



**HAL**  
open science

# Dynamique temporelle des communautés microbiennes eucaryotes en lien avec les forçages climatiques et anthropiques : approche paléolimnologique basée sur le séquençage massif d'ADN sédimentaire

Éric Capo

► **To cite this version:**

Éric Capo. Dynamique temporelle des communautés microbiennes eucaryotes en lien avec les forçages climatiques et anthropiques : approche paléolimnologique basée sur le séquençage massif d'ADN sédimentaire. Biodiversité et Ecologie. Université Grenoble Alpes, 2016. Français. NNT : 2016GREAA032 . tel-01517017

**HAL Id: tel-01517017**

**<https://theses.hal.science/tel-01517017>**

Submitted on 2 May 2017

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

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

## THÈSE

Pour obtenir le grade de

### **DOCTEUR DE LA COMMUNAUTÉ UNIVERSITÉ GRENOBLE ALPES**

Spécialité : **Biodiversité, Ecologie, Environnement**

Arrêté ministériel : 25 mai 2016

Présentée par

**Eric CAPO**

Thèse dirigée par **Isabelle DOMAIZON** et **Fabien ARNAUD**

préparée au sein du

**UMR CARTEL INRA (Université Savoie Mont-Blanc)**  
dans l'**École Doctorale SISEO**

**Dynamique temporelle des communautés  
microbiennes eucaryotes en lien avec les  
forçages climatiques et anthropiques**  
Approche paléolimnologique basée sur le  
séquençage massif d'ADN sédimentaire

Thèse soutenue publiquement le **19 décembre 2016**,  
devant le jury composé de :

**Mme Laure GUILLOU**

Directeur de recherche CNRS, UPMC (Rapporteur)

**Mme Purificación LOPEZ-GARCIA**

Directeur de recherche CNRS, Université Paris Sud (Rapporteur)

**M. Jan PAWLOWSKI**

Professeur associé, Université de Genève (Président)

**M. Thomas POMMIER**

Chargé de recherche INRA, Université Lyon 1 (Membre)

**M. Pierre SABATIER**

Maître de conférences CNRS, Université Savoie Mont Blanc (Membre)

**M. Didier DEBROAS**

Professeur, Université Clermont Ferrand (Co-encadrant de thèse)





À mon grand-père, René Capo





‘’ ...Le nain, qui jugeait quelquefois un peu trop vite, décida d’abord qu’il n’y avait personne sur la terre. Sa première raison était qu’il n’avait vu personne. Micromégas lui fit sentir poliment que c’était raisonner assez mal : « Car, disait-il, vous ne voyez pas avec vos petits yeux certaines étoiles de la cinquantième grandeur que j’aperçois très distinctement; concluez-vous de là que ces étoiles n’existent pas ? Mais dis le nain, j’ai bien tâté. Mais, répondit l’autre, vous avez mal senti »... ‘’

Micromégas, 1752, Voltaire



## Remerciements

Je remercie en premier lieu Isabelle Domaizon qui m'a fait confiance au cours de ces trois dernières années. Merci pour votre patience, tout ce que vous m'avez appris et m'apprendrez encore. Je remercie également mes co-encadrants de thèse, Fabien Arnaud et Didier Debroas qui ont beaucoup apporté à ma formation scientifique et à ces travaux de thèse.

En second lieu, je remercie la région Rhône-Alpes pour le financement de mon contrat de thèse au sein de l'ARC environnement. Je souhaite également remercier Bernard Montuelle pour son accueil au sein du laboratoire CARTELE et Emmanuel Trouvé au sein de l'école doctorale SISEO.

Je souhaite ensuite remercier les rapporteurs et membres de mon jury de thèse (dans l'ordre alphabétique) : Dr Laure Guillou (CNRS, UPMC), Dr Purificación Lopez-García (CNRS, Paris Sud), Pr Jan Pawlowski (Université de Genève), Dr Thomas Pommier (CNRS, Université Lyon 1) et Dr Pierre Sabatier (CNRS, Université Savoie Mont-Blanc). Merci beaucoup d'avoir accepté d'évaluer ce travail de thèse. Ce manuscrit de thèse a pu voir le jour grâce au soutien sans faille d'Isabelle que je remercie mille fois pour les dizaines d'heures passées à corriger et relire. Je remercie également beaucoup Didier, Valentin et Cécile pour les relectures de ce manuscrit.

La recherche ne se fait pas sans échanges et collaborations et ainsi je remercie tous les chercheurs et scientifiques avec qui j'ai pu collaborer au cours de ma thèse : le LMGE (Didier Debroas), le laboratoire EDYTEM (Fabien Arnaud, Cécile Pignol et Anne-Lise Develle), Chrono-Environnement (Typhaine Guillemot, Laurent Millet, Vincent, Bichet, Emilie Gauthier, Simon Belle), la team paleo de l'Université de McGill (Irène Gregory-Eaves, Amanda Winegardner, Joanna Gauthier, Jennifer Barrow) et la team paleo de l'Université d'Umeå. Tack Christian Bigler for our fruitful collaboration and your interest for our research. From Sweden, I also want to thank Daniela Figueora and Agneta Andersson who gave me the opportunity to work a bit on marine bacterial communities. Merci également à Franck Lejzerowicz et Jan Pawlowski (Université de Genève) pour les discussions et les collaborations. Pour le travail au laboratoire, je remercie énormément Cécile Chardon, la khaleesi du plateau bio mol, et Léandre Vasseur qui ont significativement contribué à l'acquisition des données présentées dans ces travaux.

Merci aussi à tous les scientifiques, étudiants et autres que j'ai pu croiser au cours des 3 dernières années et qui m'ont permis d'apprécier le monde de la recherche. Je pense notamment à Stéphan Jacquet, Franck Lejzerowicz, François Keck, Valentin Vasselon, Typhaine Guillemot, Amanda Winegardner, Francesco Ficetola, Thomas Pommier, Jean-Phillipe Jenny, Perrine Dranguet, Emilie Lyautey, Victor Frossard, Camille Thomas, Lucas Deschamps, Thomas Pollet, Jean-François Briand, Vincent Delafont, Stefan Lambert, Frederic Gaboyer, Sara Rivas, Nicolas Rascovan, Matthieu Mulot et j'en oublie forcément. Il y a eu également avant et pendant ma thèse de nombreux scientifiques qui

sont responsables de ma passion pour le monde microbien. Merci à vous : Paul Capdevielle, Céline Brochier, Purificacion Lopez-Garcia, Jed Fuhrman, Ashley Shade, Ramiro Logares, Ramon Massana, Patrick Forterre, Craig Venter, Claude Charpy-Roubaud, Marie-Andrée Mandrand, Yolanda del Amo, Lorena Via Ordorika, Albert Barberan, Samuel Chaffron, Lucie Bittner, Lionel Guidi, Christophe Laplace-Treuture, Alain Dutartre et bien d'autres.

Enfin, quelques mots concernant ma vie au sein du CARTELE ces 3 dernières années. Je pense que l'adjectif « dynamique » est de mise tant l'atmosphère, le cadre et les rencontres que j'ai pu faire ont contribué à ma joie d'aller au labo tous les matins. De façon formelle, je voudrais remercier les collègues au sein de la station de Thonon (Cécile, Stéphan, Agnès, Frédéric R., Bernard, Emilien, Orlane, Marie, Pascal, Laurent, Severine, Leslie et bien d'autres) et les collègues du Bourget (Emilie, Victor, Florent, David, Camille) pour les échanges et leur accueil toujours très sympathique. De façon moins formelle, je souhaite ainsi remercier les fourbouilles (ou fambouilles selon l'appartenance géographique), le club pcl, Valou la boule (dit aussi polyvalou) et Frederic Vernaz (HeisenFred, dit aussi le roux), le Docteur Keck, Stephan, Cécile, Gaël, Kalman, Teofana, Anne, Benoit Tourne-Toi et tous les autres doc, post docs, stagiaires. Pour les dizaines de kilomètres à ne pas user mes chaussures, merci beaucoup à tous mes covoitureurs intra-muros : Fred, Val, Gaël, Anne et Cécile... pour citer les plus dévoués. Je tiens également à remercier de façon non-exhaustive: les projets bourgaliku et bourciflette, les loutres qui font du piano, la chasse aux œufs de Pâques, les créations artistiques du bureau de co-PhD exceptionnel (donnant du cachet au labo : la poupée, le robot, le doll-ball, le totem Nescafé, les dinosaurus et autres stupidités), les pauses chocolats au lait, le régakeck, le Tauros chopé outre atlantique, les foires de crêtes et bien entendu le club de ping pong et ses tournois légendaires.

Malgré l'éloignement, les amis étaient là et je tiens particulièrement à remercier Jule (mon témoin, mon coloc, mon gaffas barratas), Laura, Antoine, Pierre, Jennifer, les Ecureuils, les copains de Thonon (Guillaume, Dan, Mya, Mélanie). Merci également à Ellen, ma voisine pour tous les services rendus et pour ces 2 ans de presque colocation. La séparation la plus dure sera peut être celle de minette et Mr Chat. Je remercie ma famille pour avoir supporté mon absence pendant ces 3 ans (ils ne se sont pas trop plaints en fait) plus particulièrement mon père et ma mère. Parce que sans lui, je n'en serais pas où j'en suis aujourd'hui, je tiens particulièrement à remercier mon grand-père, René Capo, pour m'avoir guidé vers cette soif de découverte de la nature et l'intérêt qu'il y a à la protéger. Merci beaucoup papi, j'espère pouvoir un jour accomplir le quart de ce que tu as accompli.

Enfin, je remercie ma femme, Marine, pour son amour et sa patience au cours des 8 dernières années. Malgré la séparation géographique et les aller-retours en train pendant 3 ans, notre amour est aujourd'hui encore plus fort avec l'accouchement imminent de nos deux magnifiques bébés-thèses jumelles. Marine Vandewalle-Capo, je t'aime autant que les souris moumou aiment le gruyère.

## Résumé

L'eutrophisation et le réchauffement climatique sont reconnus comme des forçages majeurs du fonctionnement des lacs. Toutefois les connaissances concernant la réponse des communautés microbiennes eucaryotes à ces forçages sont encore très lacunaires, alors même que les microbes eucaryotes, porteurs d'une vaste diversité taxonomique et fonctionnelle, sont des acteurs clés des réseaux trophiques lacustres. La pertinence des approches paléolimnologiques pour comprendre les impacts de ces forçages sur les communautés lacustres n'est plus à démontrer, mais aujourd'hui l'intégration des outils moléculaires pour analyser l'ADN archivé dans les sédiments offre des opportunités nouvelles pour reconstituer la dynamique passée de la biodiversité lacustre. Dans ce cadre, en s'appuyant sur le couplage entre paléolimnologie et outils de séquençage massif appliqués à l'ADN sédimentaire, ces travaux ont pour but (i) d'apporter des connaissances concernant la préservation de l'ADN des microbes eucaryotes dans les sédiments lacustres (ii) d'appliquer l'approche de paléogénétique sur des carottes sédimentaires issues de 3 lacs pour révéler la dynamique à long terme (de la décennie au millénaire) des microbes eucaryotes en lien avec l'évolution des conditions climatiques et anthropiques. Les résultats acquis sur le lac du Bourget ont permis de mettre en évidence l'efficacité d'archivage de l'ADN planctonique dans les sédiments récents pour la plupart des groupes eucaryotes (notamment chrysophycées, chytrides, chlorophytes, cercozoaires, ciliés, dinophycées). A partir d'une collection de carottes (issues du lac suédois Nylandssjön), l'effet de la diagenèse s'opérant au cours des premières années d'enfouissement a été évalué, permettant de démontrer que si la richesse taxonomique n'est pas impactée, des variations peuvent être détectées dans la structure de la communauté au cours des 10 premières années d'archivage avec une stabilisation du signal au-delà de cette période. L'approche paléogénétique a, en parallèle, été déployée d'une part à l'échelle du siècle sur deux lacs de même typologie mais ayant subi des niveaux d'eutrophisation contrastés, et d'autre part à une échelle temporelle plus longue (2200 ans) pour deux lacs de typologie contrastée (lac du Bourget, France et Igaliku, Groenland). Les résultats acquis démontrent que des réarrangements des communautés s'opèrent de manière concomitante aux périodes climatiques (réchauffement médiéval, petit âge glaciaire, réchauffement récent), et que le réchauffement climatique au cours des 30 dernières années a plus particulièrement favorisé certains groupes, notamment la richesse et l'abondance des dinophycées (en condition non eutrophe ; lacs d'Annecy et du Bourget). Toutefois l'effet de l'eutrophisation est identifié comme le facteur le plus structurant, notamment dans le lac du Bourget (cas d'eutrophisation marquée,  $\sim 120 \mu\text{gP.L}^{-1}$ ). La forte influence du niveau d'eutrophisation est détectée sur la communauté eucaryote totale et plus particulièrement sur des groupes spécifiques tels que les chlorophytes et les ciliés. Les réarrangements majeurs de la communauté sont par ailleurs marqués par la mobilisation de taxons rares dans l'assemblage microbien eucaryote suggérant le rôle de la biosphère rare dans la capacité tampon des écosystèmes. Ces travaux pluridisciplinaires comptent parmi les premières études paléogénétiques appliquées aux microbes eucaryotes lacustres, contribuant de manière inédite aux connaissances de leur dynamique temporelle à long terme. Ces études tendent à confirmer le potentiel de ces approches pour reconstituer une vaste diversité de communautés lacustres. Les perspectives qui se dessinent dans la continuité de ces travaux concernent à la fois des aspects méthodologiques autour de la calibration du signal ADN archivé et la nécessité de déployer cette approche pour des lacs (sélectionnés) de typologies et histoires écologiques variées.

Mots clefs : microbes eucaryotes, séquençage massif, ADN sédimentaire, paléolimnologie, eutrophisation, réchauffement climatique



# Abstract

Eutrophication and climate warming are key factors governing lake functioning. However, there is a lack of knowledge about the response of microbial eukaryotic communities to these forcing factors even though microbial eukaryotes represent a huge taxonomic and functional diversity within lacustrine trophic networks. Paleolimnology has a well-established reputation for providing valuable insights into the drivers of biological assemblages over long time scales. The emergence of DNA analyses of lake sediments opens up many new opportunities for the reconstruction of past lacustrine biodiversity, including taxa that do not leave distinct morphological fossils. The present work aimed (i) to gain knowledge about the preservation of microbial eukaryotes DNA in lacustrine sediments (ii) to apply DNA-based methods to dated sediments in order to reveal the long-term dynamics (centennial to millennial) of microbial eukaryotes related to climatic and anthropogenic changes. The results obtained for Lake Bourget demonstrated the good efficiency of planktonic DNA archiving in recent sediments for most of microbial groups (chrysophyceae, chytrids, chlorophytes, cercozoa, ciliates, dinophyceae ...). In complement, the use of a unique collection of freeze cores of varved sediment (Lake Nylandssjön, Sweden) allowed to assess the effects of diagenetic processes on microbial eukaryotes DNA occurring during the first years of burying. While the richness of the microbial eukaryotic community was not impacted, modifications were detected on the community structure during the first 15 years after deposition, then the DNA signal became stable. The paleoecological approach was applied to quantify centennial to millennial-scale dynamics on two deep peri-alpine lakes selected for their contrasted trophic history (lakes Bourget and Annecy, France) and two lakes with contrasted typologies (Lake Bourget, France and Lake Igaliku, Greenland). The results showed that some community rearrangements were concomitant with climate events (i.e. medieval warming, little ice age, recent warming) and that the recent climatic warming (over the last 30 years) favored more particularly some microbial groups including the dinophyceae (in terms of richness and relative abundance, in lakes Annecy and Bourget). However, the eutrophication seemed to prevail as a driver of these biological assemblages, in particular for Lake Bourget submitted to a marked eutrophication (up to  $120 \mu\text{g P. L}^{-1}$  in the 1970s). The strong impact of the eutrophication was detected both at the whole community level and for specific groups such as chlorophytes and ciliates. Major rearrangements within eukaryotes community were also marked by the mobilization of rare taxa suggesting an important role of the rare biosphere as a reservoir of diversity to buffer the impacts of environmental stress. This multidisciplinary work thus provides new insights into the long-term dynamics of microbial eukaryotes communities. Our results confirm the potential of the application of high-throughput sequencing to sedimentary DNA for the lacustrine biodiversity reconstruction. Although these approaches are promising for further revolutionizing our understanding of long-term ecological dynamics, careful calibration studies are still to be conducted, as with any paleolimnological proxy. The generalization of our results is also to be tested using a sufficient number of lakes selected for their specific typologies and ecological histories (multi-lakes approach).

