 77 OTU, à SOLA, MOLA, SOLEMIO et JULIO dans la Méditerranée Occidentale, respectivement et 41 OTU en Manche Orientale). Et quoique la faible diversité des diatomées en Manche Orientale, leur abondance relative fusse bien plus importante que dans le Golfe du Lion. En Manche Orientale, plusieurs études ont indiqué l'importance des diatomées en termes d'abondances et de biomasses, plus particulièrement en dehors du bloom de *Phaeocystis globosa* (fin d'hiver-Mi-printemps) (Breton et al., 2000,2017 ; Gomez and Souissi, 2007; Schapira et al., 2008; Lefebvre et al., 2011; Grattepanche et al., 2011a, b; Hernandez Farinas et al., 2013 ; Bonato et al., 2015).

Nous avons également observé un nombre d'OTU plus important des endosymbiotes/décomposeurs (parasites) de la matière organique (e.g MALV, Cercozoa, Oomycota, Fungi, etc.) en Manche Orientale par rapport à la Méditerranéen, ce qui indique l'importance des symbiotes/décomposeurs (parasites) dans des conditions plus eutrophes. Ceci peut être du à la biomasse élevée, à l'abondance des hôtes et de la matière organique à dégrader (Christaki et al., 2014 ;



2017). Ce type de différence a été également observé dans l’Océan Austral. Au cours d’une étude, conduite aux îles de Kerguelen dans le cadre du projet KEOPS2 (Kerguelen Ocean and Plateau compared Study), Georges et al. (2015) ont souligné également un nombre et richesse en OTU plus importants des communautés d’OUT potentiellement endosymbiotes/décomposeurs (parasites) dans les stations plus riche en biomasse phytoplanctonique.

- *Rôle des principaux groupes taxonomiques vis à vis des particularités de chaque écosystème*

Dans les écosystèmes dominés par les blooms de *Phaeocystis globosa*, tel le cas de la Manche Orientale, il a été démontré que le réseau trophique s’adapte aux cycles de vie de l’espèce qui alterne entre stade de vie nano-planctonique (cellules libres coloniales et flagellés) et micro-planctonique (colonies) (Grattepanche et al. 2011 a, b). Cette espèce peut représenter jusqu’à 90% de la biomasse du phytoplancton et régit ainsi la structuration des communautés planctoniques (par exemple Grattepanche et al. 2011 a, b et Breton et al., 2017). En effet, la présence élevée des ciliés ou des dinoflagellés dépendent effectivement du cycle de vie de *P. globosa*. Les ciliés sont des prédateurs du nanoplancton et donc des cellules libres de *P. globosa*, quant aux dinoflagellés hétérotrophes, ils consomment plutôt du microplancton (Hansen et al., 1994). En effet, des études expérimentales et microscopiques antérieures en Manche Orientale ont identifié les dinoflagellés hétérotrophes (*Gyrodinium sp.*, et *Peridinium sp.*) comme les principaux prédateurs des diatomées et les colonies de *P. globosa* (Grattepanche et al. 2011 a, b, **Fig. 24a**). Cependant, ces protistes hétérotrophes sont soumis à la prédation particulièrement par les copépodes (top-down control, **Fig. 24a**) en plus de leur régulation face à la proie en présence (bottom-up control). En conséquence ce contrôle top-down des dinoflagellés par les copépodes peut promouvoir les proliférations de *P. globosa* (Grattepanche et al. 2011a, b, **Fig. 24a**). De plus, ces études ont souligné de plus la dominance des ciliés en fin du bloom qui consomment ainsi les cellules libérées de la désintégration des colonies de *P. globosa*. Dans Christaki et al., (2017), l’étude des relations entre les parasites et le reste de la communauté protiste a montré une succession claire à l’échelle temporelle de cette communauté parasitaire et une relation négative significative avec les dinoflagellés hétérotrophes. Cependant, aucune OTU parasitaire

n'avait une connexion significative avec *P. globosa*. Et bien qu'il soit démontré que les parasites identifiés dans la Manche Orientale semble infecter les diatomées et les dinoflagellés, il n'a y a aucune observation directe pour l'instant qui l'affirme.

En Méditerranée, le phytoplancton est dominé par les cellules de petite taille, ses prédateurs majeurs sont donc les ciliés hétérotrophes (**Fig. 24b**). En plus le nanophytoplancton de taille  $<10\mu\text{m}$  n'est pas efficacement consommé par le mesozooplancton. Ce dernier exercerait donc un 'top-down' control sur les microzooplanctons et permettrait ainsi une certaine prolifération du phytoplancton de petite taille (**Fig. 24b**). D'autant plus que dans le contexte actuel, l'élévation de la température de la surface des océans, induit un appauvrissement de la surface en nutriments et favorise donc d'autant plus la prolifération des cellules de petite taille dans de vastes régions océaniques (e.g. *Behrenfeld et al., 2006; Polovina et al., 2008 ; Moran et al. 2010*).

Le « top-down control » des métazoaires sur le microzooplancton bénéficierait dans les 2 systèmes étudiés au phytoplancton. Alors qu'on a donc une certaine idée de flux de carbone via la consommation par le micro- et mesozooplankton le rôle des parasites dans ce flux est loin d'être compris. Une perspective d'actualité serait d'essayer d'identifier les hôtes des principaux endosymbiotes/parasites identifiés ainsi que d'évaluer leur rôle potentiel dans le flux de carbone.

**Tableau 6 :** (a) Bilan des caractéristiques physiques et biologiques en Manche Orientale et en Méditerranée Occidentale (b) Inventaire de la diversité obtenue via le séquençage à haut débit (Illumina) dans les deux écosystèmes : Manche Orientale et Méditerranée Occidentale à différentes périodes (2013-2015 et 2012-2014, respectivement). Les groupes taxonomiques soulignés ont été observés uniquement en Méditerranée Occidentale. Les groupes taxonomiques marqués en gras ont été observés uniquement en Manche Orientale. ND : groupe taxonomique non détecté.

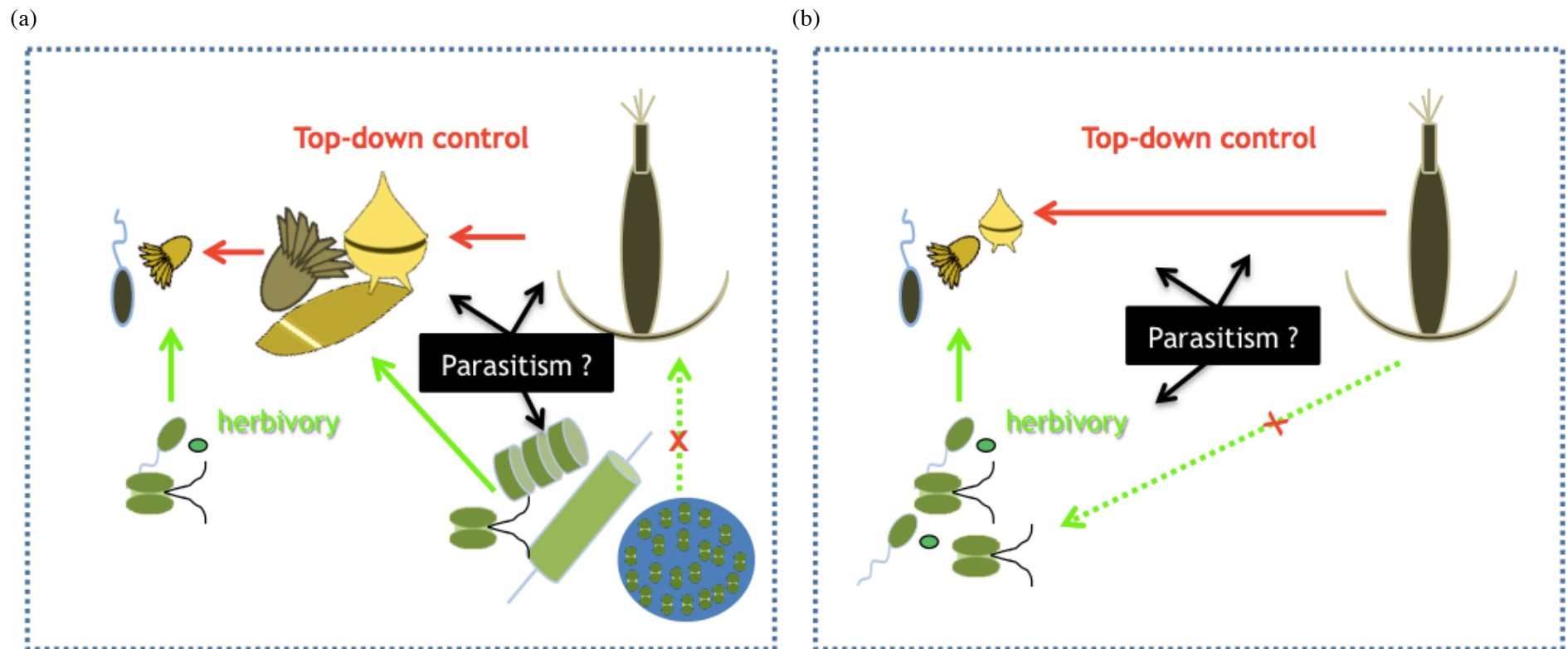
<b>(a)</b>	<b>Manche Orientale</b>	<b>Méditerranée Occidentale (Golfe du Lion)</b>
<b>Caractéristiques physiques</b>	- Mésio-eutrophe - Colonne d'eau stratifiée	- Oligotrophe - Colonne d'eau non stratifiée
<b>Facteurs de forçages</b>	- Vents (Sud-ouest et nord-est) - Forts courant de marée - Apports continentaux importants - Apports fluviaux importants (ex. la Canche, la Somme, l'Authie, etc.)	- Vents (Tramontane, Mistral et vents est / sud-est) - Courant géostrophique (courant nord) - Rhône et sa zone de dilution
<b>Facteurs abiotiques</b>	- Déséquilibre ratio N/P (0-100) - Nitrate en excès	- Déséquilibre ratio N/P > 20 - Phosphate facteur limitant
<b>Particularités biologiques</b>	- Dominance des diatomées - Bloom annuel de <i>Phaeocystis globosa</i> en printemps	- Dominance du pico- nano-phytoplancton
<b>Saisonnalité</b>	Forte saisonnalité	Faible saisonnalité
<b>Périodes d'échantillonnage</b>	2013-2015	2012-2014
<b>Nombre de stations</b>	1	4
<b>Technique de séquençage</b>	Illumina	Illumina
<b>Diversité révélée</b>	7 super-groupes 25 groupes taxonomiques 696 OTU	7 super-groupes 38 groupes taxonomiques 3350 OTU

(b)