Key words: microbial eukaryotes, high-throughput sequencing, sedimentary DNA, paleolimnology, eutrophication, climate warming





# Table des matières

Liste des figures.....	i
Liste des tableaux.....	v
Liste des encadrés.....	vi
<b>I. Chapitre introductif .....</b>	<b>1</b>
<i>I.1. Contexte et problématiques.....</i>	<i>1</i>
I.1.1. Les lacs : des écosystèmes sensibles aux changements globaux et locaux.....	1
I.1.1.1. Contexte global des effets des activités humaines sur les écosystèmes.....	1
I.1.1.2. Quelques rappels sur les lacs.....	3
I.1.1.3. Le lac, sensible aux forçages environnementaux.....	6
I.1.1.4. L'archive sédimentaire, l'opportunité d'une fenêtre temporelle pertinente pour l'étude de la réponse des communautés biologiques aux forçages environnementaux.....	9
I.1.2. Apports de la paléocéologie moléculaire aux questions d'écologie microbienne.....	14
I.1.3. Les micro-eucaryotes : diversité, rôles et facteurs de régulation dans les milieux lacustres.....	20
I.1.3.1. Du microscope au séquençage à haut-débit de l'ADN.....	20
I.1.3.2. Les micro-eucaryotes, un compartiment diversifié.....	22
I.1.3.3. Rôle des micro-eucaryotes dans les réseaux trophiques pélagiques.....	29
I.1.3.4. Etat des connaissances sur la réponse des micro-eucaryotes à l'eutrophisation et aux changements climatiques.....	33
<i>I.2. Objectifs et stratégie des travaux de thèse.....</i>	<i>35</i>
I.2.1. Objectifs de la thèse.....	35
I.2.1.1. Axe 1 : Calibration/validation de l'utilisation de l'ADN sédimentaire.....	35
I.2.1.2. Axe 2 : Dynamique temporelle des communautés microbiennes en réponse aux forçages environnementaux.....	38
I.2.2. Approche méthodologique.....	41
I.2.3. Structure de la thèse.....	47
<b>II. Efficacité d'archivage de l'ADN des microorganismes eucaryotes de la masse d'eau vers les sédiments du lac du Bourget.....</b>	<b>49</b>
<i>II.1. Abstract.....</i>	<i>50</i>
<i>II.2. Introduction.....</i>	<i>51</i>
<i>II.3. Materials and Methods.....</i>	<i>53</i>
III.3.1. Study site.....	53
II.3.2. Sampling.....	53
II.3.3. Molecular Analysis.....	54
II.3.4. Data Accessibility.....	57
II.3.5. Data Analysis and Statistical Tests.....	58
<i>II.4. Results.....</i>	<i>59</i>
II.4.1. Richness and Taxonomic Composition of Unicellular Eukaryotes in Water and Sediments.....	59
II.4.2. Phylogenetic Diversity Archived in Sediment.....	61
II.4.3. Potential Links Between Recurrence and Abundance of PUs in Water and Their Archiving in Sediment.....	63
<i>II.5. Discussion.....</i>	<i>63</i>
II.5.1. Overall Diversity of Protistan Taxa in Water and Sediment.....	63
II.5.2. Active Protists Do Not Bias the Paleosignal.....	64
II.5.3. Efficiency of Planktonic DNA Archiving in Recent Sedimentary Deposits.....	65
<i>II.6. Supporting information.....</i>	<i>68</i>
<b>III. Préservation de l'ADN sédimentaire au cours de l'enfouissement.....</b>	<b>71</b>
<i>III.1. Abstract.....</i>	<i>73</i>

<i>III.2. Introduction</i> .....	73
<i>III.3. Materials &amp; Methods</i> .....	76
III.3.1. Study site and sediment sampling.....	76
III.3.2. Molecular data .....	77
III.3.3. Bioinformatics and statistical analysis.....	77
III.3.4. Data treatment and statistical analyses .....	79
<i>III.4. Results</i> .....	81
III.4.1. DNA quantities .....	81
III.4.2. Richness and diversity .....	81
III.4.3. Community structure .....	84
<i>III.5. Discussion</i> .....	90
III.5.1. Overall, stability of the DNA signal of micro-eukaryotes .....	90
III.5.2. Only one taxonomic group varied according to the core: focus on Fungi .....	92
III.5.3. Variations in relative abundance of some OTUs in the uppermost strata .....	93
<i>III.6. Conclusion</i> .....	95
<i>III.7. Supporting information</i> .....	96
<b>IV. Impact des forçages environnementaux au cours des 2200 dernières années sur les communautés microbiennes eucaryotes des lacs Bourget (France) et Igaliku (Groenland)</b> .....	<b>101</b>
<i>IV.1. Abstract</i> .....	103
<i>IV.2. Introduction</i> .....	103
<i>IV.3. Materials and methods</i> .....	106
IV.3.1. Study sites .....	106
IV.3.2. Field sampling & laboratory sub-sampling .....	106
IV.3.3. Dating of sediment cores .....	107
IV.3.4. Molecular analysis.....	108
IV.3.5. Bioinformatics processing .....	109
IV.3.6. Data treatment and statistical analyses .....	111
<i>IV.4. Results</i> .....	114
IV.4.1. OTU richness and taxonomic composition.....	114
IV.4.2. Temporal tipping points detected for the structure of eukaryotes community .....	114
IV.4.3. Community rearrangements over time .....	118
<i>IV.5. Discussion</i> .....	123
IV.5.1. Methodological aspects .....	123
IV.5.2. Sedimentary DNA reveals past microbial eukaryotes community .....	124
IV.5.3. A temporal view on dominance and rarity .....	125
IV.5.4. First clues of synchrony between MECs shifts and climatic fluctuations .....	126
IV.5.5. The impacts of marked eutrophication seems to overpass that of other forcing factors .....	127
IV.5.6. Toward 'new' biological markers from sedimentary DNA.....	128
<i>IV.6. Additional analysis</i> .....	131
IV.6.1. Application of network analysis to the view of the temporal dynamics of microbial eukaryotes communities via the network analysis time series.....	131
<i>IV.7. Supporting information</i> .....	134
<b>V. Impact des forçages environnementaux au cours des 100 dernières années sur les communautés microbiennes eucaryotes des lacs Annecy et Bourget (France)</b> .....	<b>145</b>
<i>V.1. Abstract</i> .....	147
<i>V.2. Introduction</i> .....	148
<i>V.3. Materials and Methods</i> .....	150
V.3.1. Study sites and forcing factors .....	150
V.3.2. Coring, dating and sub-sampling .....	151
V.3.3. Molecular analysis .....	151

V.3.4. Bioinformatics processing applied for massive sequencing data .....	152
V.3.5. Specific procedures for paleogenetics work .....	153
V.3.6. Data treatments and statistical analyses .....	154
V.4. <i>Results</i> .....	158
V.4.1. Lake settings and environmental forcing factors .....	158
V.4.2. Richness, composition and structure of MECs .....	158
V.4.3. Temporal rearrangements of MEC related to environmental change .....	161
V.4.4. Network structure, modularity, connectivity and links with environmental changes .....	164
V.5. <i>Discussion</i> .....	165
V.5.1. Characterization of MECs from sedimentary archives .....	165
V.5.2. The [P] range as a major driver of MECs in both lakes .....	167
V.5.3. No return to pre-eutrophication conditions .....	169
V.5.4. Simultaneous effects of re-oligotrophication and warming since the late 1980s .....	170
V.6. <i>Conclusion</i> .....	172
V.7. <i>Additional analysis</i> .....	173
V.7.1. Detection of change points in MECs richness/diversity (bcp analysis) .....	173
V.7.2. Detection of break points in the structure of MECs (MRT analysis) .....	175
V.7.3. Focus on the temporal dynamics of rare OTUs .....	177
V.8. <i>Supporting Information</i> .....	179
V.9. <i>Analyses complémentaires</i> .....	207
V.9.1. Contexte et stratégie d'étude .....	207
V.9.2. Matériels et méthodes .....	207
V.9.3. Résultats .....	210
V.9.4. Conclusion .....	211
<b>VI. Discussion et Perspectives .....</b>	<b>213</b>
VI.1. <i>Autour de la validation du proxy ADN sédimentaire en milieu lacustre</i> .....	213
VI.1.1. De la masse d'eau au sédiment : une bonne préservation de l'ADN des communautés de micro-eucaryotes de la masse d'eau (lac profond tempéré) .....	215
VI.1.2. Les effets de la diagénèse précoce sur l'ADN des micro-eucaryotes lacustres : effets très modérés pour les sédiments varvés du Lac Nylandssjön .....	216
VI.1.3. Effet du vieillissement des sédiments sur l'ADN de micro-eucaryotes au cours de longues périodes de temps .....	218
VI.1.4. Une préservation différentielle de l'ADN des communautés biologiques du lac au cours de l'enfouissement dans les sédiments ? .....	219
VI.1.5. Focus sur quelques choix méthodologiques et les potentielles limites concernant l'approche déployée dans ces travaux .....	223
VI.2. <i>Vision à long-terme de la dynamique temporelle des communautés lacustres de micro-eucaryotes</i> ....	227
VI.2.1. Vers une étude de la dynamique temporelle des communautés multi-échelles : les réarrangements de la communauté vus à court et à long terme .....	227
VI.2.2. Dynamique temporelle des taxons abondants .....	228
VI.2.3. Contribution de la rare biosphère aux réarrangements .....	230
VI.2.4. Les données temporelles au service de la description d'un <i>core microbiome</i> eucaryote .....	233
VI.3. <i>La réponse des communautés de micro-eucaryotes aux pressions anthropiques et climatiques</i> .....	236
VI.3.1. Réponse des communautés de micro-eucaryotes lacustres à différents niveaux d'eutrophisation .....	238
VI.3.2. Effets de l'eutrophisation au sein d'un réseau trophique lacustre : quelques hypothèses issues des données paléolimnologiques concernant les interactions biotiques .....	239
VI.3.3. Les Chlorophytes, de potentiels indicateurs de l'état trophique du lac ? .....	240
VI.3.4. L'effet des fluctuations climatiques perceptibles sur les communautés eucaryotes .....	241
VI.3.5. Effets combinés des pressions anthropiques et climatiques sur les micro-eucaryotes: le cas de lacs en cours de ré-oligotrophisation exposés au réchauffement climatique .....	242
<b>VII. Conclusion .....</b>	<b>245</b>
<b>VIII. Annexe : Application des méthodes moléculaires en paléolimnologie, de nouvelles opportunités pour étudier la dynamique à long terme de la biodiversité lacustre .....</b>	<b>247</b>

<i>VIII.1. Abstract</i> .....	248
<i>VIII.2. Potential of DNA tools in paleolimnology</i> .....	248
VIII.2.1. DNA-based methods complement classical paleolimnology proxies and expand the range of organisms detectable in lake sediments. ....	249
VIII.2.2. DNA-based methods are uniquely positioned to answer critical questions in paleolimnology, some of which are not possible with classic paleoecological indicators. ....	250
<i>VIII.3. Some methodological aspects: precautions, limits and possible solutions</i> .....	251
VIII.3.1. The way DNA is archived in sediments matters (Figure VIII.2) .....	252
VIII.3.2. How do lakes characteristics affect DNA preservation?.....	253
VIII.3.3. What are the requisite precautions to ensure the authenticity of results?.....	256
<i>VIII.4. Applications of sedimentary DNA analyses to advance lacustrine ecology</i> .....	260
VIII.4.1 Both intra specific diversity and complex biological group can be described from sedimentary DNA .....	260
VIII.4.2. DNA from zooplankton resting stages .....	260
VIII.4.3. Targeting specific functions in the ecosystem .....	262
<i>VIII.5. Emerging topics &amp; challenges in applying sedimentary DNA to the field of lacustrine ecology</i> .....	262
VIII.5.1. DNA approaches to track the dynamics of invasive species .....	262
VIII.5.2. DNA approaches to reveal the presence of fish populations over time .....	263
VIII.5.3. Inferring ecological networks from molecular inventories obtained for complex assemblages .	264
VIII.5.4. Coupling spatial and temporal approach: temporal biogeography .....	265
VIII.5.5. The use of molecular data for paleoecological transfer functions.....	265
VIII.5.6. Integration of sedimentary DNA in biological collections and long-term environmental observatory: sediment core repository as a conservation tool for biodiversity .....	266
<i>VIII.6. Conclusion</i> .....	266
<b>Glossaire</b> .....	<b>267</b>
<b>Références</b> .....	<b>269</b>

# Liste des figures

Figure I.1. Schéma illustrant le niveau de risque associé aux effets irréversibles des perturbations humaines sur plusieurs variables importantes pour le fonctionnement des écosystèmes.....	2
Figure I.2. Distribution des lacs sur la planète. La valeur dL correspond à la densité moyenne des lacs par surface entre 1 et 10 km <sup>2</sup> .....	3
Figure I.3. Exemple d'un réseau trophique pélagique lacustre .....	5
Figure I.4. Schéma illustrant les effets directs et indirects des forçages environnementaux globaux et locaux sur le réseau trophique pélagique lacustre .....	7
Figure I.5. Illustration des intervalles de temps associés à chaque couche sédimentaire. ....	16
Figure I.6. Schéma de l'effet d'une perturbation sur la composition et les fonctions associées à une communauté microbienne .....	18
Figure I.7. Concentrations en phosphore estimées à partir du comptage des diatomées dans les archives sédimentaires.....	19
Figure I.8. Arbre phylogénétique des eucaryotes.....	23
Figure I.9. Schématisation des réseaux trophiques pélagiques lacustres. ....	29
Figure I.10. Assignation de rôles fonctionnels putatifs aux différents groupes de micro-eucaryotes.....	30
Figure I.11. Recherche bibliographique. ....	31
Figure I.12. Liste non exhaustive des principales publications portant sur l'étude des communautés microbiennes eucaryotes lacustres utilisant des analyses de biologie moléculaire, les questions d'intérêt, leurs régions climatiques, les techniques moléculaires utilisées et l'apport de ces publications à la compréhension des assemblages de micro-eucaryotes.....	32
Figure I.13. Schéma général de la position trophique du nano et micro-zooplancton au sein du réseau trophique planctonique et des effets potentiels directs et indirects des changements climatiques sur ces assemblages .....	34
Figure I.14. Schéma illustrant la période temporelle couverte par chacune des 10 carottes sédimentaire prélevées entre 1979 et 2007 (a) et les teneurs en carbone (b) et azote (c) .....	37
Figure I.15. Approche méthodologique appliquée dans ces travaux de thèse. ....	42
Figure I.16. Analyses GAMs. Figure illustrant la contribution relative de deux predictors aux changements dans une variable réponse.....	46
Figure II.1. Richness (number of PUs) distribution within the different phyla in water samples at 2 and 130 m depth and in sediment samples .....	58
Figure II.2. (A) Composition of the protistan community in water samples (at 2 and 130 m depth) and in sediment samples, expressed as the number of DNA reads obtained for each PU within the main phylogenetic groups; (B) focus on Cryptophyta PUs; (C) focus on Haptophyta PUs.....	60
Figure II.3. Boxplots and p value obtained for ANOVA tests performed to assess the relationships between PU recurrence in water at 2 m (a) and 130 m (b) depths and their recurrence in sediments, and, the relationships between the relative number of reads associated with PUs at 2 m (c) and 130 m (d) depth and their recurrence in the sediment .....	62