<b>Groupes taxonomiques</b>	<b>OTUs dominantes en Manche Orientale</b>	<b>OTUs dominantes en Méditerranée Occidentale</b>
Apicomplexa	<i>Gregarines sp.</i>	<i>Apicomplexa sp.</i>
Dinophyceae	<i>Gyrodinium spirale</i>	<i>Dinophyceae sp. / Gyrodinium sp.</i>
Syndiniales	<i>Dino-GroupI-Cladel</i>	<i>Dino-GroupI-Cladel</i>
Ciliophora	<i>Strombidium basimorphum</i>	<i>Strombidiidae sp.</i>
<u>Perkinsea</u>	<b>ND</b>	<i>Perkinsida sp.</i>
Apusozoa	<i>Amastigomonas mutabilis</i>	<i>Mantamonas plastica</i>
Chlorophyta	<i>Bathycoccus prasinos</i>	<i>Micromonas sp. / Bathycoccus prasinos</i>
<u>Rhodophyta</u>	<b>ND</b>	<i>Peyssonnelia capensis</i>
Streptophyta	<i>Humulus scandens</i>	<i>Cucurbita pepo</i>
Centroheliozoa	<i>Pterocystida sp.</i>	<i>Pterocystida sp.</i>
<u>Cryptophyta</u>	<b>ND</b>	<i>Goniomonas sp.</i>
Haptophyta	<i>Phaeocystis globosa</i>	<i>Chrysochromulina sp.</i>
<u>Katablepharidophyta</u>	<b>ND</b>	<i>Katablepharidales sp.</i>
Picobiliphyta	<i>Picobiliphyta sp.</i>	<i>Picobiliphyta sp.</i>
Telonemia	<i>Telonemia-Group-2-sp.</i>	<i>Telonemia-Group-2-sp.</i>
Choanoflagellida	<i>Stephanoecidae-Group-D-sp.</i>	<i>Monosigidae-Group B sp.</i>
Mesomycetozoa	<i>Ichthyosporea sp.</i>	<i>Abeoformidae-Group-MAIP I sp.</i>

Fungi	<i>Penicillium sp.</i>	<i>Ustilaginomycetes sp.</i>
Cercozoa	<i>Sandonidae sp.</i>	<i>Protaspa-lineage sp.</i>
<u>Radiolaria</u>	<b>ND</b>	<i>Spumellarida-Group-I-X sp.</i>
Bacillariophyta	<i>Leptocylindrus aporus</i>	<i>Chaetoceros tenuissimus</i>
Bicoecea	<i>Cafeteria roenbergensis</i>	<i>Caecitellus parvulus</i>
Bolidophyceae-and-relatives	<i>Bolidophyceae-and-relatives sp.</i>	<i>Bolidomonas sp.</i>
<u>Chrysomerophyceae</u>	<b>ND</b>	<i>Chrysowaernella hieroglyphica</i>
Chrysophyceae-Synurophyceae	<i>Chrysophyceae-Synurophyceae sp.</i>	<i>Clade-G-X sp. / Spumella sp.</i>
<u>Dictyochophyceae</u>	<b>ND</b>	<i>Dictyochales-X sp.</i>
<u>Eustigmatophyceae</u>	<b>ND</b>	<i>Nannochloropsis gaditana</i>
<u>Hyphochytriomyceta</u>	<b>ND</b>	<i>Rhizidiomyces apophysatus</i>
Labyrinthulea	<i>Thraustochytriaceae sp.</i>	<i>Thraustochytriaceae-X sp.</i>
<u>Labyrinthulomycetes</u>	<b>ND</b>	<i>Amphitrema wrightianum</i>
MAST	<i>MAST-1A sp.</i>	<i>MAST-4A sp.</i>
MOCH	<i>MOCH 2 sp.</i>	<i>MOCH 2 sp.</i>
<b>Ochrophyta</b>	<i>Actinoptychus undulatus</i>	<b>ND</b>
Oomycota	<i>Olpidiopsis porphyrae</i>	<i>Pythium pectinolyticum</i>
<u>Pelagophyceae</u>	<b>ND</b>	<i>Pelagophyceae-X sp.</i>
<u>Phaeophyceae</u>	<b>ND</b>	<i>Phaeophyceae-X sp.</i>
<u>Pinguiphyceae</u>	<b>ND</b>	<i>Phaeomonas parva / Pinguiochrysis pyriformis</i>

<u>Pirsonia</u>	<b>ND</b>	<i>Pirsonia-X sp.</i>
<u>Radiophycea</u>	<b>ND</b>	<i>Fibrocapsa japonica</i>



**Figure. 24 :** (a) En Manche Orientale les protistes hétérotrophes, en particulier les dinoflagellés hétérotrophes de grande taille, sont les principaux prédateurs des diatomées et les seuls consommateurs de colonies de *Phaeocystis globosa*. Notre hypothèse est que, l'accumulation de colonies de *Phaeocystis* pourrait être favorisée par le parasitisme sur les prédateurs et les concurrents potentiels de *Phaeocystis*. Les dinoflagellés peuvent alors subir un double contrôle descendant par les copépodes et les parasites. (b) En Méditerranée le phytoplancton de petite taille est consommé par du microzooplancton de petite taille à son tour consommé par le mesozooplancton. De même façon qu'en Manche Orientale, le rôle de parasites et leur potentiel de contrôle dans un contexte de changement global reste à élucider.





## **REFERENCES BIBLIOGRAPHIQUES**



- Adl SM, Simpson AG, Farmer MA, et al**, (2005). The new higher-level classification of eukaryotes with emphasis on the taxonomy of protists. *J Eukaryot Microbiol.* 52, 399- 451.
- Adl SM, Simpson AG, Lane CE, Lukeš J, Bass D, Bowser SS, et al**, (2012). The revised classification of Eukaryotes. *J Eukaryot Microbiol*, 59(5), 429-93.
- Albérola C, Millot C**, (2003). Circulation in the french Mediterranean coastal zone near Mar-seilles: the influence of wind and the Northern Current. *Continental Shelf Research* 23, 587–610.
- Alongi D**, (1998). The coastal ocean. I. The coastal zone. In: Kennish, M.J., Lutz, P.L. (Eds.), *Coastal Ecosystem Processes*. CRC Press, Boca Raton, London, New York, Washington, D.C., pp. 183-247.
- Altschul SF, Gish W, Miller W et al** (1990). Basic local alignment search tool *J. Mol. Biol.*, 215 (1990)
- Amann RI, Blinder J, Olson RJ, Chisholm SW, Devereux R, Stahl DA**, (1990). Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations *Appl. Environ. Microbiol.*, 56, pp. 1919–1925
- Amaral Zettler LA, Gómez F, Zettler E et al**, (2002). Eukaryotic diversity in Spain ' s River of Fire. *Nature*, 417, 137.
- Amaral Zettler LA, McCliment EA, Ducklow HW, Huse SM** (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.
- Aminot A, Chaussepied M** (1983). *Manuel des analyses chimiques en milieu marin*. CNEXO, Brest: 395 pp.
- Aminot A, Kérouel R** (2004). *Hydrologie des écosystèmes marins. Paramètres et analyses*. Ifremer: 336 pp.
- Aminot A, Kérouel R** (2007). *Dosage automatique des nutriments dans les eaux marines : méthodes en flux continu*. Ed. Ifremer, *Méthodes d'analyse en milieu marin* 188p.
- Andersen RA**, (2004). Biology and systematics of heterokont and haptophyte algae. *American journal of botany*, 91, 1508–22.
- André G, Garreau P, Garnier V, Fraunié P** (2005). Modelled variability of the sea surface circulation in the North-Western Mediterranean Sea and in the Gulf of Lions. *Ocean Dynamics* 55, 294–308.
- André G** (2007). *Echanges côte-large à moyenne échelle au niveau de la pente continentale du Golfe du Lion -processus et modélisation-*. Thèse de doctorat, Université du Sud Toulon-Var.
- Archibald JM**, (2005). Jumping genes and shrinking genomes-probing the evolution of eukaryotic photosynthesis with genomics. *IUBMB Life.* 57, 539-47.
- Arisue N, Hasegawa M et Hashimoto T**, (2005). Root of the Eukaryota tree as inferred from combined maximum likelihood analyses of multiple molecular sequence data. *Mol Biol Evol.* 22, 409-20.

- Azam F, Fenchel T, Field JG, Gray JS, Meyer-Reil LA & Thingstad TF** (1983). The Ecological Role of Water-Column Microbes in the Sea. *Mar Ecol Prog Ser* 10, 257-263.
- Baptiste E, Brinkmann H, Lee JA, et al**, (2002). The analysis of 100 genes supports the grouping of three highly divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba. *Proc Natl Acad Sci U S A*. 99, 1414-9.
- Barnard RL, Osborne CA, Firestone MK**, (2013). Responses of soil bacterial and fungal communities to extreme desiccation and rewetting. *ISME J*.
- Batchelder HP, Powell TM**, (2002). Physical and biological conditions and processes in the northeast Pacific Ocean. *Progress in Oceanography*, 53, 105-
- Beaugrand G, Reid PC, Ibanez F, Lindley JA, Edwards M** (2002). Reorganization of North Atlantic Marine copepod biodiversity and climate. *Science*, 296, 1692- 1694.
- Behrenfeld et al**, (2006). Climate-driven trends in contemporary ocean productivity *Nature* 444, 752-755, doi:10.1038/nature05317
- Bendschneider K, Robinson RJ** (1952). A new spectrophotometric determination of nitrate in seawater. *Journal of Marine Research*. 11: 87-96.
- Bentley D** (1985). Caractéristiques physico-chimiques des eaux du détroit du pas de Calais. *J. Rech. Oceanogr*. 10: 69-71.
- Bentley D, Lafite R, Morley NH, James R, Staham PJ, Guary J** (1993). Flux de nutriments entre la Manche et la Mer du Nord. Situation actuelle et évolution depuis dix ans. *Oceanol. Acta*, 16: 599-606.
- Bethoux J, Gentili B, Morin P, Nicolas E, Pierre C and Ruiz-Pino D**, (1999). The Mediterranean Sea: a miniature ocean for climatic and environmental studies and a key for the climatic functioning of the North Atlantic. *Progress in Oceanography*, 44(1-3):131-146. International Conference on Progress in Oceanography of the Mediterranean Sea, ROME, ITALY, NOV 17-19.
- Béthoux P** (1981). Le phosphore et l'azote en Méditerranée, bilans et fertilité potentielle. *Marine Chemistry*, 10: 141-158.
- Blanfune A, Cohu S, Mangialajo L, Lemée R, Thibaut T**, (2012). Preliminary assessments of the impact of *Ostreopsis cf. ovata* (Dinophyceae) development on macroinvertebrates in the North Western Mediterranean Sea. *CRYPTOGAMIE ALGOL*. 33, 129-136.
- Blazewicz SJ, Barnard RL, Daly RA and Firestone MK**, (2013). Evaluating rRNA as an indicator of microbial activity in environmental communities: limitations and uses the *ISME Journal*, 7, pp. 2061-2068, doi: 10.1038/ismej.2013.102
- Bolch CJS, de Salas MF**, (2007). A review of the molecular evidence for ballast water introduction of the toxic dinoflagellates *Gymnodinium catenatum* and the Alexandrium. *Harmful Algae* 6 (4) 465-485.
- Bonato S, Christaki U, Lefebvre A, Lizon F, Thyssen M, Artigas LF** (2015). High spatial variability of phytoplankton assessed by flow cytometry, in a dynamic productive coastal area, in spring: The eastern English Channel, Elsevier
- Bourrin F et Durrieu de Madron X**, (2006). Contribution to the study of coastal rivers and

associated prodeltas to sediment supply in North-western Mediterranean Sea (Gulf of Lions). *Life and Environment* 56(4), 307-314.

**Bourrin F, Durrieu de Madron X, Heussner S et Estournel C**, (2008). Impact of winter dense water formation on shelf sediment erosion (evidence from the Gulf of Lions, (NW Mediterranean). *Continental Shelf Research* 28(15), 1984–1999.

**Box GEP and Cox DR** (1964). An Analysis of Transformations *Journal of the Royal Statistical Society, Series B (Methodological)* Vol. 26, No. 2 , pp. 211-252

**Bråte J, Krabberot AK, Dolven JK, Ose RF, Kristensen T, Bjorklund KR. et al**, (2012). Radiolaria associated with large diversity of marine alveolates. *Protist* 163, 767–777. doi: 10.1016/j.protis.2012.04.004

**Breton E, Brunet C, Sautour B, Brylinski JM**, (2000). Annual variations of phytoplanktonic biomass in the Eastern English Channel: comparison by pigment signatures and microscopic counts. *Journal of Plankton Research*. 22: 1423-1440.

**Breton E, Christaki U, Bonato S, Didry M and Artigas LF**, (2017). Phytoplankton trait variation and nitrogen use efficiency in temperate coastal phytoplankton. *Mar. Ecol. Prog. Ser.* 563, 35–49. doi: 10.3354/meps11974

**Breton E, Sautour B, Brylinski JM**, (1999). No feeding on *Phaeocystis sp.* As solitary cells (post- bloom) period by the copepod *Temora longicornis* in the coastal waters of the English Channel. *Hydrobiol.* 414: 13-23.

**Brunet C**, (1994). Analyse des pigments photosynthétiques par HPLC: communautés phytoplanktoniques et productivité primaire en Manche orientale. Thèse de doctorat, Université Paris VI, 364 pp.

**Brunet C, Brylinski JM, Bodineau L, Thoumelin G, Bentley D, Hilde D** (1996). Phytoplankton dynamics during the spring bloom in the South-Eastern English Channel. *Estuarine, Coastal and Shelf Science*. 43: 469–483.

**Brunet C, Brylinski JM, Frontier S**, (1992). Productivity, photosynthetic pigments and hydrology in the coastal front of the eastern English Channel. *Journal of Plankton Research*. 14: 1541-1552.

**Brylinski JM, Lagadeuc Y, Gentilhomme V, Dupont JP, Lafite R, Dupeuble PA, Huault MF, Auger Y, Puskaric E, Wartel M, Cabioch L**, (1991). Le "fleuve Côtier": un phénomène hydrologique important en Manche Orientale. Exemple du Pas de Calais. *O. Acta*. 11: 197-203.

**Brylinski JM, Bentley D, Quisthoudt C**, (1988). Discontinuité et zooplancton (Copépodes) en Manche orientale. *J. Plankton Res.* 10: 503-513.

**Brylinski JM, Brunet C, Bentley D, Thoumelin G, Hilde D** (1996). Hydrography and Phytoplankton biomass in the Eastern English Channel in spring 1992. *Est. Coast Shelf Sci.* 43: 507- 519.

**Brylinski JM, Dupont J et Bentley D**, (1984). Conditions hydrologiques au large du Cap Gris-Nez (France): premiers résultats. *Oceanologica Acta*, 7 (3): 315-322.

**Buck KR, Barry JP, Simpson AGB**, (2000). Monterey Bay cold seep biota: Euglenozoa with chemoautotrophic bacterial epibionts. *Eur. J. Protistol.* 36:117–26

- Buitenhuis E, Bleijswijk J, van Bakker D, Veldhuis M**, (1996). Trends in inorganic and organic carbon in a bloom of *Emiliana huxleyi* in the North Sea. *Marine Ecology Progress Series* 143: 271–282.
- Burki F**, (2014). The eukaryotic tree of life from a global phylogenomic perspective. *Cold Spring Harb. Perspect.* 6(5), 1-17.
- Burki F, Shalchian-Tabrizi K, Minge M, Skjaeveland Å, Nikolaev SI, Jakobsen KS, Pawlowski J**, (2007). Phylogenomics reshuffles the eukaryotic supergroups. *PLoS ONE* 8: 790–795.
- Cadée GC & Hegeman J**, (1986). Seasonal and annual variation in *Phaeocystis pouchetii* (Haptophyceae) in the westernmost inlet of the Wadden Sea during the 1973 to 1985 period. *Netherlands Journal of Sea Research*, 20(1), 29-36.
- Calbet A and Landry MR**, (2004). Phytoplankton growth, microzooplankton grazing, and carbon cycling in marine systems. *Limnology and Oceanography* 40:51–57.
- Campbell BJ, Yu L, Heidelberg JF and Kirchman DL**, (2011). **Activity of abundant and rare bacteria in a coastal ocean** *Proc. Natl. Acad. Sci. USA*, 108, pp. 12776–12781
- Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK & Knight R** (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature methods*, 7(5), 335-336.
- Caron DA, Countway PD, Jones AC, Kim DY, Schnetzer A**, (2012). Marine protistan diversity. *Annual review of marine science*, 4, 467-493.
- Caron DA**, (2000). Symbiosis and mixotrophy among pelagic microorganisms. In *Microbial Ecology of the Oceans*, ed. DL Kirchman, pp. 495–523. New York: Wiley-Liss
- Caron DA, Countway P**, (2009a). Defining DNA-based operational taxonomic units for microbialeukaryote ecology. *Applied and Environmental Microbiology*, 75, 5797–5808.
- Caron DA, Countway P**, (2009b). Hypotheses on the role of the protistan rare biosphere in a changing world. *Aquatic Microbial Ecology*, 57, 227–238.
- Caron DA, Countway PD, Jones AC, Kim DY, Schnetzer A**, (2012). Marine Protistan Diversity. *Annual Review of Marine Science*, 4, 467–493.
- Caron DA, Swanberg NR** (1990). The ecology of planktonic sarcodines. *Rev. Aquat. Sci.* 3:147–80
- Cauwet G, Gadel F, de Souza Sierra MM, Donard O et Ewald M**, (1990). Contribution of the Rhône River to organic carbon inputs to the northwestern Mediterranean Sea. *Continental Shelf Research* 10(9-11), 1025-1037.
- Cavalier-Smith T et Chao EE**, (1996). Molecular phylogeny of the free-living archezoan *Trepomonas agilis* and the nature of the first eukaryote. *J Mol Evol.* 43, 551-62.
- Cavalier-Smith T et Chao EE**, (2003). Phylogeny and classification of phylum Cercozoa (Protozoa). *Protist.* 154, 341-58.
- Cavalier-Smith T**, (2002). The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa. *Int J Syst Evol Microbiol.* 52, 297-354.

**Chantangsi C, Leander BS**, (2010). An SSU rDNA barcoding approach to the diversity of marine interstitial cercozoans, including descriptions of four novel genera and nine novel species. *Int J Syst Evol Microbiol* 60:1962–1977

**Charvet S, Vincent WF and Lovejoy C**, (2014). Effects of light and prey availability on Arctic freshwater protist communities examined by high-throughput DNA and RNA sequencing *FEMS Microbiol. Ecol.*, 88, pp. 550–564

**Cheung MK, Au CH, Chu KH, Kwan HS, Wong CK** (2010). Composition and genetic diversity of picoeukaryotes in subtropical coastal waters as revealed by 454 pyrosequencing. *The ISME Journal*, 4, 1053-1059

**Chiba S, Aita MN, Tadokoro K, Saino T, Sugisaki H, Nakata K**, (2008). From climate regime shifts to lower- trophic level phenology: Synthesis of recent progress in retrospective studies of the western North Pacific. *Progress In Oceanography*, 77, 112-126.

**Christaki U, Genitsaris S, Monchy S, Li LL, Rachik S, Breton E, Sime-Ngando T**, (2017). Parasitic eukaryotes in a meso-eutrophic coastal system with marked phytoplankton blooms *Front. Mar. Sci.* 4:416. doi: 10.3389/fmars.2017.00416.

**Christaki U, Kormas KA, Genitsaris S, Georges C, Sime-Ngando T, Viscogliosi E et al**, (2014). Winter-summer succession of unicellular eukaryotes in a meso-eutrophic coastal system. *Microb. Ecol.* 67, 13–23. doi: 10.1007/s00248-013-0290-4.

**Christaki U, Obernosterer I, Van Wambeke F, Veldhuis M., Garcia N & Catala P** (2008). Microbial food web structure in a naturally iron-fertilized area in the Southern Ocean (Kerguelen Plateau). *Deep Sea Research Part II: Topical Studies in Oceanography*, 55(5), 706-719.

**Christensen V and Pauly D**, (1995). *Fish production, catches and the carrying capacity of the world oceans*. Naga, the ICLARM Quarterly, 18(3), pp. 34-40.

**Clarke KR, Gorley RN** (2006). Primer v6: User Manual/Tutorial PRIMER-E, Plymouth

**Cloern JE**, (2001). Our evolving conceptual model of the coastal autrophication problem. *Marine Ecology Progress Series*. 210: 223-253.

**Cohu S, Thibaut T, Mangialajo L, Labat JP, Passafiume A, Blanfuné A, Simon N, Cottalorda JM, Lemée R**, (2011). Occurrence of the toxic dinoflagellate *Ostreopsis* cf. *ovata* in relation with environmental factors in Monaco (NW Mediterranean). *Mar. Pol. Bull.* 62, 2681-2691.

**Conley DJ**, (2000). Biogeochemical nutrient cycles and nutrient strategies. *Hydrobiologia*. 410: 87-96.

**Corliss JO**, (1991). Historically important events, discoveries, and works in Protozoology from de mid-17th to the mid-20th century.

**Cottonec G, Brunet, Sautour B, Thoumelin G** (2001). Nutritive value and selection of food particles by copepods during a spring bloom of *Phaeocystis* sp. in the English Channel, as determined by pigment and fatty acid analysis. *J. Plankton. Res.* 23: 693-703.

**Crawford DW, Lindholm T**, (1997). Some observations on vertical distribution and migration of the phototrophic ciliate *Mesodinium rubrum* (= *Myrionecta rubra*) in a stratified brackish inlet. *Aquat Microb Ecol* 13:267–274

**Cushing DH**, (1989). A difference in structure between ecosystems in strongly stratified waters and in those that are only weakly stratified. *J. Plankton Res.* 11:1-13

**Dawson SC, Hagen KD**, (2009). Mapping the protistan 'rare biosphere'. *J Biol* 8:105.

**Day Jr, Hall JW, CAS, Kemp WM, Yanez-Arancibia A**, (1989). *Estuarine ecology*, John Wiley and Sons (Eds), New York, 558 pp.