Figure II.S1. Supporting information presenting the flow chart for bioinformatic analyses using PANAM pipeline .....	68
Figure III.1. Experimental procedure used in the present study. ....	80
Figure III.2. Boxplots of DNA quantities obtained from the 12 varves in the cores C and C+6. ....	81
Figure III.3. Barplots of diversity indices: number of OTUs, Chao1 estimator, Shannon diversity, Simpson diversity, Pielou evenness obtained for the 48 samples. ANOVA p-values were calculated for each metric by comparing the values from the cores C and C+6. ....	83
Figure III.4. Hierarchical clustering tree computed using Bray-Curtis index (calculated from the OTU relative abundance dataset) and Jaccard index (calculated from presence/absence dataset) from the 48 molecular inventories. ....	84
Figure III.5. Values of Bray-Curtis dissimilarity (A) and Jaccard distance (B) obtained for each couple of replicates for the 12 analysed strata. ....	85
Figure III.6. Relative abundance (number of DNA sequences) of the 12 OTUs that contributed to more than 1 % to the dissimilarity between the molecular inventories obtained for the cores C and C+6, in each of the 12 analysed varves .....	86
Figure III.7. Distribution of DNA sequences in the molecular inventories with the main microbial eukaryotes groups. ....	88
Figure III.8. Plot of procrustean rotation analysis illustrated the potential differences in community structure for molecular inventories from core C and core C+6. ....	89
Figure III.S1. Quantities of DNA extracted for each sample (in $\mu\text{g.g}^{-1}$ of wet sediment).....	96
Figure III.S2. Distribution of the richness and relative abundance of fungal groups in the 48 molecular inventories. ....	96
Figure IV.1. For Lake Bourget (A) and Lake Igaliku (B) : Temporal dynamics of the relative abundance of each phylogenetic group of microbial eukaryotes and breakpoints detected by MRT analysis considering the whole list of OTUs and separately non-pigmented OTUs (NPI) and pigmented OTUs (PI). ....	115
Figure IV.2. Time series of Hill's numbers (OTU richness, transformed Shannon and Simpson indices) for Lake Bourget (A) and Lake Igaliku (B). ....	117
Figure IV.3. Non-metric multidimensional scaling analysis of microbial eukaryotes communities based on Bray-Curtis dissimilarity are presented for Lake Bourget (A) and Lake Igaliku (B) time series .....	119
Figure IV.4. Temporal relative abundances of the non-pigmented, pigmented and unknown OTUs, as well as the OTUs that contributed to more than 2 % of the total dissimilarity over time, are presented for Lake Bourget (A) and Lake Igaliku (B). ....	119
Figure IV.5. Temporal dynamics of oscillating OTUs in Lake Bourget and Lake Igaliku. Oscillating OTUs shifted between abundant ( $> 0.1\%$ ) and rare states ( $< 0.1\%$ of the number of DNA sequences, on average for each period). ....	121
Figure IV.6. Composition and temporal dynamics of the chlorophytan groups in Lake Bourget (A) and in Lake Igaliku (B) based on their relative abundance in terms of number of DNA sequences. ....	122
Figure IV.7. Co-occurrence networks for Lake Bourget (A) and Lake Igaliku (B). The nodes represented each one OTU. ....	130
Figure IV.8. Consensus network. ....	132

Figure IV.S1. Dating of sediment cores.....	134
Figure IV.S2. Figure showing from the MRT analysis, the relative error RE and the cross-validated relative error CVRE for the 3 studied datasets (all OTUs, non-pigmented OTUs, pigmented OTUs) in Lake Bourget (A) and Lake Igaliku (B).....	135
Figure IV.S3. Distribution of richness (in terms of OTUs number) within each phylogenetic group in Lake Bourget (A) and Lake Igaliku (B).....	136
Figure IV.S4. Temporal changes over the last century for (A) the anomalies of air temperature (annual average in 10-1 °C) ; Data from Auer et al. 2007 (B) Daphnia-inferred total phosphorus concentrations (annual mean concentrations over the 0-20 m water depths ± 95 confidence intervals) in µg P.L-1.....	137
Figure IV.S5. Non-metric multidimensional scaling analysis based on Bray-Curtis (relative abundance of OTUs) and Jaccard (presence/absence of OTUs) indices.....	138
Figure V.1. Temporal changes of (A) air temperature anomalies (annual average expressed in 10-1 °C; data from Auer et al., 2007) (B) Daphnia-inferred total phosphorus concentrations (annual mean concentration over the 0-20 m water depths ± 95 confidence interval expressed in µgP L-1; data from Alric et al., 2013, Berthon et al., 2014) (C) relative abundance (in terms of number of DNA sequences) of each microbial eukaryotes group, illustrated by heatmaps. Data cover the last century and are presented for lakes Annecy and Bourget. ....	157
Figure V.2. Non-metric multidimensional scaling analysis of microbial eukaryotes communities based on Bray-Curtis dissimilarity obtained for Lake Annecy (A) and Lake Bourget (B) time series. ....	159
Figure V.3. GAMs analysis results for Lake Annecy and Lake Bourget.....	160
Figure V.4. Co-occurrence networks for Lake Annecy and Lake Bourget MECs. ....	163
Figure V.5. Time series of richness, diversity and evenness indices for Lake Annecy and Lake Bourget. ....	174
Figure V.6. Breakpoints detected in the structure of MECs by MRT analysis.....	175
Figure V.7. Temporal dynamics of oscillating OTUs in Lake Annecy and Lake Bourget. Oscillating OTUs shifted between abundant (> 0.1 %) and rare states (< 0.1 % of the number of DNA sequences, on average for each period).....	178
Figure V.8. Changements temporels au cours du siècle dernier dans les lacs Annecy et Bourget de la quantité de copies de séquences d'ADN de Chlorophytes et la proportion relative des Chlorophytes au sein des assemblages de micro-eucaryotes.....	212
Figure V.S1. Bray-Curtis analysis pseudo-autocorrelation.....	179
Figure V.S2. The figure presents GAMs analysis results for Lake Annecy and Lake Bourget for the 4 indices describing the whole MECs and the 9 most abundant groups. ....	180
Figure V.S3. Topological roles of OTUs within the Annecy and Bourget network (A & B). Plot of intra vs inter-modular connectivity values illustrating the proportion of connectors, module hubs, network hubs and peripherals in each network.....	190
Figure V.S4. This figure illustrates the correlations between modules eigengene values and the environmental factors (phosphorus concentration and air temperature) in Lake Annecy and Lake Bourget. ....	191
Figure V.S5. Consensus network. (A) Only the OTUs retrieved within the 5 consensus modules were represented. Correspondence of the (B) Annecy set-specific and (C) Bourget set-specific with the consensus modules. ....	192



Figure VI.1. Synthèse des principaux résultats portant sur la préservation de l'ADN des micro-eucaryotes au cours de son archivage dans les sédiments lacustres. ....	214
Figure VI.2. Pourcentage de richesse commune et spécifique pour chaque réplicat de séquençage obtenu à partir de 2 extraits d'ADN issus d'une même strate sédimentaire .....	226
Figure VI.3. Abondance relative des OTUs les plus abondantes dans la série temporelle de la communauté microbienne eucaryote du lac Annecy au cours des 100 dernières années. ....	229
Figure VI.4. Courbes de Raréfaction/Extrapolation obtenus pour la richesse spécifique de chaque inventaire moléculaire des analyses du chapitre IV. ....	231
Figure VI.5. Schéma illustrant la définition des niveaux de rareté considéré dans les inventaires moléculaires. Bien que les conditionally rare taxa n'est pas été étudié en tant que tel, les traitements des seuils de rareté ont permis de les intégrer dans les réseaux de cooccurrence. ....	232
Figure VI.6: Schéma simplifié illustrant les périodes temporelles au cours desquelles les variations des deux forçages étudiés semblent avoir eu un effet sur la communauté de micro-eucaryotes des lacs étudiés. ....	237
Figure VIII.1. Schematic view of the lacustrine pelagic food web highlighting how DNA analysis can expand the range of organisms in paleolimnology. ....	249
Figure VIII.2. Possible DNA analyses according to the goal of the study and the way DNA is archived. ....	252

# Liste des tableaux

Tableau I.1. Synthèse de la littérature portant sur l'utilisation de l'ADN sédimentaire des milieux lacustres pour l'étude de la dynamique temporelle des communautés microbiennes.	11
Tableau I.2. Caractéristiques des carottes sédimentaires utilisées dans ces travaux de thèse.	41
Table II.1. The upper part of the table provides a description of the analyzed samples regarding the period covered by each sample (periods), the depth of sampling in the water column and in the sediment core (depth), the phosphorous concentration in the lake for each period ([P] in water in $\mu\text{g/L}$ ), and the number of samples analyzed. The concentration of phosphorous in lake water is provided as an indicator of the trophic state of the lake: long-term changes in lake total phosphorus concentrations were reconstructed from a diatom-inferred transfer function (Berthon <i>et al.</i> , 2013) based on the identification and quantification of diatoms' frustules preserved in sediment (all values with asterisks); total phosphorus concentrations for recent periods were also verified from direct measurement on lake water (chemical survey performed by the observatory on alpine lakes; all values with crosses). The lower part of the table present the number of PUs detected from massive sequencing in water and sediment samples, as well as percentages of PUs from water retrieved in sediment samples.	56
Table II.2. Taxonomic composition of active eukaryotes detected in the first two centimeters of sediment, expressed as the relative number of reads affiliated to the main phylogenetic groups.	58
Table II.S1. Number of sequences, OTUs, PUs, and values of Chao index associated to each sample.	69
Table II.S2. Number of reads obtained for the different phyla in water at 2 m and 130 m depth and in sediment samples.	70
Table II.S3. (simplified): Number of PUs and percentage of shared PUs between water and sediment samples: (A) number of PUs in all samples, (B) number of PUs in sediment samples, (C) number of PUs in all water samples and percentage of these PUs retrieved in sediments, (D- E) for each water layer (i.e. 2 m and 130 m depth) the table presents the total number of PUs (# PUs), the number of PUs found specifically during mixing or stratification (# specific PUs), and the percentage of these 'period-specific' PUs that is retrieved in sediment. Uncult: uncultured ES: environmental samples.	70
Table III.1. The table presents (i) richness and relative abundance of the main microbial eukaryotic groups (14 most diverse and abundant groups) and (ii) the significance of the comparison test (ANOVA) performed to compare these metrics for C and C+6; differences were considered significant when p-values were inferior or equal to 0.01	82
Table III.S1. Distribution of reads and OTUs numbers in the 48 molecular inventories within the main microbial eukaryotic groups.	97
Table III.S2. Results obtained from SIMPER analysis comparing the molecular inventories of the two cores considering the 6 most recent strata	99
Table IV.1. The table presents (i) for each lake, the numbers of DNA sequences, OTUs and oscillating OTUs within each taxonomic group (ii) the numbers of common OTUs in the two lakes and core OTUs found in all samples.	113

Table IV.S2. Richness and diversity indices obtained for each sample.....	139
Table IV.S3. For Lake Bourget (A) and Lake Igaliku (B): results obtained from SIMPER analysis considering the periods defined by ANOSIM analysis. ....	142
Table V.1. The table presents for each lake the number of DNA sequences and OTUs within each microbial eukaryotic group and, the numbers of common OTUs and total OTUs in the two lakes.....	156
Table V.2. Properties of the two individual networks built for Lake Annecy and Lake Bourget MECs. The topological indices were calculated considering only the edges values superior to 0.5.....	162
Table V.3. Distribution of total OTUs and oscillating OTUs within each microbial eukaryote group for each lake independently .....	176
Tableau V.4. Caractéristiques des 9 souches d’algues utilisées pour la vérification de la spécificité des amorces nucléotidiques.....	208
Tableau V.5. Affiliation taxonomique associée aux séquences obtenues pour les échantillons environnementaux de chaque lac .....	211
Table V.S1. Richness, diversity and evenness indices obtained for each sample.....	193
Table V.S2. For each lake independently, the table presents for each OTU, the percentage of DNA sequences, the occurrence through time, the modularity features (intraM = intramodular connectivity, interM = intermodular connectivity), the topological role within network and the taxonomic affiliation. ....	194
Table VIII.1. Our review of the literature focusing on the study of sedimentary DNA analyses in freshwater systems. ....	255
Table VIII.2. Main analytical steps, risks and possible solutions for implementation of DNA analyses in paleolimnological studies. ....	259

## Liste des encadrés

Encadré I.1. Dynamique annuelle de la structure de la masse d’eau des lacs.....	5
---	---

# I. Chapitre introductif

## I.1. Contexte et problématiques

### I.1.1. Les lacs : des écosystèmes sensibles aux changements globaux et locaux

#### I.1.1.1. Contexte global des effets des activités humaines sur les écosystèmes

L'impact des activités humaines sur les écosystèmes s'est aggravé depuis le début de l'ère industrielle au 19<sup>ème</sup> siècle (Vitousek 1997). Ainsi, l'espèce humaine transforme les paysages par la déforestation, l'agriculture et l'élevage intensif ainsi que par l'exploitation des ressources naturelles de la planète comme le pétrole et les minéraux. Toutes ces actions ont des effets directs sur les communautés biologiques des écosystèmes en perturbant les réseaux trophiques et en entraînant l'érosion de la biodiversité (Walther *et al.* 2002, Rockström *et al.* 2009, Steffen *et al.* 2015, Newbold *et al.* 2016, Figure I.1). Les forçages environnementaux locaux, notamment la pollution des milieux aquatiques (par les eaux usées et les rejets industriels et agricoles), sont listés parmi les facteurs pouvant altérer la stabilité des écosystèmes et leurs propriétés émergentes. De plus, les activités d'origine anthropique sont reconnues pour accélérer le réchauffement climatique planétaire (IPCC 2013, Abram *et al.* 2016). Il semble de plus en plus évident que, sous l'effet de l'Homme, le réchauffement de la surface de la Terre s'est en effet accéléré au cours des trois dernières décennies à tel point que les modèles climatiques les plus désastreux annoncent des effets irréversibles sur les écosystèmes et la démographie humaine (via par exemple le réchauffement des eaux, l'augmentation du niveau moyen des océans ou la multiplicité des événements extrêmes soudains). L'impact du réchauffement climatique sur les milieux aquatiques d'eau douce a été peu évoqué lors des discussions de la COP21 (Conférence de Paris sur le climat 2015) notamment concernant les systèmes lacustres. Pourtant, les lacs sont également impactés par le réchauffement climatique et les activités humaines locales (Williamson *et al.* 2009, Moss 2012) bien que les effets de ces forçages environnementaux sur la biodiversité lacustre soient encore aujourd'hui mal évalués. Afin de comprendre, voire de prédire, l'impact des augmentations de température sur le fonctionnement du lac et ses services écosystémiques, il est nécessaire d'apporter des connaissances sur la réponse biologique des lacs aux multiples forçages environnementaux auxquels ils sont soumis.