**Debroas D, Domaizon I, Humbert JF, Jardillier L, Lepère C, Oudart A, Taïb N (2017)**. Overview of freshwater microbial eukaryotes diversity: a first analysis of publicly available metabarcoding data. *FEMS Microbial Ecol* 93(4). Doi : 10.1093/femsc/fix023

**Diez 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. *Applied and Environmental Microbiology*, 67, 2932–2941.

**Dolan JR**, (2000). Tinitinnid ciliate diversity in the Mediterranean Sea: Longitudinal patterns related to water column structure in late spring-early summer. *Aquatic Microbial Ecology* 22: 69-78.

**Dolan JR, McKeon K**, (2005). The reliability of grazing rate estimates from dilution experiments: Have we over-estimated rates of organic carbon consumption by microzooplankton? *Ocean Science* 1:1-7.

**Dolan JR, Pérez MT**, (2000). Costs and benefits of mixotrophy: the case of marine oligotrichs. *Freshw Biol* 45:227–238

**Dolédec S, Chessel D** (1994). Co-inertia analysis: an alternative method for studying species-environment relationships. *Freshw Biol* 31: 277–294

**Dray S and Dufour AB** (2007). The ade4 package: implementing the duality diagram for ecologists. *Journal of Statistical Software* 22(4):1-20.

**Dray S, Chessel D, Thioulouse J** (2003). Co-inertia analysis and the linking of ecological data tables. *Ecology* 84: 3078–3089

**Ducklow HW, Steinberg DK, and Buesseler KO**, (2001). Upper ocean carbon export and the biological pump. *Oceanography* 14(4):50–58.

**Dugdale RC and Wilkerson FP**, (1988). Nutrient sources and primary production in the Eastern Mediterranean. *Oceanol. Acta*, 9, Spec. issue: 179-184.

**Dupont JP, Lafite R, Huault MF, Lamboy M, Brylinski JM, Guéguéniat P**, (1991). La dynamique des masses d'eau et des matières en suspension en Manche orientale. *Oceanol. Acta*. 11:177-186.

**Duris P, Gohau G**. (2011). Chap.1 L'ordre de la création. In: *Histoire des sciences de la vie*.

**Durrieu de Madron X, Abassi A, Heussner S, Monaco A, Aloisi JC, Radakovitch O, Giresse P, Buscail R et Kerherve P**, (2000). Particulate matter and organic carbon budgets for the Gulf of Lion (NW Mediterranean). *Oceanologica Acta* 23(6), 717–730.

**Durrieu de Madron X, Radakovitch O, Heussner S, Loye-Pilot MD, Monaco A**, (1999). Role of the climatological and current variability on the shelf-slope exchanges of particulate matter: Evidence from the Rhône continental margin (NW Mediterranean). *Deep Sea*



Research I 46, 1513–1538.

**Dyal PL, Hope S, Roberts DM & Embley TM**, (1995). Use of the PCR and fluorescent probes to recover SSU rRNA gene sequences from single cells of the ciliate protozoan *Spathidiutn*. *Molecular ecology*, 4(4), 499-504.

**Edgcomb V, Orsi W, Bunge J, Jeon S, Christen R, et al. (2011)**. Protistan microbial observatory in the Cariaco Basin, Caribbean. I. Pyrosequencing vs Sanger insights into species richness. *The ISME Journal* 5: 1344-1356.

**Embley TM et Martin W**, (2006). Eukaryotic evolution changes and challenges. *Nature*. 440, 623-30.

**Estournel C, Durrieu de Madron X, Marsaleix P, Auclair F, Julliand C, Vehil R**, (2003). Observation and modeling of winter coastal oceanic circulation in the Gulf of Lion under wind conditions influenced by the continental orography (FETCH experiment). *Journal of Geophysical Research* 108 (C3), 8059

**Estournel C, Broche P, Marsaleix P, Devenon J, Auclair F and Vehil R**, (2001). The Rhone river plume in unsteady conditions: Numerical and experimental results. *Estuarine Coastal and Shelf Science*, 53(1): 25–38.

**Falkowski PG, Barber RT, Smetacek V** (1998). Biogeochemical Controls and Feedbacks on Ocean Primary Production. *Science*. 281 (5374), 200-206.

**Fehling J, Stoecker DK, Baldauf SL**, (2007). Photosynthesis and the eukaryote tree of life. In: Falkowski PG, Knoll AH eds. *Evolution of primary producers in the sea*. New York: Academic Press. 75–107

**Ferrera I, Giner CR, Reñé A, Camp J, Massana R, Gasol J.M. and Garcés E**, (2016). Evaluation of Alternative High-Throughput Sequencing Methodologies for the Monitoring of Marine Picoplanktonic Biodiversity Based on rRNA Gene Amplicons *Front. Mar. Sci*, doi: 10.3389/fmars.00147

**Field CB, Behrenfeld MJ, Randerson JT, Falkowski P**, (1998). Primary production of the biosphere: integrating terrestrial and oceanic components. *Science* 281:237–40

**Flexas M, Durrieu de Madron X, Garcia M, Canals M, Arnau P**, (2002). Flow variability in the Gulf of Lions during the MATER HFF experiment (March-May 1997). *Journal of Marine Systems* 33–34, 197–214.

**Fokin S**, (2013) Otto Bütschli (1848–1920). Where we will genuflect? *Protistology*.

**Fuda J, Etiopie G, Millot C, Favali P, Calcara M, Smriglio G and Boschi E** (2002). Warming, salting and origin of the Tyrrhenian Deep Water. *Geophysical Research Letters*, 29(19).

**Fukuda HR, Sohrin T, Nagata and I Koike** (2007). Size distribution and biomass of nanoflagellates in meso- and bathypelagic layers of the subarctic Pacific. *Aquat. Microb.Ecol.*46:203–207

**Genitsaris S, Monchy S, Breton E, Lecuyer E, Christaki U**, (2016). Small-scale variability of protistan planktonic communities relative to environmental pressures and biotic interactions at two adjacent coastal stations *Mar. Ecol. Prog. Ser.*, 548, pp. 61-75, doi: 10.3354/meps11647

- Genitsaris S, Monchy S, Viscogliosi E, Sime-Ngando T, Ferreira S, Christaki U**, (2015). Seasonal variations of marine protist community structure based on taxon-specific traits using the eastern English Channel as a model coastal system *FEMS Microbiol. Ecol.*, 91.
- Gentilhomme V, Lizon F**, (1998). Seasonal cycle of nitrogen and phytoplankton biomass in a well-mixed coastal system (Eastern English Channel). *Hydrobiologia*. 361: 191-199.
- Gentile G, Giuliano L, D'Auria G, Smedile F, Azzaro M et al** (2006). Study of bacterial communities in Antarctic coastal waters by a combination of 16S rRNA and 16S rDNA sequencing *Environ. Microbiol.*, 8, pp. 2150–2161.
- Georges C, Monchy S, Genitsaris S and Christaki U** (2014). Protist community composition during early phytoplankton blooms in the naturally iron-fertilized Kerguelen area (Southern Ocean). *Biogeosciences* 11, 5847–5863. doi: 10.5194/bg-11-5847-2014
- Gest H**, (2004b). The discovery of microorganisms by Robert Hooke and Antoni van Leeuwenhoek, Fellows of The Royal Society. *Notes and Records of the Royal Society*, 58, 187–201.
- Gilbert JA, Field D, Swift P, Newbold L, Oliver A, Smyth T, Somerfield PJ, Huse S, Joint I** (2009). The seasonal structure of microbial communities in the Western English Channel, *Environmental Microbiology*, 11(12)
- Gilbert JA, Steele JA, Caporaso JG, Steinbrück L, Reeder J, Temperton B, et al** (2012). Defining seasonal marine microbial community dynamics. *ISMEJ*. 6, 298–308. doi: 10.1038/ismej.2011.107
- Gilles A, Meglécz E, Pech N et al**, (2011). Accuracy and quality assessment of 454 GS-FLX Titanium pyrosequencing. *BMC genomics*, 12, 245.
- Goffart A, Hecq JH, Legendre L**, (2002). Changes in the development of the winter-spring phytoplankton bloom in the Bay of Calvi (NW Mediterranean) over the last two decades: A response to changing climate? *Mar Ecol Prog Ser* 23, 45–60
- Gómez F, Gorsky G**, (2003). Annual microplankton cycles in Villefranche Bay, Ligurian Sea, NW Mediterranean, *Journal of Plankton Research*, 25, 323-339
- Gómez F & Souissi S** (2007). Unusual diatoms linked to climatic events in the northeastern English Channel. *Journal of Sea Research*, 58(4), 283-290.
- Gómez F**, (2005). A list of dinoXagellates in the world oceans. *Acta Bot Croatica* 84:129–212
- González HE, Giesecke R, Vargas A, Pavez M, Iriarte J, Santibáñez, P, Castro L, Escribano R, Pagès F**, (2004). Carbon cycling through the pelagic foodweb in the northern Humboldt current off Chile (23°S). *ICES J. Mar. Sci.* 61: 572-584.
- Gran-Stadniczenko S, Supraha L, Egge ED, Edvardsen B**, (2016). Haptophyte Diversity and Vertical Distribution Explored by 18S and 28S Ribosomal RNA Gene Metabarcoding and Scanning Electron Microscopy. *Journal of Eukaryotic Microbiology*, 64(4). doi.org/10.1111/jeu.12388
- Grattepanche JD, Breton E, Brylinski JM et al**, (2011a). Succession of primary producers and micrograzers in a coastal ecosystem dominated by *Phaeocystis globosa* blooms *J. Plankton Res.*, 33, pp. 37–50.

**Grattepanche JD, Vincent D, Breton E & Christaki U** (2011). Microzooplankton herbivory during the diatom *Phaeocystis* spring succession in the eastern English Channel. *Journal of Experimental Marine Biology and Ecology*, 404(1), 87-97.

**Grattepanche JD, Vincent D, Breton E et al**, (2011b). Phytoplankton growth and microzooplankton grazing during a spring bloom in the eastern English Channel *J. Exp. Mar. Biol. Ecol.*, 404, pp. 87–97

**Guarini S, Schiöth HB, Mioni C, Cainazzo MM, Ferrazza G, Giuliani D, et al**, (2002). MC<sub>3</sub> receptors are involved in the protective effect of melanocortins in myocardial ischaemia/reperfusion-induced arrhythmias. *Naunyn-Schmiedeberg's Arch Pharmacol* 366:177–182.

**Guéguéniat P, Dubois PB, Salomon JC, Masson M, Cabioch L** (1995). FLUXMANCHE radiotracer measurements : a contribution to the dynamics of the English Channel and North Sea. *J. Mar. Sys.* 6: 483-494.

**Guillou L, Bachar D, Audic S et al**, (2013). The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy *Nucleic Acid Res.*, 41, pp. 597–604

**Guillou L, Eikrem W, Chretiennot-Dinet MJ, Le Gall F, Massana R**, (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*

**Guillou L, Viprey M, Chambouvet A et al**, (2008). Widespread occurrence and genetic diversity of marine parasitoids belonging to Syndiniales (Alveolata). *Environmental microbiology*, 10, 3349– 3365.

**Guiselin N**, (2010). Etude de la dynamique des communautés phytoplanctoniques par microscopie et cytométrie en flux, en eaux côtières de la Manche orientale. Thèse de doctorat, Université du littoral côte d'opale: 190 pp.

**Habura A, Pawlowski J, Hanes SD, et al**, (2004b). Unexpected foraminiferal diversity revealed by small-subunit rDNA analysis of Antarctic sediment. *J Eukaryot Microbiol.* 51, 173-9.

**Hackett JD, Yoon HS, Li S, et al**, (2007). Phylogenomic analysis supports the monophyly of cryptophytes and haptophytes and the association of rhizaria with chromalveolates. *Mol Biol Evol.* 24, 1702-13.

**Hackett JD, Scheetz TE, Yoon HS et al**, (2005). Insights into a dinoflagellate genome through expressed sequence tag analysis. *BMC Genomics.* 6, 80.

**Hammer O, Harper DAT, Ryan PD**, (2001). PAST: paleontological statistics software package for education and data analysis *Paleontol. Electron*, pp. 4-9

**Hansen FC**, (1995). Trophic interactions between zooplankton and *Phaeocystis* cf. *globosa*. *Helgol. Meeresunters.* 49: 283–293.

**Hansen PJ & Calado AJ** (1999). Phagotrophic Mechanisms and Prey Selection in Free-living Dinoflagellates1. *J Euk Microbiol* 46, 382-389.

**Hansen PJ**, (1991). Quantitative importance and trophic role of heterotrophic dinoflagellates in a coastal pelagic food web. *Mar Ecol Prog Ser* 73:253–261

- Harada A, Ohtsuka S and Horiguchi T**, (2007). Species of the parasitic genus *Duboscquella* are members of the enigmatic marine alveolate group, I. *Protist* 158, 337–347. doi: 10.1016/j.protis. .03.005
- Hernandez-Farinas T, Soudant D, Barille L, Belin C, Lefebvre A, Bacher C** (2014). Temporal changes in the phytoplankton community along the French coast of the eastern English Channel and the southern Bight of the North Sea, *Ices Journal Of Marine Science*, 71 (4), 821-833
- Hoch T, Garreau P**, (1998). Phytoplankton dynamics in the English Channel: a simplified three dimensional approach. *Journal of Marine Systems*. 16: 133-150.
- Hooke R**, (1665). Obferv. XX. Of blue Mould, and of the firft Principles of Vegetation arifing from Putrefaction. In: *Micrographia* , pp. 125–131.
- Hooff RC and Peterson WT**, (2006). Copepod biodiversity as an indicator of changes in ocean and climate conditions of the northern California current ecosystem. *Limnol. Oceanogr.* 51, 2607–2620. Doi: 10.4319/lo.2006.51.6.2607
- Hoppenrath, M, Elbrächter M & Drebes G**, (2009). *Marine phytoplankton*. Stuttgart: Schweitzerbart'sche Verlagsbuchhandlung.
- Horner RA, Postel JR, Halsband-Lenk C, Pierson JJ, Pohnert G, Wichard T**, (2005) Winter-spring phytoplankton blooms in Dabob Bay, Washington. *Prog Oceanogr* 67:286–313
- Houliez E, Lizon F, Artigas LF, Lefebvre S, Schmitt FG**, (2013). Spatio-temporal variability of phytoplankton photosynthetic activity in a macrotidal ecosystem (the Strait of Dover, eastern English Channel). *Estuarine, Coastal and Shelf Science*, 129, 37-48.
- Houliez E, Lizon F, Thyssen M, Artigas LF, Schmitt FG** (2012). Spectral fluorometric characterization of Haptophyte dynamics using the FluoroProbe: an application in the eastern English Channel for monitoring *Phaeocystis globosa*. *Journal of Plankton Research*, 34, 136–151.
- Hu SK, Campbell V, Connell P, Gellene AG, Liu Z, Terrado R and Caron D**, (2016). Protistan diversity and activity inferred from RNA and DNA at a coastal site in the eastern North Pacific FEMS *Microbiol. Ecol.* 92: fiw050.
- Hugoni M, Taib N, Debroas D, Domaizon I, Dufournel IJ, Bronner G, Salter I, Agogué H, Mary I and Galand PE**, (2013). Structure of the rare archaeal biosphere and seasonal dynamics of active ecotypes in surface coastal waters *PNAS.*, 110 (15), pp. 6004-6009.
- James TY, Kauff F, Schoch CL, et al**, (2006). Reconstructing the early evolution of Fungi using a six-gene phylogeny. *Nature*. 443, 818-22.
- Jeong HJ**, (1999). The Ecological Roles of Heterotrophic Dinoflagellates in Marine Planktonic Community1. *J Euk Microbiol* 46, 390-396.
- Jing H, Rocke E, Kong L, Xia X, Liu H and Landry MR** (2015). Protist communities in a marine oxygen minimum zone off Costa Rica by 454 pyrosequencing *Biogeosciences Discuss.*, 12, 13483-13509
- Jones MDM, Forn I, Gadelha C et al**, (2011). Discovery of novel intermediate forms redefines the fungal tree of life. *Nature*, 474, 200–3.

- Jordan R, Chamberlain A**, (1997). Biodiversity among haptophyte algae. *Biodiversity & Conservation*, 6, 131–152.
- Jardillier L, Zubkov MV, Pearman J, & Scanlan DJ** (2010). Significant CO<sub>2</sub> fixation by small prymnesiophytes in the subtropical and tropical northeast Atlantic Ocean. *The ISME journal*, 4(9), 1180-1192.
- Jürgens K, Massana R & Kirchman, DL**, (2008). Protist grazing on marine bacterioplankton. *Microbial ecology of the oceans*, 2nd edn. Wiley-Liss, New York, 383-442.
- Karl DM, Bidigare RR, Letelier RM**, (2001). Long-term changes in plankton community structure and productivity in the North Pacific Subtropical Gyre: The domain shift hypothesis. *Deep Sea Research Part II: Topical Studies in Oceanography*, 48, 1449-1470.
- Karpov SA, Mamkaeva MA, Aleoshin VV et al**, (2014). Morphology, phylogeny, and ecology of the aphelids (Aphelidea, Opisthokonta) and proposal for the new superphylum Opisthosporidia. *Frontiers in Microbiology*, 5.
- Kenkel NC** (2006). On selecting an appropriate multivariate analysis. *Can J Plant Sci* 86: 663–676
- Kolodziej K, Stoeck T**, (2007). Cellular identification of a novel uncultured marine stramenopile (MAST-12 Clade) small-subunit rRNA gene sequence from a norwegian estuary by use of fluorescence in situ hybridization-scanning electron microscopy. *Applied and environmental microbiology*, 73, 2718–26.
- Kuhn M. Contributions from Wing J, Weston S, Williams A, Keefer C, Engelhardt A, Cooper T, Mayer Z, Kenkel B, the R Core Team, Benesty M, Lescarbeau R, Ziem A, Scrucca I, Tang Y and Candan C**, (2016). **Caret: Classification and Regression Training** R package version 6.0-68
- Kunin V, Engelbrektson A, Ochman H, Hugenholtz P**, (2010). Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. *Environmental microbiology*, 12, 118–23.
- Lafite R, Shimwell S, Grochowski N, Dupont JP, Nash L, Salomon JC, Cabioch L, Collins M, Gao S**, (2000). Suspended particulate matter fluxes through the Straits of Dover, English Channel: observations and modelling. *Oceanol. Acta*, 23: 687-699.
- Lamy A, Millot C, Molines JM**, (1981). Bottom pressure and sea level measurements in the Gulf of Lions. *Journal of Physical Oceanography* 11, 394–410.
- Lamy D, Obernosterer I, Laghdass M, Artigas F, Breton E, Grattepanche J, Lecuyer E, Degros N, Lebaron P, Christaki U**, (2009). Temporal changes of major bacterial groups and bacterial heterotrophic activity during a *Phaeocystis globosa* bloom in the eastern English Channel. *Aquat. Microb. Ecol.*, 58, 95–107.
- Landry MR AND OTHERS**, (1997). Iron and grazing constraints on primary production in the central equatorial Pacific: An EqPac synthesis. *Limnol. Oceanogr.* 42: 405–418.
- Langlais C**, (2007). Etude de la variabilité interannuelle des échanges côte-large: Simulation haute résolution de la dynamique du Golfe du Lion. Thèse de Doctorat, Université du Sud Toulon Var.
- Lara E, Chatzinotas A et Simpson AG**, (2006). Andalucia (n. gen.)-the deepest branch within jakobids (Jakobida; Excavata), based on morphological and molecular study of a new flagellate from soil. *J Eukaryot Microbiol.* 53, 112-20.

- Lara E, Moreira D et Lopez-Garcia P**, (2009). The Environmental Clade LKM11 and Rozella Form the Deepest Branching Clade of Fungi. *Protist*.
- Leakey RJG, Burkill PH & Sleight MA** (1994). A comparison of fixatives for the estimation of abundance and biovolume of marine planktonic ciliate populations. *Journal of plankton Research*, 16(4), 375-389.
- Lecointre G, Le Guyader H**, (2006). *Classification phylogénétique du vivant*. 3ème édition. Paris, France.
- Lee J, Godon C, Lagniel G, et al**, (1999). Yap1 and Skn7 control two specialized oxidative stress response regulons in yeast. *J Biol Chem*. 274, 16040-6.
- Lefebvre A, Guiselin N, Barbet F, and Artigas FL** (2011). Long-term hydrological and phytoplankton monitoring (1992–2007) of three potentially eutrophic systems in the eastern English Channel and the Southern Bight of the North Sea. *ICES Journal of Marine Science*, 68, 2029–2043.
- Legendre L, Rassoulzadegan F**, (1995). Plankton and nutrient dynamics in marine waters. *Ophelia* 41:153-172
- Legendre P, Legendre L** (1998). *Numerical ecology*, 2nd edn. Elsevier, Amsterdam
- Leguerrier D**, (2005). Construction et étude d'un modèle de réseau trophique de la vasière de Brouage (Bassin de Marennes-Oléron, France). Prise en compte de la saisonnalité et des échanges physiques pour la synthèse constructive des connaissances sur une zone intertidale d'une région tempérée. Université de La Rochelle La Rochelle
- Licandro P and Ibanez F**, (2000). Changes of zooplankton communities in the Gulf of Tigullio (Ligurian sea, western Mediterranean) from 1985 to 1995. Influence of hydroclimatic factors. *Journal of Plankton research*, 22(12), 2225-22253.
- Liu H, Probert I, Uitz J et al**, (2009). Extreme diversity in noncalcifying haptophytes explains a major pigment paradox in open oceans. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 12803–12808.
- Lizon F, Seuront L, Lagadeuc Y** (1998). Photoadaptation and primary production study in a tidally- mixed coastal waters using a lagrangian model. *Mar. Ecol. Prog. Ser.* 169 : 43-54.
- Logares R, Audic S, Bass D, Bittner L, Boutte C, Christen R, Claverie JM, Decelle J, Dolan JR, Dunthorn M, Edvardsen B, Gobet A, Kooistra WH, Mahe F, Not F, Ogata H, Pawlowski J, Pernice MC, Romac S, Shalchian-Tabrizi K, Simon N, Stoeck T, Santini S, Siano R, Wincker P, Zingone A, Richards TA, de Vargas C and Massana R**, (2014). Patterns of rare and abundant marine microbial eukaryotes *Curr. Biol.*, 24, pp. 813–821.
- Lopez Garcia P, Rodríguez-Valera F, Pedrós-Alió C, Moreira D**, (2001). Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. *Nature*, 409, 603–7.
- López-García P et Moreira D**, (2008). Tracking microbial biodiversity through molecular and genomic ecology. *Research in Microbiology*. 159, 67-73.
- López-García P, Philippe H, Gail F, Moreira D**, (2003). Autochthonous eukaryotic diversity in hydrothermal sediment and experimental microcolonizers at the Mid- Atlantic Ridge *Proc. Natl. Acad. Sci. USA*, 100, pp. 697–702