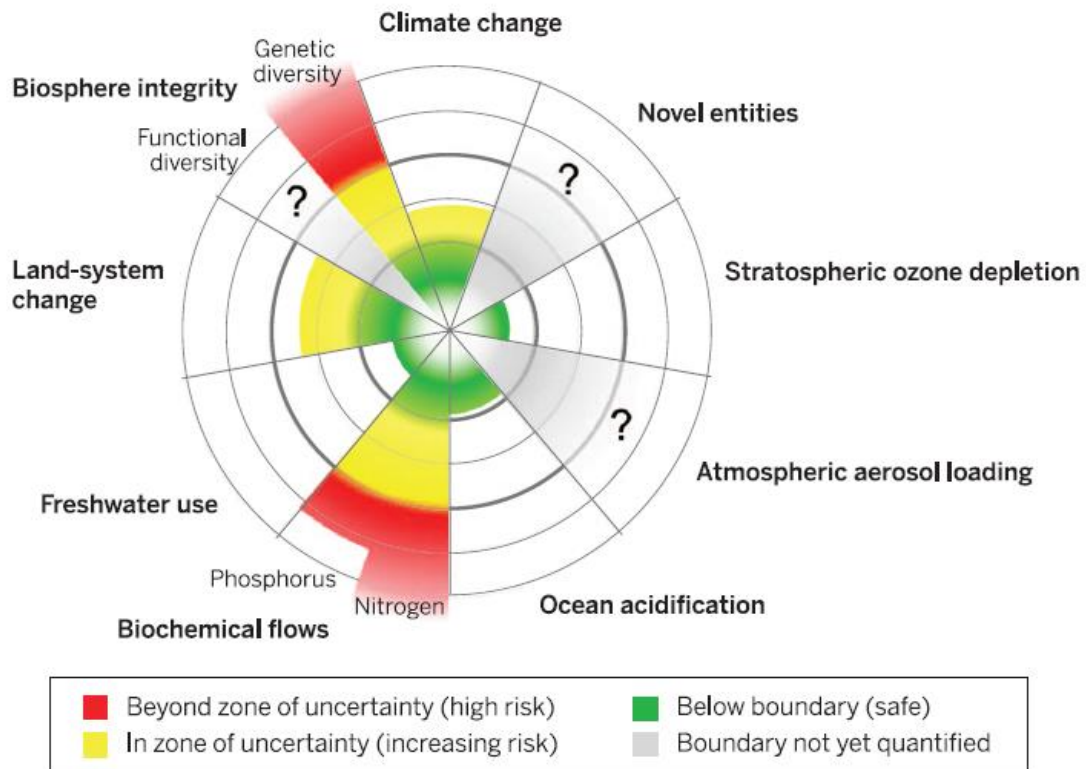


Figure I.1. Schéma illustrant le niveau de risque associé aux effets irréversibles des perturbations humaines sur plusieurs variables importantes pour le fonctionnement des écosystèmes (Figure issue de Steffen *et al.* 2015). Les ressources exploitables sans risques de perturbations sur le fonctionnement des écosystèmes sont représentées dans la zone verte. Dans la zone jaune, le risque que les ressources induisent des perturbations sur les écosystèmes est inconnu alors que les ressources présentes dans la zone rouge ont des forts risques d'induire des effets irréversibles sur les écosystèmes. Enfin, les ressources pour lesquelles un changement a été observé mais dont l'effet n'a pu être associé à un niveau de risque sont représentées par un point d'interrogation.

### I.1.1.2. Quelques rappels sur les lacs

Le nombre de lacs sur Terre est estimé à 304 millions (0,8 % de la superficie de la Terre, Wetzel 2001, Downing *et al.* 2006, Figure I.2). Une grande majorité du volume total d'eau douce à la surface de la Terre est contenue dans moins de 300 lacs avec notamment 20 % du volume total dans le lac Baïkal (Russie). Il existe une très grande variété de formes, superficies et profondeurs caractérisant chaque lac et qui dépendent fortement de leurs origines géologiques (retrait de glaciers, séismes ou activités volcaniques). La plupart des lacs des régions tempérées et boréales sont nés de la dernière déglaciation (il y a environ 25 000 ans); c'est notamment le cas des grands lacs d'Amérique du Nord et des 3 lacs péri-alpins du Sud-est de la France: le lac d'Annecy, le lac du Bourget et le Léman.

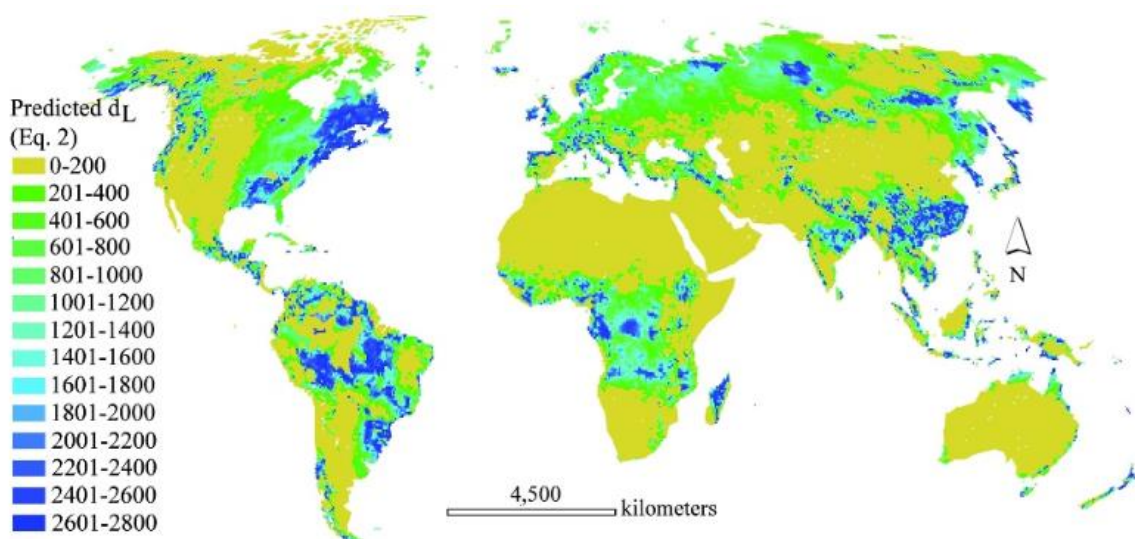


Figure I.2. Distribution des lacs sur la planète. La valeur  $d_L$  correspond à la densité moyenne des lacs par surface entre 1 et 10 km<sup>2</sup> (d'après Downing *et al.* 2006).

L'écosystème lacustre est caractérisé par une variabilité spatiale importante (verticale et horizontale) au niveau de ses conditions physico-chimiques, la composition de ses communautés biologiques et de ses flux énergétiques et nutritifs (Wetzel 2001). La zonation verticale de la communauté, de la colonne d'eau vers le sédiment, est fortement liée à l'éclairement de la masse d'eau et à sa structure thermique (Encadré 1). La zone pélagique est une zone de production d'oxygène et de matière organique biogénique, dominée par les processus autotrophes\*. Ces processus sont réalisés principalement par des microorganismes photosynthétiques, incluant les cyanobactéries et les micro-algues eucaryotes, constituant un compartiment clé du réseau trophique pélagique (Figure I.3). D'autres organismes tels que les ciliés, les cladocères et les rotifères sont des prédateurs\* d'autres organismes (notamment des micro-algues de petite taille et des bactéries). Un certain nombre de parasites\* est également présent au sein du réseau trophique pélagique notamment les chytrides et les perkinsés (Kagami *et al.* 2007, Mangot *et al.* 2011). La matière organique qui sédimente à partir des couches supérieures est recyclée dans les couches plus profondes (processus de décomposition et de re-minéralisation) par des microorganismes hétérotrophes\* ou mixotrophes\* favorisant la respiration microbienne de ces communautés. Le sédiment constitue enfin un compartiment important pour le recyclage et l'archivage au cours du temps de la matière inorganique et organique dont les cellules mortes de la communauté biologique du lac. La zonation horizontale de l'écosystème lacustre est, elle, caractérisée par (i) un gradient d'apport de ressources nutritives et énergétiques du bassin versant vers le lac (plus forte dans la zone littorale que dans la zone pélagique), (ii) la dominance de macrophytes dans la zone littorale, (iii) la présence de remontée d'éléments nutritifs des couches profondes vers les couches supérieures (« upwelling », principalement dans la zone pélagique) et (iv) une exposition similaire de la masse d'eau aux facteurs climatiques externes (vent, pluie, lumière).

La communauté biologique du lac est impactée directement et indirectement par les forçages environnementaux qu'ils soient périodiques (comme les variations saisonnières des conditions environnementales), ponctuels (tempêtes, upwellings) ou continus (réchauffement climatique, eutrophisation) ce qui peut avoir un impact sur les services écosystémiques du lac.

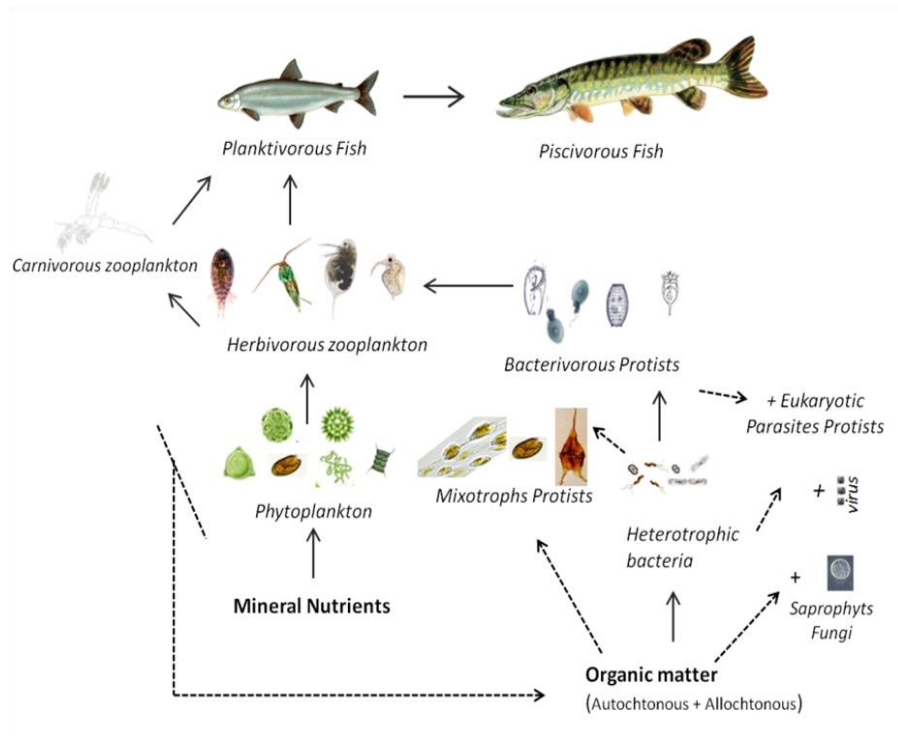


Figure I.3. Exemple d'un réseau trophique pélagique lacustre (Domaizon et collègues)

Encadré I.1. Dynamique annuelle de la structure de la masse d'eau des lacs (Wetzel 2001). La structure de la masse d'eau d'un lac diffère en fonction de la typologie du lac, de son bassin versant et des forçages environnementaux. Une alternance entre périodes de brassage et stratification de la masse d'eau est observée dans la plupart des lacs de régions tempérées. Sous l'action du vent, le brassage des eaux superficielles entraîne l'homogénéisation thermique de la masse d'eau, la redistribution des nutriments et la ré-oxygénation des couches les plus profondes du lac. Une stratification thermique peut apparaître en été et/ou en hiver, phénomène caractérisé par la présence de différentes couches thermiques au sein de la masse d'eau. Par exemple, lors d'une stratification thermique estivale, la masse d'eau est structurée par des couches superficielles (épilimnion) brassées, oxygénées, fortement éclairées et à températures élevées et des couches inférieures (hypolimnion) non brassées, anoxiques et peu ou pas éclairées. Le nombre de lacs monomictiques (une seule stratification par an; exemple d'une stratification estivale pour les lacs Léman, Bourget et Annecy) et dimictiques (deux stratifications) est élevée dans les régions tempérées. Certains lacs sont caractérisés par plusieurs périodes de stratification au cours d'une même saison (lacs polymictiques). Il existe également des lacs méromictiques subissant un brassage incomplet possédant une épaisse couche d'eau stagnante et anoxique appelée monimolimnion (exemple du lac Pavin, France).



### *1.1.1.3. Le lac, sensible aux forçages environnementaux*

Les lacs sont de bons modèles pour étudier la réponse de communautés biologiques aux forçages environnementaux grâce à leur distribution cosmopolite à l'échelle planétaire dans des écorégions aux conditions climatiques variées (polaires, tempérées et tropicales) (Downing *et al.* 2006, Adrian *et al.* 2009, Williamson *et al.* 2009). Dans les études de comparaison multi-lacs, il est toutefois important de prendre en compte la diversité de la morphométrie des lacs et de leurs bassins versants qui déterminent en partie les conditions physico-chimiques et la structure thermique de ces milieux. Parmi les perturbations auxquelles sont soumis les écosystèmes lacustres, l'eutrophisation des eaux (enrichissement excessif en nutriments) et le réchauffement climatique sont les deux facteurs principaux structurant fortement les communautés biologiques (Figure I.4).

L'eutrophisation des eaux, notamment l'enrichissement en phosphore ( $> 20\text{-}30 \mu\text{g.L}^{-1}$ , Downing & McCauley 1992, Carpenter & Bennett 2011), est bien connue pour impacter fortement les communautés pélagiques des lacs entraînant des modifications dans la structure des réseaux trophiques (Jeppesen *et al.* 2000, Moss 2012). Un des effets de l'eutrophisation est l'apparition d'efflorescences algales (appelées aussi « blooms », Smith 2003, Kong & Gao 2005, Paerl *et al.* 2011). Ces blooms, souvent soudains et temporaires, ont des conséquences sur la physico-chimie des lacs, via l'augmentation de la turbidité, la diminution de son éclaircissement et la désoxygénation des couches d'eaux profondes. Ainsi, la communauté biologique de la masse d'eau du lac peut être marquée par la transition d'un lac à eau claire avec de nombreux macrophytes vers un lac à eau turbide dominée par des populations de micro-algues très abondantes (Jeppesen *et al.* 2000, Mooij *et al.* 2005). Les micro-algues (cyanobactéries et algues eucaryotes) impliquées dans ces blooms peuvent également produire des toxines engendrant la mort d'espèces zooplanctoniques et piscicoles. Une étude intégrant différents compartiments biologiques de 71 lacs peu profonds a par exemple mis en évidence les changements dans la structure trophique, la richesse spécifique et la diversité biologique de la communauté pélagique en lien avec les modifications des concentrations en phosphore (Jeppesen *et al.* 2000). Ainsi, alors que la richesse des espèces zooplanctoniques et des macrophytes submergés diminue graduellement avec l'augmentation des concentrations en phosphore, une relation unimodale a été mise en évidence pour la richesse du phytoplancton, des macrophytes flottants et de la population piscicole, cette richesse étant maximale entre 100 et 400  $\mu\text{gP.L}^{-1}$  puis diminuant pour des concentrations en phosphore plus élevées (Jeppesen *et al.* 2000).