- Ludwig W, Meybeck M, Abousamra F** (2003). Riverine transport of water, sediments, and pollutants to the Mediterranean Sea. UNEP MAP Technical report Series 141, UNEP/MAP Athens, 111 pp
- Ludwig W, E Dumont, M Meybeck, and S Heussner**, (2009). River discharges of water and nutrients to the Mediterranean Sea: Major drivers for ecosystem changes during past and future decades? *Prog. Oceanogr.*, 80, 199–217, doi:10.1016/j.pocean.2009.02.001.
- Mackas L, Thomson RE, Galbraith M**, (2001). Hanges in the zooplankton community of the British Columbia continental margin, 1985-1999, and their covariation with oceanographic conditions. *Canadian Journal of Fisheries and Aquatic Sciences*, 58, 685-702.
- Maclean RC, Richardson DJ, LePardo R et al**, (2004). The identification of *Naegleria fowleri* from water and soil samples by nested PCR. *Parasitol Res.* 93, 211-7.
- Mann DG** (1999). The species concept in diatoms. *Phycologia*, 38(6), 437-495.
- Mantoura RFC, Martin J-M, Wollast R**, (1991). Ocean margin processes in global change. WileyInterscience, New York: 469 pp.
- Margulies M, Egholm M, Altman WE et al**, (2005). Genome sequencing in microfabricated highdensity picolitre reactors. *Nature*, 437, 376–80.
- Marquis E, Niquil N, Delmas D, Hartmann HJ, Bonnet D, Carlotti F, Herbland A, Labry C, Sautour B, Laborde P, Vézina A, Dupuy C**, (2007) Inverse analysis of the planktonic food web dynamics related to phytoplankton bloom development on the continental shelf of the Bay of Biscay, French coast. *Estuarine, Coastal and Shelf Science* 73, 223-235
- Marty JC, Chiavérini J, Pizay MD, Avril B** (2002). Seasonal and interannual dynamics of nutrients and phytoplankton pigments in the western Mediterranean Sea at the DYFAMED times-series station (1991-1999). *Deep Sea Research Part II: Topical Studies in Oceanography*, 49, 1965-1985
- Massana R**, (2011). Eukaryotic picoplankton in surface oceans. *Annu Rev Microbiol* 65: 1–110. doi: 10.1146/annurev-micro-090110-102903
- Massana R, Castresana J, Balague V, Guillou L, Romari K**, (2004). Phylogenetic and ecological analysis of novel marine stramenopiles. *Appl. Environ. Microbiol.* 70:3528–34
- Massana R, Del Campo J, Sieracki ME, Audic S, Logares R**, (2013) Exploring the uncultured microeukaryote majority in the oceans: reevaluation of ribogroups within stramenopiles. *The ISME journal*
- Massana R, Gobet A, Audic S, Bass D, Bittner L, Boutte C, Chambouvet A, Christen R, Claverie J, Decelle J, Dolan JR, Dunthorn M, Edvardsen B, Forn I, Forster D, Guillou L, Jaillon O, Kooistra W H C F, Logares R, Mahé F, Not F, Ogata H, Pawlowski J, Pernice MC, Probert I, Romac S, Richards T, Santini S, Shalchian-Tabrizi K, Siano R, Simon N, Stoeck T, Vaultot D, Zingone A & de Vargas C**, (2015). Marine protist diversity in European coastal waters and sediments as revealed by high-throughput sequencing *Environ. Microbiol.*, 17, pp. 4035–4049
- Massana R, Pedrós-Alió C**, (2008) Unveiling new microbial eukaryotes in the surface ocean. *Current opinion in microbiology*, 11, 213–218.

- Mayer Z, Kenkel B, the R Core Team, Benesty M, Lescarbeau R, Ziem A, Scrucca I, Tang Y and Candan C** (2016) caret: Classification and Regression Training R package version 6.0-68
- McLusky DS**, (1989). The estuarine ecosystem, 2nd ed. Blackie, Glasgow. 215 pp.
- McManus GB, Griffin PM, Pennock, JR**, (2004). Bacterioplankton abundance and growth in a river-dominated estuary: relationships with temperature and resources. *Aquat. Microb. Ecol.* 37: 23- 32.
- McMinn A, Scot FJ**, (2005) 3. DinoXagellates. In: Scott FJ, Marchant HW (eds) Antarctic marine protists. *Austral Biol Res Study*, Canberra, pp 202–250
- Medinger R, Nolte V, Pandey RV, Jost S, Ottenwalder B, Schlotterer C and Boenigk J**, (2010). Diversity in a hidden world: potential and limitation of next-generation sequencing for surveys of molecular diversity of eukaryotic microorganisms, *Mol. Ecol.*, 19, 32–40.
- Middelboe, M, Kroer N, Jørgensen, NOG, Pakulski, D**, (1998). Influence of sediment on pelagic carbon and nitrogen turnover in a shallow Danish Estuary. *Aquat. Microb. Ecol.* 14: 81-90.
- Millot C**, (2013). Levantine Intermediate Water characteristics: an astounding general misunderstanding! *Scientia Marina*, 77(2): 217–232.
- Millot C**, (1990). The Gulf Of Lions Hydrodynamics. *Continental Shelf Research* 10(9-11), 885-894.
- Millot C**, (1999). Circulation in the Western Mediterranean Sea. *Journal of Marine Systems*, 20(1-4): 423–442.
- Mills HJ, Martinez RJ, Story S, et al**, (2005). Characterization of microbial community structure in Gulf of Mexico gas hydrates: comparative analysis of DNA- and RNA-derived clone libraries. *Appl Environ Microbiol.* 71, 3235-3247.
- Mitchell A, Romano GH, Groisman B, Yona A, Dekel E, Kupiec M, Dahan O and Pilpel Y**, (2009). Adaptive prediction of environmental changes by microorganisms. *Nature*, 460 pp: 220–224.
- Modigh M**, (2001). Seasonal variations of photosynthetic ciliates at a Mediterranean coastal site. *Aquat Microb Ecol* 23:163–175
- Moeseneder MM, Arrieta JM and Herndl GJ**, (2005). A comparison of DNA- and RNA-based clone libraries from the same marine bacterioplankton community *FEMS Microbiol. Ecol.*, 51, pp. 341–352
- Molinero JC, Ibanez F, Nival P, Buecher E, Souissi S** (2005). North Atlantic climate and northwestern Mediterranean plankton variability. *Limnol Oceanogr*, 50(4), 1213-1220
- Monaco A, Courp T, Heussner S, Carbonne J, Fowler SW et Deniaux B**, (1990). Seasonality and composition of particulate fluxes during ECOMARGE I, western Gulf of Lions. *Continental Shelf Research* 10(9-11), 959-987.
- Monchy S, Grattepanche JD, Breton E, Meloni D, Sancier G, Chabé M, Delhaes L, Viscogliosi E, Sime-Ngando T, Christaki U**, (2012). Microplanktonic community structure in a coastal system relative to a Phaeocystis bloom inferred from morphological and tag pyrosequencing methods. *PLoS One* 7:e39924.



- Moon-van der Staay SY, Wachter R De, Vaultot D**, (2001). Oceanic 18S rDNA sequences from picoplankton reveal unsuspected eukaryotic diversity. *Nature*, 409, 607–610.
- Moran et al**, (2010). Increasing importance of small phytoplankton in a warmer ocean. *Global Change Biology* (2010) 16, 1137–1144, doi: 10.1111/j.1365 2486.2009.01960.x
- Moreira D & López-García P**, (2002). The molecular ecology of microbial eukaryotes unveils a hidden world. *Trends in microbiology*, 10(1), 31-38.
- Moreira D, von der Heyden S, Bass D et al**, (2007). Global eukaryote phylogeny: Combined small- and large-subunit ribosomal DNA trees support monophyly of Rhizaria, Retaria and Excavata. *Molecular phylogenetics and evolution*, 44, 255–66.
- Mullin JB, Riley JP** (1955). The colorimetric determination of silicate with special reference to sea and natural waters. *Analytica Chimica Acta*. 12: 162-176.
- Murphy J, Riley JP** (1962). A modified single solution method for the determination of phosphate in natural waters. *Analytica Chimica Acta*. 12 : 31-36.
- Nelson DM, Tréguer P, Brzezinski MA, Leynaert A & Quéguiner B**, (1995). Production and dissolution of biogenic silica in the ocean: revised global estimates, comparison with regional data and relationship to biogenic sedimentation. *Global Biogeochemical Cycles*, 9(3), 359-372.
- Nittis K and Lascaratos A** (1998). Diagnostic and prognostic numerical studies of LIW formation. *Journal of Marine Systems*, 18(1-3): 179–195.
- Not F, Gausling R, Azam F, Heidelberg JF, & Worden, AZ** (2007). Vertical distribution of picoeukaryotic diversity in the Sargasso Sea. *Environmental Microbiology*, 9(5), 1233-1252.
- Not F, del Campo J, Balagué V et al**, (2009). New insights into the diversity of marine picoeukaryotes *PloS One*, 4 (e7143).
- Nowack EC, Melkonian M et Glockner G**, (2008). Chromatophore genome sequence of *Paulinella* sheds light on acquisition of photosynthesis by eukaryotes. *Curr Biol*. 18, 410-8.
- Orsi W, Biddle JF and Edgcomb V**, (2013). Deep sequencing of seafloor eukaryotic rRNA reveals active fungi across marine subsurface provinces *PLoS ONE*, 8: e56335.
- Ovchinnicov I and Plakhin E** (1984). Formation of Mediterranean Intermediate Water in the Rhodes Cyclonic Gyre. *Okeanologiya*, 24(3): 417–420.
- Parfrey LW, Barbero E, Lasser E, et al**, (2006). Evaluating support for the current classification of eukaryotic diversity. *PLoS Genet*. 2, e220.
- Patron NJ, Inagaki Y et Keeling PJ**, (2007). Multiple gene phylogenies support the monophyly of cryptomonad and haptophyte host lineages. *Curr Biol*. 17, 887-91.
- Pawlowski J et Burki F**, (2009). Untangling the phylogeny of amoeboid protists. *J Eukaryot Microbiol*. 56, 16-25.
- Penna A, Casabianca S, Guerra AF, Vernesi C, Scardi M**, (2017). Analysis of phytoplankton assemblage structure in the Mediterranean Sea based on high-throughput sequencing of partial 18S rRNA sequences. *Marine Genomics* 36, 49–55.

- Peperzak L, Colijn F, Gieskes WWC, Peeters JCH** (1998). Development of the diatom-Phaeocystis spring bloom in the Dutch coastal zone of the North Sea: the silicon depletion versus **the daily** irradiance threshold hypothesis. *Journal of Plankton Research*, 20, 517-537
- Pereira Quiroga AI**, (2005). Evolution des populations phytoplanctoniques de la Baie de Banyuls-sur-Mer, avec une attention particulière pour les genres possédant des espèces potentiellement toxiques. *Thèse de doctorat en Océanologie biologique*, Université Pierre et Marie Curie, Paris, 186 pp.
- Pernetta JC and JD Milliman (eds.)** (1995). Land-Ocean Interactions in the Coastal Zone Implementation Plan. IGBP Report No. 33, Stockholm Sweden, 215 p.
- Petrenko A** (2003). Variability of circulation features in the Gulf of Lions NW Mediterranean Sea. *Oceanologica Acta* 26 (5), 323–338.
- Petz W** (2005). Ciliates. In: Scott, F.J., Marchant, H.J. (Eds.), *Antarctic Marine Protists*. ABRIS and AAD Publishers, pp. 347–448.
- Philippe H et Germot A**, (2000). Phylogeny of eukaryotes based on ribosomal RNA: longbranch attraction and models of sequence evolution. *Mol Biol Evol.* 17, 830-4.
- Philippe H, Germot A, Moreira D**, (2000). The new phylogeny of eukaryotes. *Current Opinion in Genetics & Development*, 10, 596–601.
- Pitta P, Giannakouru A, Christaki U** (2001). Planktonic ciliates in the oligotrophic Mediterranean Sea : Longitudinal trends of standing stocks, distributions and analysis of food vacuole contents, *Aquat. Microb. Ecol.*, 24, 297-311
- Piwosz K, Wiktor JM, Niemi A, Tatarek A, Michel C**, (2013). Mesoscale distribution and functional diversity of picoeukaryotes in the first-year sea ice of the Canadian Arctic. *The ISME journal*, 7, 1461–71.
- Polovina et al**, (2008). Ocean's least productive waters are expanding. *Geophysical Research Letters*. Volume 35, Issue 3 DOI: 10.1029/2007GL031745
- Pomeroy JW, Gray DM, Brown T, Hedstrom NR, Quinton WL, Granger RJ & Carey SK**, (2007). The cold regions hydrological model: a platform for basing process representation and model structure on physical evidence. *Hydrological Processes*, 21(19), 2650-2667.
- Pomeroy LR** (1974). The Ocean's Food Web, A Changing Paradigm. *BioSci* 24, 499-504.
- Prescott DM**, (2000). Genome gymnastics: unique modes of DNA evolution and processing in ciliates. *Nat Rev Genet.* 1, 191-8.
- Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO** (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* Jan;41(Database issue):D590-6.
- Rabouille C, Conley DJ, Dai MH, Cai WJ, Chen CTA, Lansard B, Green R, Yin K, Harrison PJ, Dagg M et McKee B**, (2008). Comparison of hypoxia among four river-dominated ocean margins: The Changjiang (Yangtze), Mississippi, Pearl, and Rhone rivers. *Continental Shelf Research* 28(12), 1527-1537.
- Raghukumar S**, (2002) Ecology of the marine protists, the Labyrinthulomycetes (Thraustochytrids and Labyrinthuloids). *Europ J Protistol* 38: 127—145

- Rassoulzadegan F**, (1993). Protozoan patterns in the Azam-Ammerman's bacteria-phytoplankton mutualism. edn., Barcelona: Spanish Society for Microbiology
- Rédou V, Ciobanu M C, Pachiadaki MG, Edgcomb V, Alain K, Barbier G, et al.** (2014). In-depth analyses of deep subsurface sediments using 454-pyrosequencing reveals a reservoir of buried fungal communities at record-breaking depths. *FEMS Microbiol. Ecol.* 90, 908–921. doi: 10.1111/1574-6941.12447
- Reyes-Prieto A, Weber AP et Bhattacharya D**, (2007). The origin and establishment of the plastid in algae and plants. *Annu Rev Genet.* 41, 147-68.
- Rixen, M, Beckers J, Levitus S, Antonov J, Boyer T, Maillard C, Fichaut M, Balopoulos E, Iona S, Dooley H, Garcia M, Manca B, Giorgetti A, Manzella G, Mikhailov N, Pinardi N and Zavatarelli M**, (2005). The Western Mediterranean Deep Water : A proxy for climate change. *Geophysical Research Letters*, 32(12).
- Rodriguez-Ezpeleta N, Brinkmann H, Burey SC, et al**, (2005). Monophyly of primary photosynthetic eukaryotes: green plants, red algae, and glaucophytes. *Curr Biol.* 15, 1325-30.
- Salomon JC, Breton M** (1993). An atlas of long-term currents in the Channel. *O. Acta.* 16: 439-448.
- Samari C, Millot C et Prieur L**, (1995). Aspects of the seasonal and mesoscale variabilities of the Northern Current in the western Mediterranean Sea inferred from the PROIG-2 and PROS-6 experiments. *Deep-Sea Research* 42(6), 893–917
- Salter I, Galand PE, Fagervold SK, Lebaron P and others** (2015) Seasonal dynamics of active SAR11 ecotypes in the oligotrophic Northwest Mediterranean Sea. *ISME J* 9:347–360
- Sanders RW** (1995). Seasonal distributions of the photosynthetic ciliates *Laboea strobila* and *Myrionecta rubra* (= *Mesodinium rubrum*) in an estuary of the Gulf of Maine. *Aquat Microb Ecol* 9:237–242
- Sanders RW, Berninger UG, Lim EL, Kemp PF & Caron DA**, (2000). Heterotrophic and mixotrophic nanoplankton predation on picoplankton in the Sargasso Sea and on Georges Bank. *Marine Ecology Progress Series*, 192(1), 103-118.
- Sarthou G, Timmermans KR, Blain S, Tréguer P**, (2005). Growth physiology and fate of diatoms in the ocean: a review. *J.Sea.Res.* 53 (1-2), 25-42.
- Schapira M, Vincent D, Seuront L, Gentilhomme V**, (2008). Temporal patterns of phytoplankton assemblages, size spectra and diversity during the wane of a *Phaeocystis globosa* spring bloom in hydrologically contrasted coastal waters. *Journal of Marine Biological Association of United Kingdom.* 88, 649-662.
- Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Weber CF** (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and environmental microbiology*, 75(23), 7537-7541.
- Schoeman V, Becquevort S, Stefels J, Rousseau V, Lancelot C**, (2005). *Phaeocystis* blooms in the global ocean and their controlling mechanisms: a review. *J. Sea Res.* 53: 43-66.
- Schoemann, V, Becquevort S, Stefels J, Rousseau V & Lancelot C**, (2005). *Phaeocystis* blooms in the global ocean and their controlling mechanisms: a review. *J Sea Res* 53, 43-66.

- Seenivasan R, Sausen N, Medlin LK, Melkonian M**, (2013) *Picomonas judraskeda* gen. et sp. nov.: the first identified member of the Picozoa phylum nov., a widespread group of picoeukaryotes, formerly known as “picobiliphytes”. (RF Waller, Ed.). *PloS one*, 8, e59565.
- Send U, Font J, Krahnemann G, Millot C, Rhein M and Tintore J**, (1999). Recent advances in observing the physical oceanography of the western Mediterranean Sea. *Progress in Oceanography*, 44(1-3) :37–64.
- Seuront L, Vincent D, Mitchell JG** (2006). Biologically-induced modification of seawater viscosity in the eastern English Channel during a *Phaeocystis globosa* bloom. *Journal of Marine Systems*. 6 : 118- 133.
- Sheldon RW, Prakash A & Sutcliffe WH**, (1972). The size distribution of particles in the ocean. *Limnol. Oceanogr*, 17(3), 327-340.
- Shendure J, Ji H**, (2008). Next-generation DNA sequencing. *Nature biotechnology*, 26, 1135–45.
- Sherr BF, Sherr E, Caron DA, Vault D and Worden AZ** (2007). Oceanic protists. *Oceanography* 20: 130–134.
- Sherr EB, Sherr BF** (1994) Bacterivory and herbivory: key roles of phagotrophic protists in pelagic food webs. *Microbial Ecology*, 28, 223–235.
- Shi X, Marie D, Jardillier L, Scanlan D, Vault D**, (2009) Groups without cultured representatives dominate eukaryotic picophytoplankton in the oligotrophic South East Pacific Ocean. *PLoS One*, 4, e7657.
- Simpson AG**, (2003). Cytoskeletal organization, phylogenetic affinities and systematics in the contentious taxon Excavata (Eukaryota). *Int J Syst Evol Microbiol*. 53, 1759-77.
- Simpson AG, Inagaki Y et Roger AJ**, (2006). Comprehensive multigene phylogenies of excavate protists reveal the evolutionary positions of "primitive" eukaryotes. *Mol Biol Evol*. 23, 615-25.
- Sintes E, Martinez-Taberner A, Moya G, Ramon G**, (2004). Dissecting the microbial food web: structure and function in the absence of autotrophs. *Aquatic Microbial Ecology* 37:283-293
- Skovgaard A, Massana R, Balague V and Saiz E**, (2005). Phylogenetic position of the copepod-infesting parasite *Syndinium turbo* (Dinoflagellata, Syndinea). *Protist* 156, 413–423. doi: 10.1016/j.protis. .08.002
- Slapeta J, Moreira D et Lopez-Garcia P**, (2005). The extent of protist diversity: insights from molecular ecology of freshwater eukaryotes. *Proc Biol Sci*. 272, 2073-81.
- Steenkamp ET, Wright J et Baldauf SL**, (2006). The protistan origins of animals and fungi. *Mol Biol Evol*. 23, 93-106.
- Stoeck T, Hayward B, Taylor GT, Varela R, Epstein SS**, (2006). A multiple PCR-primer approach to access the microeukaryotic diversity in environmental samples. *Protist*. 157 (1), 31-43.
- Stoeck T, Zuendorf A, Breiner HW, Behnke A**, (2007). A molecular approach to identify active microbes in environmental eukaryote clone libraries. *Microb Ecol*. 53(2), 328-339.