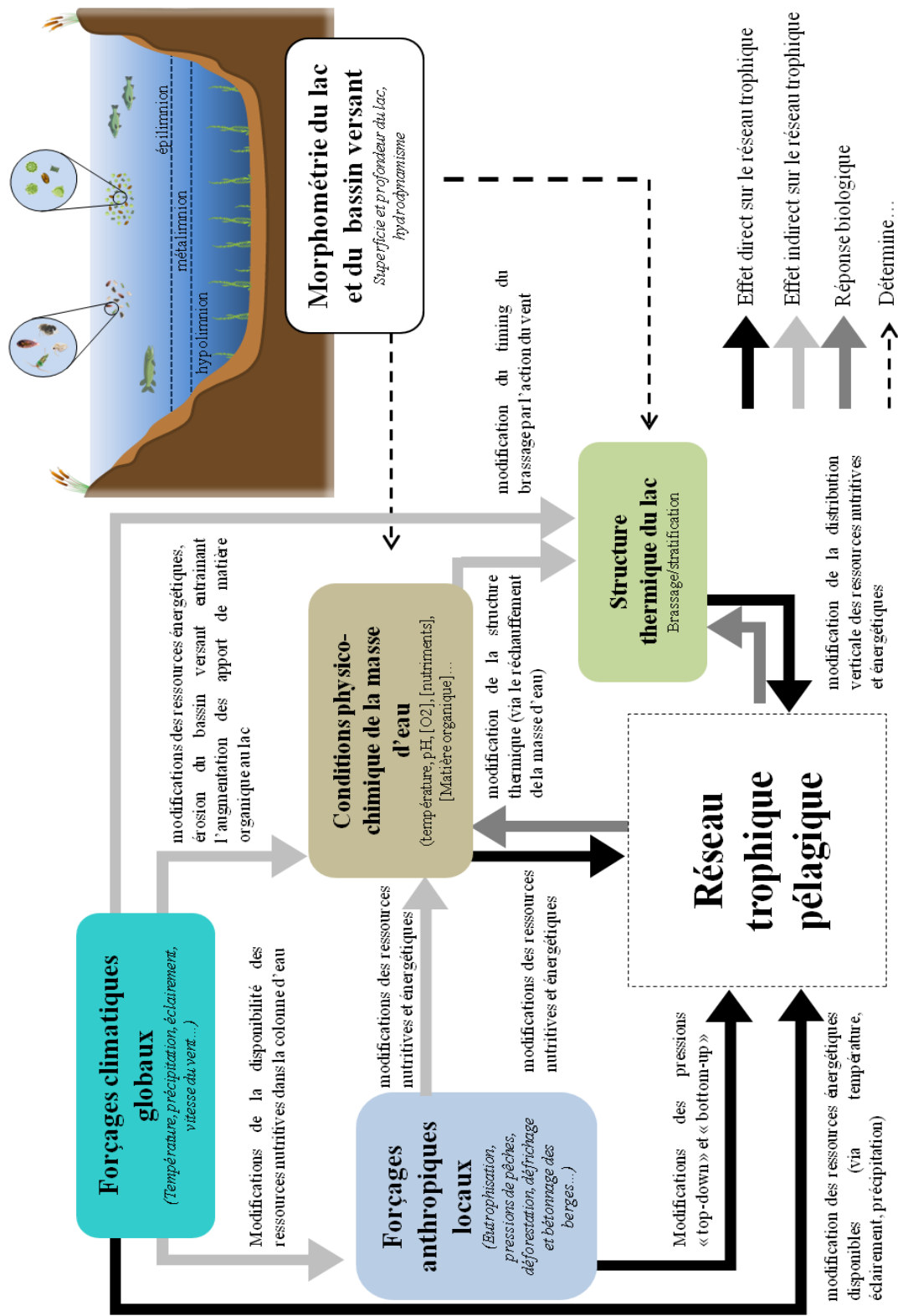


Figure I.4. Schéma illustrant les effets directs et indirects des forçages environnementaux globaux et locaux sur le réseau trophique pélagique lacustre

Le climat, notion intégrant toutes les conditions atmosphériques (température, précipitations, éclairage, pression atmosphérique, humidité et vitesse du vent) a également un fort impact sur les caractéristiques physico-chimiques et biologiques des milieux lacustres (Williamson *et al.* 2009, Moss 2012, Figure **L4**). Ainsi, le réchauffement climatique global peut augmenter significativement la température des eaux superficielles via l'accroissement de l'éclairage dans la masse d'eau (Wetzel 2001). De plus, les caractéristiques chimiques du lac telles que la concentration en nutriments, le pH ou la conductivité peuvent être modifiées par le réchauffement de la masse d'eau. Un des effets indirects du réchauffement climatique sur la physico-chimie du lac est le lessivage de matière organique vers le lac (via l'érosion du bassin versant), rendant la masse d'eau turbide et modifiant ainsi la qualité et la quantité de la matière organique dans la masse d'eau (Chipman *et al.* 2016). La réponse du compartiment biologique au climat semble plus complexe que pour les caractéristiques physico-chimiques car elle dépend à la fois de caractéristiques abiotiques (comme la disponibilité en ressources nutritives et énergétiques) et biotiques (la prédation, le parasitisme ou la compétition pour la ressource...). Le réchauffement climatique est connu pour structurer la communauté biologique du lac de façon directe notamment par l'apport de rayons UV (délétères pour une partie des organismes) et de façon indirecte via par exemple l'impact sur la structure thermique du lac (réchauffement des eaux superficielles entraînant une stratification de la masse d'eau). A partir d'études réalisées sur plusieurs lacs localisés dans des écorégions contrastées (Amérique du Nord, Afrique, Asie du Sud-est, Europe), Paerl & Huisman (2008) suggèrent en effet que le réchauffement climatique est un catalyseur pour les blooms algaux (notamment cyanobactériens) entraînant une plus forte stratification thermique et une plus grande absorption d'UV par les microorganismes phototrophes. De nombreux lacs des régions anthropisées ont subi une eutrophisation au cours du XX<sup>ème</sup> siècle et sont aujourd'hui en phase de restauration grâce aux moyens mis en œuvre par les gestionnaires des lacs. Cependant, malgré la diminution des intrants en nutriments dans ces écosystèmes, les communautés biologiques restent différentes de celles observées avant l'eutrophisation (Jeppesen *et al.* 2005, Battarbee *et al.* 2005, Sondergaard *et al.* 2007). Cela peut être en partie expliqué par l'apparition d'effets couplés des multiples forçages (effets synergétiques ou antagonistes) pouvant impacter les communautés biologiques, en particulier la modification des concentrations en nutriments ou le réchauffement de la masse d'eau (Gallina *et al.* 2013, Rigosi *et al.* 2014). Le retour des communautés à leur état pré-eutrophisation nécessite peut être également un temps plus long que les périodes de re-oligotrophisation considérées dans les études actuelles (généralement 1 ou 2 décennies).

#### *I.1.1.4 L'archive sédimentaire, l'opportunité d'une fenêtre temporelle pertinente pour l'étude de la réponse des communautés biologiques aux forçages environnementaux*

Le sédiment représente l'archive naturelle d'un lac car il intègre les changements d'état de la dynamique temporelle de ces communautés biologiques via la préservation de marqueurs géochimiques et biologiques (Moss *et al.* 2003, Smol *et al.* 2005, Smol 2008, Leavitt *et al.* 2009, Davidson & Jeppesen 2013). A partir de sédiments aquatiques, plusieurs types de biomarqueurs peuvent être utilisés pour retracer des changements au niveau des populations et des communautés notamment les restes d'organismes zooplanctoniques (carapaces de cladocères et de copépodes, capsules céphaliques de chironomes...) et de micro-algues (frustules de diatomées, cystes de chrysophycées...) (Raghukumar *et al.* 2004, Smol *et al.* 2005, Millet *et al.* 2009, 2010, 2014, Gauthier *et al.* 2010, Alric & Perga 2011, Alric *et al.* 2013, Berthon *et al.* 2013, 2014). En parallèle, plusieurs études utilisent les pigments, les lipides et l'ADN comme biomarqueurs (Leavitt & Hodgson 2002, Coolen *et al.* 2009, Battarbee *et al.* 2012, Randlett *et al.* 2014). L'ADN sédimentaire est utilisé depuis une dizaine d'années pour retracer la dynamique temporelle de communautés microbiennes (Coolen *et al.* 2008, Hou *et al.* 2014, Epp *et al.* 2015, Stoof-Leichsenring *et al.* 2015, Table **I.1** pour synthèse des études sur l'ADN sédimentaire en milieu lacustre). Cette approche, dite de paléoécologie moléculaire, permet d'étudier les changements dans le compartiment microbien, au niveau de la population (Domaizon *et al.* 2013, Stoof-Leichsenring *et al.* 2015) et de la communauté (Coolen *et al.* 2013), à des échelles de temps largement supérieures (des dizaines à des milliers d'années) à celles classiquement étudiées en écologie microbienne.

Les études rétrospectives ont permis d'acquérir de nombreuses connaissances concernant les modifications physico-chimiques et biologiques des écosystèmes passés. Ainsi, des travaux ont mis en évidence l'extinction il y a 50 000 ans de la mégafaune (mammifères et oiseaux) (Barnosky *et al.* 2004) quand d'autres ont révélé les larges oscillations climatiques au cours de plusieurs millénaires notamment le réchauffement pendant la période médiévale (Mann 2003, Kaufman *et al.* 2009, Büntgen *et al.* 2011, Mckay & Kaufman 2014). Traditionnellement développés pour l'étude des écosystèmes sur des échelles de plusieurs millénaires, millions ou milliards d'années, de nombreuses études rétrospectives s'intéressent de plus en plus à l'Anthropocène et l'impact des activités humaines sur les écosystèmes depuis le début de l'ère industrielle (fin 19<sup>ème</sup> siècle). L'archive sédimentaire des milieux lacustres a ainsi été utilisée pour révéler l'acidification des lacs d'Europe du Nord et d'Amérique du Nord dans les années 1980s (Battarbee *et al.* 1999). De nombreuses problématiques concernant le fonctionnement des lacs sont encore non résolues et

bénéficieront des apports de la paléocéologie (Seddon *et al.* 2014) notamment concernant les changements à long terme de la biodiversité lacustre, des réarrangements au sein des communautés biologiques et des facteurs qui peuvent être à l'origine de ces modifications (implantation de l'Homme dans le bassin versant, augmentation des activités humaines et des pollutions locales, réchauffement climatique global).

La vision par l'archive possède un fort potentiel dans les études rétrospectives mais n'est pas parfaite car les caractéristiques physico-chimiques et les traces du vivant peuvent être modifiées à plusieurs étapes du processus d'archivage. Ainsi une dégradation des teneurs en carbone et azote a été mise en évidence au cours des 5 premières années d'archivage dans les sédiments (Gälman *et al.* 2008, Rydberg & Martinez-Cortizas 2014). Concernant les communautés biologiques, l'archivage des restes d'organismes est connu pour être impacté par des processus taphonomiques définis comme les processus affectant la préservation des fossiles, de leur formation jusqu'à leur découverte (Behrensmeyer & Kidwell 1985, Alric & Perga 2011, Behrensmeyer *et al.* 2000). Le signal ADN provenant de la communauté biologique de la masse d'eau peut également être modifié : pendant la sédimentation dans la masse d'eau, à l'interface eau-sédiment et au cours de l'enfouissement dans les sédiments. La forme sous laquelle l'ADN arrive dans les sédiments (intracellulaire ou extracellulaire) a un impact sur sa dégradation potentielle car l'ADN intracellulaire est protégé au sein de la cellule alors que l'ADN extracellulaire est soumis à différents processus de dégradation via l'action de nucléases microbiennes, l'hydrolyse (dépuration, déamination, formation de sites sans base) et l'oxydation (Dell'Anno *et al.* 2002, Corinaldesi *et al.* 2005, 2008, Willerslev & Cooper 2005). L'efficacité d'archivage de la communauté biologique dépend également des conditions environnementales au lieu d'archivage. Par exemple, certains lacs possèdent une zone benthique anoxique à faibles températures et pH et exempte de bioturbation permettant une bonne préservation de l'ADN (Coolen & Gibson 2009, Torti *et al.* 2015). Une large part de l'ADN extracellulaire est également préservée de la dégradation grâce à son adsorption sur des minéraux et de la matière organique permettant l'archivage du signal d'une grande partie des organismes ne possédant pas de formes de résistance. Même si plusieurs études moléculaires ont révélé la présence de plusieurs groupes microbiens sur des pas de temps longs (100 ans à 270 000 ans ; Domaizon *et al.* 2013, Hou *et al.* 2014, Randlett *et al.* 2014, Savichtcheva *et al.* 2015, Tableau I.1), on peut cependant constater qu'il existe un manque de connaissances concernant la préservation différentielle de l'ADN des groupes de micro-eucaryotes.

Tableau I.1. Synthèse de la littérature portant sur l'utilisation de l'ADN sédimentaire des milieux lacustres pour l'étude de la dynamique temporelle des communautés microbiennes (modifié à partir du tableau présent dans Domaizon et al. *in rev*, voir Annexe).

Taxons	Régions climatiques	Périodes (en années)	Questions	Forçages environnementaux étudiés	Techniques moléculaires	Références
Micro-eucaryotes	tempérée	Sédiments superficiels	Amplification d'ADN sédimentaire	-	CS (18S rRNA gene)	Takishita <i>et al.</i> 2007
		60	Préservation de l'ADN sédimentaire	-	454 pyrosequencing (V4 18S; ~350 bp)	Capo <i>et al.</i> 2015
	tempérée, polaire	~2 000	Dynamique des communautés	climat, nutriments	Illumina MiSeq sequencing (V7 18S ; ~280 bp)	Capo <i>et al.</i> in press
	tempérée	10 000	Dynamique des communautés	climat, salinité	454 sequencing (V1V3 region of 18S rRNA gene, 563 bp)	Coolen <i>et al.</i> 2013
Phytoplancton	alpine	~ 3000	Dynamique des populations/communautés	climat, salinité, nutriments	PCR-DGGE (23S rDNA ~410 bp) + qPCR (~ 300 bp)	Hou <i>et al.</i> 2014
Phytoplancton	polaire	~18 500	Dynamique des communautés	climat	PCR-DGGE (23S rDNA ~410 bp) + Sequencing	Li <i>et al.</i> 2016
Dinoflagellés		3 000	Dynamique des communautés	climat	PCR-DGGE (18S rRNA gene)	Boere <i>et al.</i> 2009
Diatomées	aride	Sédiments superficiels	Amplification d'ADN sédimentaire, ADN vs analyses microscopiques	-	PCR-DHPLC (SSU gene)	Epp <i>et al.</i> 2011
		200	Dynamique des populations/communautés	climat	CS ( <i>rbcl</i> gene) ~650 - 750 bp	Stoof-Leichsenring <i>et al.</i> 2012
	polaire	~11 500	Dynamique des communautés	climat, salinité	454 pyrosequencing (gene <i>rbcl</i> , ~100 bp)	Epp <i>et al.</i> 2015
Diatomées ( <i>Staurisira</i> sp.)	boréale	7 000	Dynamique des populations	climat, distance géographique	CS ( <i>rbcl</i> gene; ~75 bp)	Stoof-Leichsenring <i>et al.</i> 2015
Haptophytes, Diatomées	polaire	~10 000	Dynamique des populations et des communautés	salinité	PCR-DGGE (498 bp and 556 bp)	Coolen <i>et al.</i> 2004
Haptophytes	tempérée	270 000	Dynamique des communautés	climat	qPCR-DGGE (18S rRNA gene ; ~500 bp)	Randlett <i>et al.</i> 2014
Chytrides (Fungi)	boréale	30	Dynamique des populations, parasitisme	climat, nutriments	qPCR (ITS ; ~150 bp)	Kyle <i>et al.</i> 2015a
Bactéries methanotrophiques de type I et II	polaire	~ 10 000	Dynamique des populations	salinité	qPCR-DGGE (specific primers on 16S rRNA gene)	Coolen <i>et al.</i> 2008
Bactéries methanotrophiques	tempérée	~1 500	Détection de processus anoxiques associés au méthane,	climat, nutriments	16S rRNA gene ; qPCR ( ~110 to 200 bp)	Belle <i>et al.</i> 2014, 2015
<i>Bifidobacterium</i> sp., <i>Bacteroides-Prevotella</i>	boréale	~ 3 000	Détection des activités humaines	-	PCR (~300bp, 500 bp, 650bp) and nested PCR	Madeja <i>et al.</i> 2010
Bactéries & Archées	tempérée	~ 11 000	Dynamique des communautés	climat	16S rRNA ; DGGE (V3 region ~250 bp)	Nelson <i>et al.</i> 2007
Cyanobactéries ( <i>Synechococcus</i> sp.)	tempérée	100	Dynamique des populations et des communautés	climat, nutriments	16S rRNA ITS ; (up to ~1000 bp) & qPCR (up to ~300bp)	Domaizon <i>et al.</i> 2013
Cyanobactéries		~200	Dynamique des communautés	climat, utilisation du bassin versant	16S rRNA; qPCR	Pal <i>et al.</i> 2015
Picocyanobactéries ( <i>Synechococcus</i> sp.)	alpine	~ 3 000	Dynamique des populations et des communautés	climat, salinité, nutriments	CS (ITS ; 1100 bp) + qPCR (~288bp)	Hou <i>et al.</i> 2014
Cyanobactéries ( <i>Planktothrix</i> sp.)	tempérée	100	Dynamique des populations et des communautés	climat, nutriments	16S rRNA gene & ITS ; qPCR(up to 300 bp) & CS (up to ~1000 bp)	Savichtcheva <i>et al.</i> 2015
Cyanobactéries ( <i>Planktothrix</i> sp.)		70	Préservation de l'ADN sédimentaire	-	qPCR (16S rRNA ITS ; up to ~300bp)	Savichtcheva <i>et al.</i> 2011
Cyanobactéries ( <i>Planktothrix</i> sp.)	boréale	300	Amplification d'ADN sédimentaire	-	qPCR ( <i>ociB</i> gene ; from 50 to ~380 bp)	Kyle <i>et al.</i> 2015b
<i>Planktothrix</i> sp. (cyanobacteria)	boréale	30	Dynamique des populations, parasitisme	-	qPCR ( <i>ociB</i> gene ; from 50 to ~380 bp)	Kyle <i>et al.</i> 2015a

D'autre part, l'utilisation de données issues du séquençage de l'ADN nécessite de connaître et prendre en compte les limites et biais notamment liés aux choix de la portion de gène ciblée (pouvant impacter le reflet de la communauté obtenue ; pour exemple concernant les communautés de micro-eucaryotes lacustres voir Taïb *et al.* (2013) et la base de données de séquences utilisée pour l'affiliation taxonomique des séquences nucléotidiques. Le séquençage à haut débit de l'ADN environnemental est un outil au fort potentiel pour étudier la diversité microbienne d'un écosystème. Il faut cependant également prendre en compte ses limites notamment liées à la méthode de séquençage utilisée, la profondeur de séquençage de chaque échantillon, la stringence des analyses bioinformatiques et la méthode/le seuil de regroupement des séquences d'ADN en unités opérationnelles taxonomiques (OTUs) (McElroy *et al.* 2014, Schmidt *et al.* 2014, Esling *et al.* 2015, Majaneva *et al.* 2015, Kopylova *et al.* 2016, D'Amore *et al.* 2016). Ces connaissances sont également indispensables dans les études de paléoécologie moléculaire basées sur le séquençage à haut débit.

L'approche paléoécologique a été récemment appliquée à trois lacs péri-alpins (lac du Bourget, lac d'Annecy et le Léman). Ces lacs, monomictiques et chauds, situés dans le Sud-est de la France, possèdent des caractéristiques communes (en termes de géomorphologie, histoire humaine et exposition climatique) et font l'objet d'un suivi de leur état écologique depuis le début des années 1970s par les gestionnaires des lacs et les scientifiques de la station INRA de Thonon-les-Bains (France) et également de l'institut Forel (Suisse) (Anneville *et al.* 2004, Jacquet *et al.* 2014, Perga *et al.* 2015). Les travaux de paléoécologie initiés au sein du projet ANR IPER-RETRO ([http://www6.inra.fr/iper\\_retro](http://www6.inra.fr/iper_retro)) ont permis d'étudier la réponse des communautés biologiques de ces lacs aux forçages anthropiques et climatiques au cours du dernier siècle. La stratégie appliquée a été de cibler des groupes fonctionnels spécifiques au sein du réseau trophique notamment parmi le pico- (*Synechococcus*), le nano- et le micro-phytoplancton (cyanobactéries et diatomées) et le zooplancton (cladocères) (Perga *et al.* 2015). Des outils classiques (comptages de carapaces de cladocères et frustules de diatomées) et moléculaires (quantification de l'abondance des cyanobactéries notamment l'espèce *Planktothrix rubescens* et des pico-cyanobactéries par qPCR et clonage-séquençage) ont été utilisés pour étudier l'impact du climat et de l'eutrophisation sur la communauté biologique lacustre. Les résultats des différentes études confirment que l'eutrophisation est un facteur clé dans la régulation des communautés aquatiques (ici diatomées, cyanobactéries et cladocères des 3 lacs) et que le statut trophique d'un lac a un impact sur la sensibilité des diatomées et des cladocères aux forçages climatiques. Les effets du climat sur les diatomées et les

cladocères semblent moins forts dans le lac Annecy, plus limité en nutriments, que dans les lacs du Bourget et Léman où l'eutrophisation a été plus marquée (Alric *et al.* 2013, Berthon *et al.* 2014). L'anthropisation du lac Bourget est le facteur le plus fort dans la structuration de la communauté cyanobactérienne du lac du Bourget avec une dominance des genres *Anabaena* et *Microcystis* en conditions eutrophes puis un changement de dominance en faveur de *Planktothrix rubescens* et des pico-cyanobactéries en conditions mésotrophes (Domaizon *et al.* 2013, Savichtcheva *et al.* 2015). Les effets du réchauffement climatique sur le compartiment cyanobactérien semblent être secondaires notamment au regard de la dynamique temporelle de *Planktothrix rubescens*. Afin de compléter les données acquises lors du programme IPER-RETRO, basées sur l'étude des groupes fonctionnels spécifiques, une approche intégrant des compartiments biologiques possédant différents niveaux trophiques permettrait d'étudier l'impact des forçages environnementaux sur le réseau trophique. La diversité génétique des communautés microbiennes lacustres est aujourd'hui relativement bien décrite notamment au sein des assemblages eucaryotes (Lepère *et al.* 2013, Debroas *et al.* 2015, Filker *et al.* 2016). Les questions d'écologie microbienne bénéficieront de la complémentarité des approches de limnologie et paléolimnologie notamment concernant la réponse des communautés microbiennes aux forçages environnementaux.



### **I.1.2. Apports de la paléocéologie moléculaire aux questions d'écologie microbienne**

L'approche de paléocéologie moléculaire est utile pour répondre aux questions de limnologie notamment concernant la variabilité génétique au sein de populations sur de longues périodes ou encore la coévolution hôte/parasite au cours du temps. Le potentiel de l'ADN sédimentaire comme outil rétrospectif pour répondre aux questions de limnologie est décrit plus largement dans l'Annexe (partie VIII). Cette partie illustre ici plus particulièrement la contribution de la paléocéologie moléculaire aux questions d'écologie microbienne concernant la description des communautés microbiennes et leurs réponses aux forçages environnementaux.

La recherche en écologie a été longtemps dédiée à l'étude des macro-organismes, leur abondance au sein des écosystèmes et l'influence des conditions environnementales sur leur dispersion et répartition géographique. Du fait de leurs tailles microscopiques et des moyens techniques disponibles pour les étudier, les micro-organismes n'ont été que tardivement inclus dans les études d'écologie au niveau de l'écosystème, le domaine de l'écologie microbienne s'étant fortement développé depuis le début des années 1960s (Bertrand *et al.* 2011). Ainsi, les recherches en écologie microbienne ont tenté de transposer les concepts d'écologie animale ou végétale au modèle microbien. Cette transposition demande différents ajustements, notamment en lien avec la notion d'espèce biologique ou l'échelle spatiale et temporelle à prendre en compte dans les études, et de coupler des données concernant les micro et macro-organismes afin de comprendre la dynamique des écosystèmes et leur réponse aux forçages (Jessup *et al.* 2004, Prosser *et al.* 2007, Barberán *et al.* 2014). Malgré les avancées dans les outils utilisés en écologie microbienne pour caractériser la dynamique spatio-temporelle des communautés microbiennes, basées sur des paramètres morphologiques, génétiques et métaboliques, de nombreuses questions restent aujourd'hui non résolues (Nemergut *et al.* 2014, Shade & Gilbert 2015, Nesme *et al.* 2016). Une partie des problématiques est commune avec les questions concernant les macro-organismes tels que : Quels microorganismes sont présents dans les écosystèmes ? À quelles abondances ? Quelles sont les interactions entre microorganismes ? Comment réagissent-ils aux changements environnementaux du système ? Certaines questions sont plus spécifiques au compartiment microbien notamment : Comment les communautés microbiennes changent au cours du temps, à quelles échelles temporelles ? Quelle est la capacité de dispersion des microorganismes dans les écosystèmes ? Quels sont leurs fonctions dans les écosystèmes ? Quels sont leurs rôles dans les processus biogéochimiques ? Quelle est le rôle de la biosphère rare dans la réponse aux changements environnementaux ? Dans cette partie, nous nous

intéressons aux apports potentiels de la paléoécologie moléculaire aux questions d'écologie microbienne portant sur la diversité des micro-eucaryotes, leur capacité de dispersion et leurs rôles dans la réponse de l'écosystème aux perturbations environnementales. La description de la diversité microbienne via le séquençage à haut débit de l'ADN permet d'affiner nos connaissances sur les questions évoquées ci-dessus.

La diversité biologique d'une communauté intégrant selon la définition de Frontier et al. (2008), sa richesse taxonomique (nombre de taxons) et sa composition (la répartition numérique des taxons dans l'inventaire) est une caractéristique à prendre en compte dans les études portant sur l'évaluation de l'état écologique d'un écosystème et de son fonctionnement (Bent & Forney 2008, Haegeman *et al.* 2013, Grattepanche *et al.* 2014, Nemergut *et al.* 2014, Gibbons & Gilbert 2015, Shade 2016). La dynamique temporelle des communautés est un paramètre important pour évaluer la diversité microbienne d'un écosystème. Ainsi, une partie des réarrangements dans les communautés microbiennes (perte et ajout de taxons et changements dans la composition) est liée à des fluctuations périodiques des conditions environnementales pouvant agir sur des pas de temps contrastés (heures, journées, saisons voire années) telles que les cycles jour-nuit d'éclairement, la variation saisonnière des conditions météorologiques ou la modification de l'intensité lumineuse (Gilbert *et al.* 2012, Giovannoni & Vergin 2012, Ottesen *et al.* 2014, Fuhrman *et al.* 2015). Malgré ces changements d'états cycliques, la diversité et la structure des communautés microbiennes restent pourtant relativement stables oscillant autour d'une ligne de base (Shade *et al.* 2012, Cram *et al.* 2015, Fuhrman *et al.* 2015). Ainsi, le prélèvement régulier d'échantillons environnementaux permet d'évaluer la part de réarrangements de ces communautés associée à ce phénomène et de découpler la part de variabilité qui peut être attribuée aux perturbations environnementales. Les recherches en écologie microbienne pourraient bénéficier de la complémentarité avec l'approche de paléoécologie moléculaire. En effet, **l'étude de l'ADN sédimentaire apporte une vision temporelle intégrative des communautés d'un écosystème au cours d'une ou plusieurs années** (Figure I.5.). Les inventaires moléculaires obtenus à partir des strates sédimentaires peuvent être utilisés pour étudier l'impact des forçages environnementaux en masquant la variabilité temporelle de la communauté liés aux fluctuations périodiques (journalières, saisonnières...).

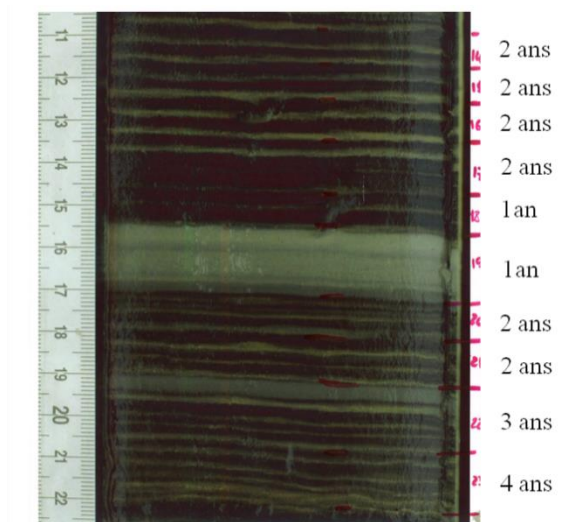


Figure I.5. Illustration des intervalles de temps associés à chaque couche sédimentaire. Carotte sédimentaire prélevée dans le lac du Bourget.

L'étude de la biogéographie des microbes dans les écosystèmes est un sujet dont l'intérêt a été redécouvert avec l'application croissante de techniques de biologie moléculaire pour décrire la diversité microbienne des écosystèmes (Brown *et al.* 2012, Barberán *et al.* 2012, Logares *et al.* 2013, Chow & Suttle 2015, De Vargas *et al.* 2015, Sunagawa *et al.* 2015). Plusieurs hypothèses existent concernant la dispersion des microorganismes à l'échelle planétaire notamment les protistes. Les micro-eucaryotes sont connus pour leur forte capacité de dispersion via le vent (Finlay 2002) ou encore les oiseaux migrateurs (Wilkinson *et al.* 2012). Certains scientifiques pensent que les protistes sont cosmopolites et ont une variabilité limitée même à l'échelle spatiale planétaire (modèle « ubiquiste » de Finlay et Fenchel, Finlay 2002, Finlay & Fenchel 2004). D'autres études penchent en faveur du modèle d'endémicité modéré de Foissner (1999, 2006, 2008) postulant que la capacité de dispersion des organismes et les filtres environnementaux expliquent le niveau d'endémicité des communautés microbiennes. L'utilisation des méthodes moléculaires permet d'apporter de nouvelles connaissances sur la distribution des protistes dans les écosystèmes (Bates *et al.* 2013, de Vargas *et al.* 2015, Filker *et al.* 2016) et permettra probablement de confirmer ou réfuter le dogme « *Everything is everywhere but the environment select* » de Bass Becking, (1934) (de Wit & Bouvier 2006) pour les micro-eucaryotes. En milieu lacustre, une étude récente sur la biogéographie des protistes et champignons dans 13 lacs d'altitude (> 1800 m) a révélé une forte empreinte régionale dans les communautés et une capacité de dispersion inférieure à 10 000 km (Filker *et al.* 2016) confirmant l'effet de distance inter-lacs rapporté par Lepère *et al.* (2013) sur la

richesse de ces communautés à une échelle spatiale régionale. A une échelle plus restreinte, la distribution des micro-eucaryotes semblent être majoritairement impactée par les forçages locaux. Simon et al. (2015) a en effet mis en évidence cette sélection par les conditions environnementales dans des plans d'eau provenant de la même écorégion (bassin parisien, Simon *et al.* 2015). Les patrons de biogéographie des microorganismes sont généralement décrits sans considération de la variabilité temporelle. **L'approche de paléoécologie moléculaire pourrait permettre d'intégrer cette caractéristique temporelle pour la compréhension des patrons de biogéographie des micro-eucaryotes** tels que déjà mis en évidence pour des populations lacustres de *Staurosira* sp. (diatomées) par Stoof-Leichsenring *et al.* (2015).