- Stoecker DK** (1987). Photosynthesis found in some single-cell marine animals. *Oceanus* 30, 49-53.
- Stoecker DK**, (1999). Mixotrophy among dinoflagellates. *J. Eukaryot. Microbiol.* 46:397–401
- Stoecker DK, Gifford DJ, Putt**, (1994). M.Preservation of marine planktonic ciliates: lossess and cell shrinkage during fixation. *Mar. Ecol. Prog. Ser.*;110:293-299
- Sun Y, Cai Y, Liu L, Yu F, Farrell ML, McKendree W, & Farmerie W** (2009). ESPRIT: estimating species richness using large collections of 16S rRNA pyrosequences. *Nucleic acids research*, 37(10), e76- e76.
- Tagkopoulos I, Liu Y-C, Tavazoie S**, (2008). Predictive behavior within microbial genetic networks. *Science* 320. pp: 1313–1317.
- Taylor FJR (ed)**, (1987a). *The Biology of DinoXagellates*. Bot Monogr 21:i–xii, 1–785
- The Medatlas Group, Gaillard P, Ravazzola P, Kontolios C, Arrivet L, Athanassoulis GA, Stefanakos CN, Gerostathis TP, Cavaleri L, Bertotti L, Sclavo M, Ramieri E, Dentone L, Noel C, Viala C et Lefevre JM**, (2004). Wind and wave atlas of the Mediterranean sea. Rapport technique.
- Thronsdén J**, (1976). Occurrence and productivity of small marine flagellates. *Norw.J.Bot.*23:269-293.
- Tillmann U, Hesse KJ** (1998). On the quantitative importance of heterotrophic microplankton in the northern German Wadden Sea. *Estuaries* 21:585–596
- Tréguer P, Pondaven P**, (2000). Global change: Silica control of carbon dioxide. *Nature*, 406(6794), 358- 359.
- Triadó-Margarit X, Casamayor EO**, (2012). Genetic diversity of planktonic eukaryotes in high mountain lakes (Central Pyrenees, Spain). *Environmental microbiology*, 14, 2445–56.
- Turley CM** (1999). The changing Mediterranean Sea-a sensitive ecosystem ? *Progress in Oceanography* 44, 387-400.
- Tunin-Ley A, Ibañez F, Labat JP, Zingone A, Lemée R**, (2009). Phytoplankton biodiversity and NW Mediterranean Sea warming: Changes in the dinoflagellate genus *Ceratium* in the 20th century. *Mar. Ecol. Prog. Ser.* 375, 85–99. doi: 10.3354/meps07730.
- Tunin-Ley A, Labat JP, Gasparini S, Mousseau L, Lemée R**, (2007). Annual cycle and diversity of species and infraspecific taxa of *Ceratium* (Dinophyceae) in the Ligurian Sea, northwest Mediterranean. *J. Phycol.* 43, 1149–1163. doi: 10.1111/j.1529-8817.00417.x.
- Ulses C** (2006). Dynamique océanique et transport de la matière particulaire dans le Golfe du Lion: Crue, tempête et période hivernale. Thèse de doctorat, Université Paul Sabatier.
- Ulses C, Estournel C, Durrieu de Madron X et Palanques A**, (2008) Suspended sediment transport in the gulf of lions (nw mediterranean): Impact of extreme storms and floods. *Continental Shelf Research* 28, 2048-2070. 10.1016/j.csr.2008.01.015.

- Umami SF, Beran A**, (2003) Seasonal variations in the dynamics of microbial plankton communities: first estimates from experiments in the Gulf of Trieste, northern Adriatic Sea. *Mar Ecol Prog Ser* 447:1–16
- Vandenkoornhuysse P, Baldauf SL, Leyval C, et al**, (2002). Extensive fungal diversity in plant roots. *Science*. 295, 2051.
- Vantrepotte V**, (2003). Caractérisation Bio-Optique des eaux côtières en Manche orientale pour l'estimation de la production primaire et le suivi des poussées phytoplanctoniques : application à la télédétection satellitaire « couleur de l'eau » en milieu côtier. Thèse de doctorat, Université du Littoral Côte d'Opale, 272 pp.
- Vaquer A, Troussellier M, Courties C. & Bibent, B**, (1996). Standing stock and dynamics of picoplankton in the Thau lagoon (Northwest Mediterranean coast). *Limnol. Oceanogr.* 41:1821–8.
- Vaulot D, Eikrem W**, (2008). The diversity of small eukaryotic phytoplankton ( $\leq 3 \mu\text{m}$ ) in marine ecosystems. *FEMS microbiology reviews*, 32, 795–820.
- Velegrakis AF, Michel D, Collins MB, Lafite R, Oikonomou EK, Dupont JP, Huault MF, Lecouturier M, Salomon JC, Bishop C** (1999). Sources, sinks and resuspension of suspended particulate matter in the eastern English Channel. *Cont. Shelf Res.* 19: 1933-1957.
- Verity P, Wassmann P, Frischer ME, Howard-Jones MH, Allen AE**, (2002). Grazing of phytoplankton by microzooplankton in the Barents Sea during early summer. *J Mar Syst* 38:109–123
- Vladstein O, Stibor H, Lippert B, Loseth K, Roederer W, Sundt-Hansen L, Olsen Y**, (2004). Moderate increase in the biomass of omnivorous copepods may ease grazing control of planktonic algae. *Mar Ecol Prog Ser* 270:199–207.
- Walker G, Dorrell R, Schlacht A, Dacks J**, (2011a). The eukaryotes in 2011. *84.237.21.152*, 1–26.
- Ward JH**, (1663) Hierarchical Grouping to Optimize an Objective Function, *Journal of the American Statistical Association*, 58, 236–244.
- WK Li, McLaughlin FA, Lovejoy C & Carmack EC**, (2009). Smallest algae thrive as the Arctic Ocean freshens. *Science*, 326(5952), 539-539.
- Woese CR, Fox GE.**,(1977). Phylogenetic structure of the prokaryotic domain: The primary kingdoms. *Proceedings of the National Academy of Sciences*, 74, 5088–5090.
- Woese CR, Kandler O, Wheelis ML**, (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proceedings of the National Academy of Sciences*, 87, 4576–4579.
- Woods ED, Armstrong FAJ, Richards FA** (1967). Determination of nitrate in sea water by cadmium- copper reduction to nitrite. *Journal of Marine Biological Association of United Kingdom*. 47: 23-31.
- Worden AZ**, (2006). Picoeukaryote diversity in coastal waters of the Pacific Ocean. *Aquat. Microb. Ecol.* 43:165–75
- Wu P and Haines K** (1996). Modeling the dispersal of Levantine Intermediate Water and its role in Mediterranean deep water formation. *Journal Of Geophysical Research - Oceans*, 101(C3) :6591–6607.

**Zarauz L & Irigoien X**, (2008). Effects of Lugol's fixation on the size structure of natural nano–microplankton samples, analyzed by means of an automatic counting method. *Journal of plankton research*, 30(11), 1297-1303.

**Zhu F et al**, (2005). Mapping of picoeukaryotes in marine ecosystems with a quantitative PCR of the 18S rRNA gene. *FEMS Microbiol. Ecol.* 52, 79–92.

**Zinger L, Gobet A, Pommier T**, (2011). Two decades of describing the unseen majority of aquatic microbial diversity. *Molecular Ecology*, 21, 1878–96.

**Zinser ER, Lindell D, Johnson ZI, Futschik ME, Steglich C, Coleman ML, Wright MA, Rector T, Steen R, McNulty N, Thompson LR, Chisholm SW**, (2009). Choreography of the transcriptome, photophysiology, and cell cycle of a minimal photoautotroph, *Prochlorococcus*. *PLoS ONE* 4: e5135.

# **ANNEXES**





## VALORISATION DE LA THESE

### 1. Publications acceptées

Christaki U, Genitsaris S, Monchy S, Li LL, **Rachik S**, Breton E, Sime-Ngando T (2017). Parasitic eukaryotes in a meso-eutrophic coastal system with marked phytoplankton blooms *Front. Mar. Sci.* 4:416. doi: 10.3389/fmars.2017.00416

**Sara Rachik**, Urania Christaki, Luen Luen Li, Savvas Genitsaris, Elsa Breton, Sébastien Monchy **Diversity overview and potential activity patterns of planctonic eukaryotic microbes in a mesoeutrophic coastal area (Eastern English Chanel)** (*PlosOne*)

### 2. Publications en preparation

**Sara Rachik**, Sébastien Monchy, Bernard Queguiner, Elsa Breton, Urania Christaki **Small and (sub)mesoscale variability of protistan plankton communities in oligotrophic coastal ecosystems (NW Mediterranean, Gulf of Lions) assessed by NGS** (en préparation)

### 3. Publications annexes

Elsa Breton, Juliette Fauchot, **Sara Rachik**, Luis Felipe Artigas, Jean-Michel Brylinski<sup>0</sup>, Pascal Claquin, Vincent Cornille, Olivier Crispi, Muriel Crouvoisier, Valérie Gentilhomme, Savvas Genitsaris, Jean-David Grattepanche, Emilie Grosstefan, Oriane Joly, Stéphane Karasiewicz, Eric Lecuyer, Sébastien Lefebvre, Jean-Paul Lehodey, Fabrice Lizon, Arnaud Louchard, Sébastien Monchy, David Pecqueur<sup>3</sup>, Adrien Poquet, Christophe Salmeron, Benoit Sautour, and Urania Christaki **One step forward: developing a trait-based approach for understanding the seasonal succession of phytoplankton species in coastal waters of the eastern English Channel** (en préparation)

### 4. Présentations orales

**Sara RACHIK**, Urania CHRISTAKI, Luen Luen LI, Savvas GENITSARIS, Elsa BRETON, Sébastien MONCHY **Diversity overview and potential activity patterns of planctonic protist in a mesoeutrophic coastal area (Eastern English Chanel)** Conference EEC-2017 "Océanographie et Géosciences en Manche orientale – Passé, Présent, futur", Wimereux (France), Laboratoire d'Océanologie et Géosciences (CNRS UMR 8187 LOG), 6 au 8 juin 2017 (oral)

### 5. Autres présentations ou posters

Elsa Breton, **Sara Rachik**, Sébastien Monchy, Luis Felipe Artigas, Jean-Michel Brylinski, Vincent Cornille, Olivier Crispi, Muriel Crouvoisier, Valérie Gentilhomme, Savvas Genitsaris, Jean-David Grattepanche, Emilie Grosstefan, Eric Lecuyer, Arnaud Louchard, David Pecqueur, Christophe Salmeron, Benoit Sautour, et Urania Christaki Colloque Evoleco, Bordeaux, 5 au 7 décembre 2017 (Oral)

**Rachik S**, Genitsaris S, Breton E, Li LL, Sime-Ngando T, Monchy S, Christaki U. Symbionts and decomposers in a meso-eutrophic coastal system with marked phytoplankton succession ASLO meeting, Hawaii, 27 Feb-3 March 2017 (poster)

Genitsaris S, **Rachik S**, Li L, Breton E, Kormas K, Sime-Ngando T, Monchy S & Christaki U Symbionts And Decomposers: The Hidden Diversity of a Well-Studied Area ASLO 2016, Summer Meeting, Santa Fe, New Mexico 5-10 June 2016 (oral)

B. Quéguiner, ... **S. Rachik et al.**, Relationships between the structures of planktonic ecosystems in the north-western Mediterranean, the physical settings and the biogeochemical cycles – Feedback on the long-term observation of the pelagic environment, MERMEX 2015 Workshop, 7-10 April 2015, Marseille (oral)