La réponse des communautés aux perturbations (soudaines ou continues) en termes de stabilité, résistance et résilience n'est aujourd'hui que partiellement comprise. Alors que certaines communautés sont résistantes aux perturbations, d'autres sont affectées et peuvent retourner à leur état initial (résilience) ou non (Figure **I.6**, Allison & Martiny 2008, Strickland *et al.* 2009, Shade *et al.* 2012). Bien qu'une perturbation puisse entraîner un changement compositionnel dans la communauté (gain et perte de taxons, modification dans la répartition numérique des taxons), cette modification n'est pas toujours associée à un changement fonctionnel de la communauté. Si cette communauté à une forte richesse spécifique et un fort niveau d'équitabilité (répartition numérique équilibrée), elle peut rester stable grâce à un niveau élevé de redondance fonctionnelle (Girvan *et al.* 2005, Allison & Martiny 2008, Strickland *et al.* 2009, Wittebolle *et al.* 2009, Shade *et al.* 2012). La forte proportion de taxons rares retrouvés au sein des inventaires moléculaires est également une caractéristique importante de la diversité d'une communauté microbienne (Sogin *et al.* 2006, Pedrós-Alió 2006, 2012, Jones & Lennon 2010, Gibbons *et al.* 2013, Shade *et al.* 2014, Shade & Gilbert 2015) notamment au sein des assemblages eucaryotes (Logares *et al.* 2014, Debroas *et al.* 2015). En effet, ces taxons rares semblent avoir un rôle important dans la réponse de la communauté face aux forçages environnementaux (phénomène étudié particulièrement pour les bactéries; Gilbert *et al.* 2012, Shade *et al.* 2014, Shade & Gilbert 2015). Au sein des assemblages microbiens eucaryotes, la mobilisation de la rare biosphère ne semble pas avoir lieu sur des pas de temps journaliers (Mangot *et al.* 2013) mais plutôt supérieurs (mensuels et saisonniers; Nolte *et al.* 2010, Debroas *et al.* 2015). L'activité biologique de la rare biosphère aquatique, décrite par approche métatranscriptomique (basée sur des ARNr), a également été liée aux forçages environnementaux saisonniers voire aux perturbations continues (Wilhelm

*et al.* 2014, Logares *et al.* 2014, Debroas *et al.* 2015). **Par son ouverture sur des cadres temporels très longs, l'approche de paléoécologie moléculaire met en perspective les recherches sur la dynamique temporelle des communautés microbiennes notamment à propos de leur stabilité, résilience et résistance en réponse à des changements d'état du système.** Ainsi, l'archive sédimentaire permet d'étudier la réponse de communautés biologiques sur des pas de temps cohérents avec le changement d'état du système lacustre par exemple plusieurs dizaines d'années pour l'eutrophisation et la ré-oligotrophisation d'un lac (voir le cas des 3 grands lacs péri-alpins dans la Figure I.7) et plusieurs décennies/centenaires pour les fluctuations climatiques.

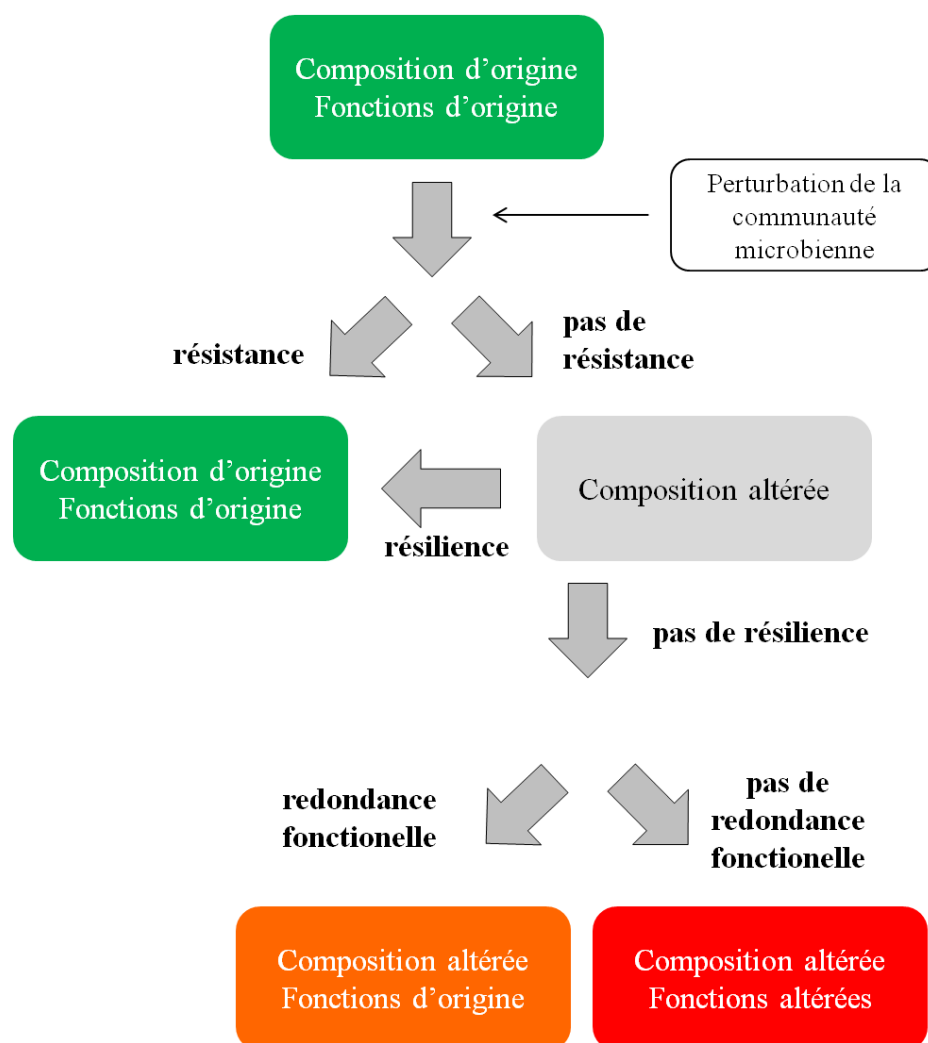


Figure I.6. Schéma de l'effet d'une perturbation sur la composition et les fonctions associées à une communauté microbienne (modifié à partir d'Allison & Martiny 2008).

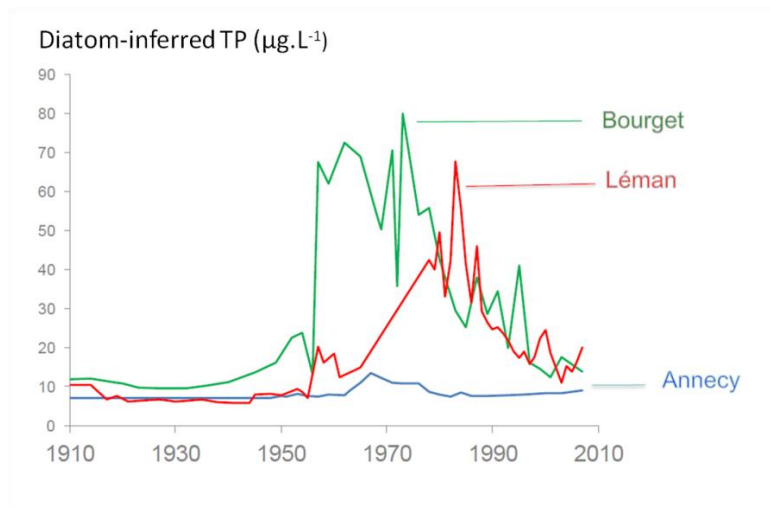


Figure I.7. Concentrations en phosphore estimées à partir du comptage des diatomées dans les archives sédimentaires des lacs d'Annecy, du Bourget et Léman (Alric *et al.* 2013).

D'autres questions d'écologie microbienne peuvent être étudiées grâce à l'approche de paléoécologie moléculaire comme la variabilité et la diversification génétique, la coévolution hôte/parasite ou encore le rôle du compartiment microbien dans les cycles biogéochimiques sur des pas de temps relativement longs (plusieurs décennies à plusieurs millénaires). Ces travaux de thèse s'intègrent dans une démarche de compréhension de la réponse du compartiment biologique face aux forçages locaux anthropiques et globaux climatiques. Plusieurs études ont mis en évidence l'effet des forçages sur les micro-eucaryotes et plus particulièrement sur des groupes spécifiques phytoplanctoniques (Coolen *et al.* 2004, Takishita *et al.* 2007, Boere *et al.* 2009, Epp *et al.* 2011, 2015, Stoof-Leichsenring *et al.* 2012, 2015, Šimek *et al.* 2013, Hou *et al.* 2014, Randlett *et al.* 2014, Li *et al.* 2016). Hou et ses collègues (2014) ont par exemple mis en évidence les changements d'abondance de 3 groupes phytoplanctoniques lacustres (majoritairement représentés par *Synechococcus*, *Isochrysis* et *Gloeotilopsis*) au cours des 3000 dernières années en lien avec les changements temporels dans la disponibilité en nutriments, la température, l'intensité lumineuse et la salinité dans le lac Kusai (Hou *et al.* 2014). Pour intégrer une grande diversité taxonomique et fonctionnelle, la communauté des micro-eucaryotes semble être un bon candidat car elle est composée de prédateurs et proies, de parasites et hôtes ainsi que de producteurs et décomposeurs de matière organique permettant d'étudier l'impact de forçages sur des interactions biotiques multiples au sein du réseau trophique.

### **I.1.3. Les micro-eucaryotes : diversité, rôles et facteurs de régulation dans les milieux lacustres**

#### *I.1.3.1. Du microscope au séquençage à haut-débit de l'ADN*

Bien que le domaine de l'écologie microbienne soit relativement récent (quelques dizaines d'années), la découverte des microorganismes date de plus de 350 ans. Antoni van Leeuwenhoek (1632-1723) fut le premier à découvrir, par hasard, ceux qu'il appela des petits « animalcules » à l'aide d'un microscope rudimentaire. Depuis, l'observation microscopique a permis de décrire plusieurs centaines voire milliers d'espèces microbiennes notamment au sein du plancton lacustre. Bien qu'appliquée depuis plusieurs décennies, notamment pour la bio-indication (la description de la qualité d'un milieu par l'étude de sa composante biologique), l'observation microscopique ne permet pas de révéler la diversité cryptique au sein d'un assemblage biologique. Ainsi, plusieurs espèces microbiennes possédant une morphologie similaire peuvent avoir des caractéristiques génétiques et écologiques différentes. Un exemple marquant est la description de plusieurs variants génétiques de taxons phytoplanctoniques qui possèdent une morphologie similaire au microscope (Xiao *et al.* 2014). De plus, les pico-eucaryotes et pico-cyanobactéries (0.2 - 2  $\mu\text{m}$ ) sont, de par leur taille, difficilement différenciables par microscopie (Worden & Not 2008). Les méthodes culture-dépendantes ne permettent de détecter que des espèces cultivables dans des milieux particuliers et certains microbes sont difficilement cultivables en culture pure car ils dépendent fortement d'un consortium microbien. Ainsi il est établi aujourd'hui, au regard de la diversité microbienne révélée par les outils de biologie moléculaire, que seulement une très faible proportion de la diversité de la communauté microbienne est révélée par la culture en laboratoire de microorganismes (0.01 - 1 %) (Perry *et al.* 2004, Bertrand *et al.* 2011).

Les avancées en biologie moléculaire au cours des 20 dernières années ont permis de compléter les connaissances précédemment acquises concernant la morphologie, le métabolisme et l'écologie des microorganismes en apportant des connaissances sur la diversité taxonomique et fonctionnelle des communautés microbiennes (Bertrand *et al.* 2011). Différentes techniques ont été développées comme l'amplification par PCR, le séquençage Sanger ou les techniques d'empreintes moléculaires pour la caractérisation de la diversité taxonomique, l'étude de la dynamique spatio-temporelle des communautés microbiennes ou encore la mise en évidence de gènes fonctionnels dans les communautés. Le séquençage à haut-débit de l'ADN est aujourd'hui la technologie la plus utilisée en écologie microbienne pour caractériser la diversité microbienne dans des écosystèmes variés, allant des sources

hydrothermales aux intestins humains en passant par les lacs et les fonds marins (Bertrand *et al.* 2011). Plusieurs approches utilisant cette technologie ont été développées pour décrypter la très large diversité taxonomique et fonctionnelle des communautés microbiennes. Ainsi l'**approche métabarcoding** (ou métagénétique) est basée sur le séquençage d'une portion d'un seul gène présent chez tous les microbes d'intérêt (gène codant pour la grande sous-unité ARN) pour décrire la diversité taxonomique d'une communauté. L'**approche métagénomique** correspond au séquençage de tout l'ADN extrait à partir d'un échantillon environnemental et permet ainsi d'obtenir des informations concernant la diversité taxonomique (via la présence de marqueurs phylogénétiques) et fonctionnelle (par exemple des gènes impliqués dans la photosynthèse, la respiration ou dans les cycles de nutriments). L'approche **métatranscriptomique** est basée sur l'extraction de transcrits d'ARNr qui sont des molécules considérées comme reflétant l'activité au sein d'une communauté microbienne. Cette approche nécessite une transformation de l'ARN en ADN complémentaire en vue d'une amplification PCR. Un exemple de la contribution de ces technologies aux connaissances en écologie est la description de la communauté bactérienne des écosystèmes (Eiler *et al.* 2012, Barberán *et al.* 2012, Caporaso *et al.* 2012, Shade *et al.* 2013, Christner *et al.* 2014). Récemment, les données moléculaires provenant du projet Tara Océans ont par exemple permis d'estimer le nombre de procaryotes présents dans les océans à plus de 35 000 OTUs (Unités Taxonomiques Opérationnelles; Sunagawa *et al.* 2015).

Malgré leurs rôles clés dans les réseaux trophiques aquatiques, les micro-eucaryotes ont été globalement moins étudiés que les bactéries et les archées (Caron *et al.* 2009a, Grattepanche *et al.* 2014). Les avancées dans les méthodes de biologie moléculaire ont permis de décrire la diversité des micro-eucaryotes dans la plupart des écosystèmes : les sols (Bates *et al.* 2013), les milieux marins (de Vargas *et al.* 2015, Massana *et al.* 2014, 2015) et les milieux lacustres (Debroas *et al.* 2015, Filker *et al.* 2016, voir Table I.1). Une étude récente a par exemple estimée la richesse planctonique eucaryote océanique à 150 000 taxons (de Vargas *et al.* 2015). L'utilisation d'outils moléculaires, notamment du séquençage à haut débit, a également permis au cours de la dernière décennie d'améliorer nos connaissances concernant la rare biosphère. Les micro-eucaryotes rares ont ainsi été retrouvés comme couvrant la majorité de la richesse des inventaires moléculaires et ils sont potentiellement impliqués dans la réponse de la communauté microbienne aux forçages environnementaux (Jones & Lennon 2010, Logares *et al.* 2014, 2015, Debroas *et al.* 2015).



### *I.1.3.2 Les micro-eucaryotes, un compartiment diversifié*

La définition du groupe des micro-eucaryotes diffère en fonction des travaux. Ainsi certains auteurs incluent des métazoaires dans les micro-eucaryotes (Finlay & Fenchel 2004, Bik *et al.* 2012) quand d'autres en excluent tous les métazoaires et les embryophytes (Lara *et al.* 2015, Kammerlander *et al.* 2015, Triadó-Margarit & Casamayor 2015). Dans cette thèse, les micro-eucaryotes incluent exclusivement les protistes (eucaryotes unicellulaires) et les champignons. Les micro-eucaryotes représentent une large part de la diversité eucaryote connue (plus de 85 % des groupes eucaryotes; Pawlowski *et al.* 2012, Pawlowski 2013, Burki 2014, Figure **I.8**). La classification des microbes au sein des eucaryotes, longtemps basée uniquement sur des caractéristiques morphologiques, a bénéficié des avancées de l'approche phylogénomique qui prend en compte plusieurs dizaines voire plusieurs centaines de gènes. Ainsi, l'arbre phylogénétique des eucaryotes proposé par Burki (2014) s'appuie sur des analyses phylogénomiques, les données génomiques disponibles et les caractéristiques morphologiques des micro-eucaryotes. Elle est considérée comme une des classifications les plus robustes à ce jour. Les eucaryotes sont distribués au sein de sept super-groupes (i) les Opisthocontes (ii) les Amoebozoa (iii) les Excavés (iv) les Archaeplastidés (v) les Rhizariens (vi) les Alvéolés et (vii) les Stramenopiles. Une description succincte de la morphologie et de l'écologie des différents groupes de microbes au sein des eucaryotes est proposée ci-dessous. Cette partie se base sur les informations provenant de plusieurs ouvrages de références (de Reviers 2002, Perry *et al.* 2004, Munn 2011, Bertrand *et al.* 2011).

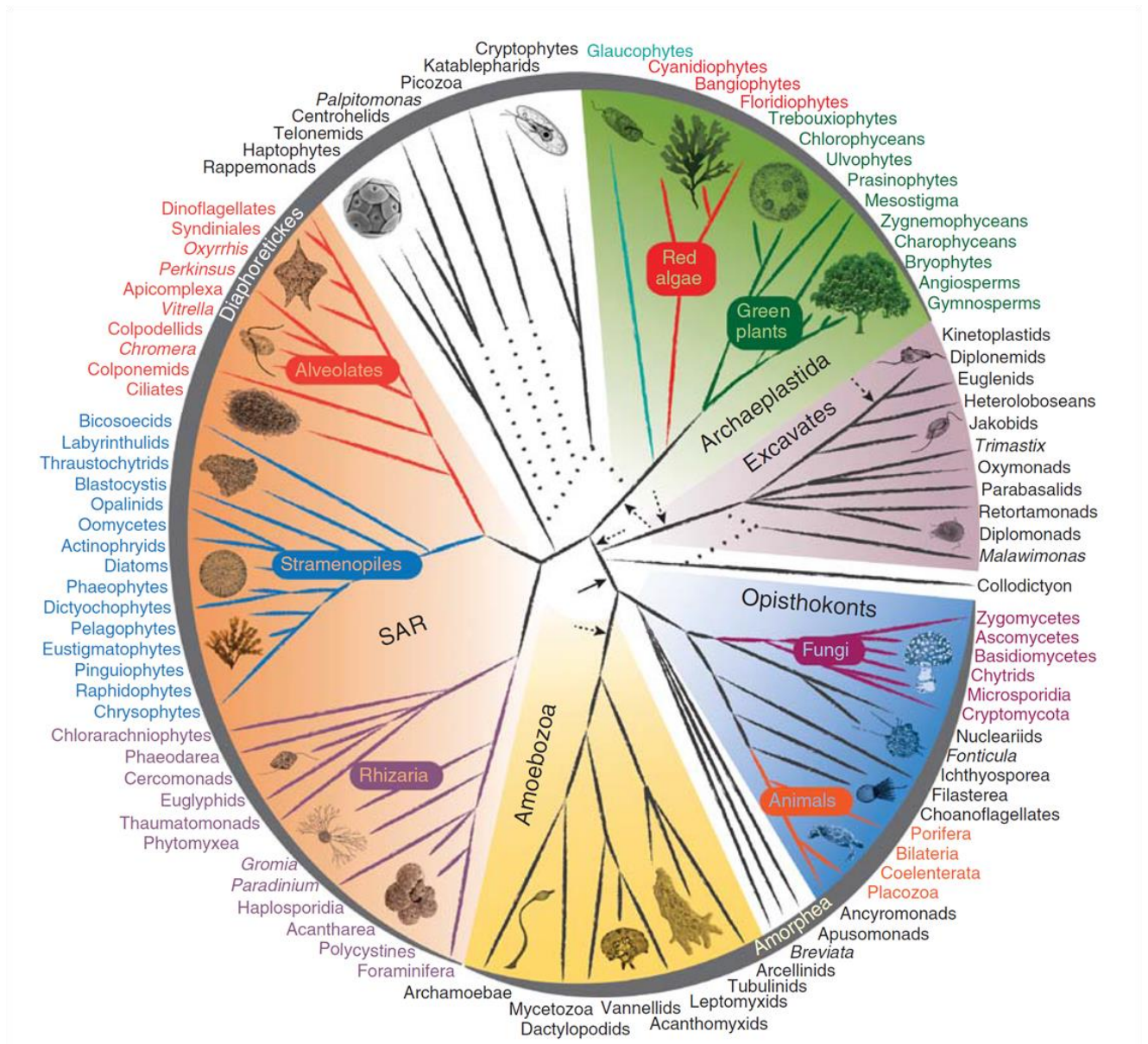


Figure I.8. Arbre phylogénétique des eucaryotes (Burki 2014)

#### (a) Les Opisthocontes (Opisthokonta, Opisthokonts)

Les Opisthocontes sont composés de groupes exclusivement pluricellulaires (les métazoaires, 850 000 espèces), unicellulaires (les choanoflagellés, les nucléarides, les ichtyosporés) ou incluant les deux types comme les champignons. La synapomorphie\* (caractère dérivé partagé par tous les membres du groupe) la plus remarquable des Opisthocontes est la présence d'un seul undulipodium\* (flagelle) présent à l'arrière de la cellule et ayant un rôle propulseur. Les **champignons** sont des hétérotrophes osmotrophes\* qui sont parfois commensaux\* (se nourrissant de la matière organique d'autres organismes sans leur nuire) jouant ainsi un rôle majeur dans la décomposition de la matière organique complexe dans l'environnement. Certains champignons, notamment les chytrides et les microsporidés, sont parasites\* d'organismes photosynthétiques ou non (Lom 2002, Kagami *et al.* 2007, Wolinska *et al.* 2009, Rasconi *et al.* 2012). D'autres agissent comme mutualistes\* et symbiontes\* pouvant notamment former des lichens avec des algues. Le nombre de champignons est estimé à 1 500 000 mais seulement 100 000 morpho-espèces ont été décrites. Les champignons sont présents dans tous les habitats terrestres et aquatiques (benthiques et pélagiques d'eau douce et salée). Les **choanoflagellés** sont des phagotrophes\*, distribués de façon ubiquiste dans les environnements aquatiques. Ils possèdent une coquille extérieure, appelée lorica, et des tentacules leurs permettant de capturer des bactéries.

#### (b) Les Amoebozoa (Ameobobiontes)

Comme les Opisthocontes, les Amoebozoa possèdent un seul undulipodium mais celui-ci se situe à l'avant de la cellule. Cette particularité morphologique, la similarité dans l'architecture génomique mitochondrial et les analyses phylogénomiques attestent de la robustesse de ce groupe taxonomique (Burki 2014). Les **amoebozoa** (plus de 750 espèces décrites) sont des hétérotrophes se nourrissant et se déplaçant à l'aide de pseudopodes\* (déformations de la membrane plasmique) leur conférant un mouvement caractéristique dit amiboïde. Certains amoebozoa sont connus pour brouter les bactéries associées aux particules en sédimentation. Ce groupe est majoritairement composé par les archamoebae (unicellulaires, connus dans les sols) et les mycétobiontes (alternant des générations uni et pluricellulaires, préférant vivre dans les milieux riches en matière organique en décomposition et phagocyter les bactéries, les champignons et la matière organique détritique).

### (c) Les Excavés (Excavates)

La plupart des **Excavés** sont caractérisés par la présence d'hydrogénosomes\* et de mitosomes\*, vestiges de mitochondries. Ils possèdent également un sillon d'alimentation par lequel les particules alimentaires transitent. Couplées aux analyses phylogénomiques, ces caractéristiques morphologiques attestent de la monophylie des Excavés (Moreira *et al.* 2007). Dans ce groupe d'unicellulaires comprenant environ 650 espèces, la plupart des taxons sont des flagellés hétérotrophes parfois commensaux, symbiontes ou parasites et comprennent également des amiboïdes. Les **parabasaliés** sont principalement commensaux et symbiotiques et sont notamment retrouvés dans le tractus digestif des arthropodes et des vertébrés où ils consomment des bactéries. Les eugléniens et les kinétoplastidés (parfois regroupés sous le nom d'Euglenozoa) diffèrent par la présence d'un amas d'ADN à la base de flagelle chez les kinétoplastidés. Les **eugléniens** (ou euglénoïdes) sont particulièrement communs en eau douce enrichie en matière organique et peuvent être phototrophes, phagotrophes ou saprophytes. Les bodonidés, flagellés hétérotrophes à vie libre ingérant des bactéries via un cytospharynx\*, et les Trypanosomatidés se nourrissant de matière organique en solution sont les deux groupes composant les **kinétoplastidés**. Au sein des Excavés, d'autres groupes de protistes flagellés peuvent être retrouvés notamment des amiboflagellés (au sein des hétéroloboséens) et des symbiontes/parasites de métazoaires (parmi les diplomonadines).

### (d) Les Archaeplastidés (Archaeplastida)

Les Archaeplastidés comprennent des organismes unicellulaires et d'autres multicellulaires. Tous les organismes de ce groupe sont caractérisés par leur capacité de photosynthèse acquise par endosymbiose\* primaire (les chloroplastes sont directement dérivés d'une endosymbiose cyanobactérienne). Plus de 300 000 espèces sont décrites au sein du groupe des Viridiplantae qui incluent les **chlorophytes** (toutes unicellulaires, majoritairement en vie libre) et les **streptophytes** (comprenant des organismes unicellulaires et les végétaux supérieurs). Les groupes comportant des espèces pluricellulaires incluent les ulvophycées, les charophycées et les embryophytes. Bien que certains groupes comme les zignématophycées possèdent des organismes à forme uni et pluri-cellulaires, les autres groupes de Viridiplantae, tels que les prasinophycées, les chlorophycées et les trebouxiophycées sont toutes unicellulaires. En milieu marin, les assemblages phytoplanctoniques sont caractérisés par la prédominance de pico-eucaryotes photosynthétiques en particulier au sein des prasinophycées. Les

**glaucophytes** sont des unicellulaires présents principalement en milieu d'eau douce mais généralement peu abondantes. Enfin, les **rhodophytes** sont des algues uni ou pluricellulaires estimées à environ 6 000 espèces, essentiellement retrouvées en milieu marin, les unicellulaires étant majoritairement pélagiques.

(e) Les Rhizariens (Rhizaria)

Le regroupement d'organismes au sein du groupe des Rhizariens est basé exclusivement sur des caractéristiques moléculaires car aucun caractère morphologique partagé n'a pu être établi entre ces différents composants. Parmi les Rhizariens, le groupe des **cercozoaires** est composé d'organismes amiboïdes, flagellés ou amiboflagellés. La plupart sont hétérotrophes mais certains ont acquis la photosynthèse comme les chlorarachniophytes (par endosymbiose secondaire avec une algue verte) quand d'autres sont mixotrophes. Alors que les chlorarachniophytes sont retrouvés en milieu marin, les autres cercozoaires sont plutôt communs en eau douce et dans les sols. Les **radiolaires** sont des prédateurs planctoniques exclusivement marins. Les **foraminifères** sont des organismes amiboïdes possédant des tests percés d'orifices par lesquels s'insèrent les pseudopodes. Ils sont abondants dans le benthos marin mais aussi retrouvés dans le plancton marin et d'eau douce.

(f) Les Alveolés (Alveolata)

Les Alveolés doivent leur nom à la présence de vésicules aplaties situées sous le plasmalemme appelée alvéoles. Le groupe des **ciliés** (Ciliates, Ciliophora) est composé de plus de 10 000 morpho-espèces unicellulaires (Foissner *et al.* 2008). Leur synapomorphie est d'être recouvert de milliers de cils\*. Ces cils sont disposés en files longitudinales (cinéties) ou agglutinés en mèche (cirres) et ont un rôle important dans la capture des aliments. Beaucoup de ciliés possèdent également une « bouche cellulaire » relié à un cytopharynx. Ces organismes sont généralement hétérotrophes (certains mixotrophes) et comprennent de nombreux prédateurs de bactéries. Alors que certains hébergent des organismes photosynthétiques (comme des cyanobactéries et des chlorophytes), d'autres ont la capacité de voler les plastes d'autres organismes (kleptoplastie). Les ciliés sont nombreux dans les milieux continentaux (eaux douces et saumâtres et sols). En milieu aquatique, ils sont connus pour avoir un rôle dans la formation d'efflorescences algales mais aussi comme lien dans le réseau trophique entre les bactéries et les micro-métazoaires. Les ciliés sont des

microorganismes reconnus pour être de bons indicateurs de l'état écologique des milieux aquatiques (Pawlowski *et al.* 2016). Les **dinoflagellés** sont des organismes majoritairement unicellulaires, ayant acquis la photosynthèse par endosymbiose secondaire ou tertiaire. Plus de 2000 espèces de dinoflagellés sont décrites. Les alvéoles des dinoflagellés sont composés de plaques de cellulose constituant la thèque et deux flagelles leur permettant de se déplacer. La majorité des dinoflagellés sont des espèces du plancton marin et d'eau douce, pouvant occasionnellement former des efflorescences (les marées rouges). Environ la moitié sont des organismes photosynthétiques (appelés dinophycées). Dans ce groupe, l'hétérotrophie et la mixotrophie sont des modes trophiques également très répandus. Certaines espèces sont même prédatrices ou parasites. Environ 5000 espèces d'**apicomplexés**, incluant de nombreux parasites intracellulaires, ont été décrites. Ces microbes sont caractérisés par la présence d'un complexe apical constitué d'un conoïde et des vésicules qui sécrètent des enzymes favorisant la pénétration dans les cellules de l'hôte. La plupart des apicomplexés possèdent un apicoplaste, acquis par endosymbiose avec une rhodophyte. Les **perkinsés** forment également un groupe d'unicellulaires parasites retrouvés notamment dans le plancton marin et d'eau douce (Bråte *et al.* 2010, Mangot *et al.* 2011). Ces microorganismes, seulement récemment détectés dans les milieux lacustres peuvent être retrouvés comme très abondants dans les assemblages eucaryotes et sont capables de parasiter des espèces phytoplanctoniques ayant donc un rôle important dans le réseau trophique lacustre (Mangot *et al.* 2011, 2013).

#### (g) Les Stramenopiles (Heterokontes)

Le groupe des Stramenopiles est composé de petits flagellés possédant deux flagelles recouverts de mastigonèmes\* tubulaires tripartites. De nombreux Stramenopiles possèdent une structure en hélice au niveau du corps basal et du flagelle. La monophylie de ce groupe comprenant plus de 100 000 espèces est confortée par les données de phylogénie moléculaire (Burki 2014). Une large diversité de taxons, unicellulaires ou non, est retrouvée au sein de ce groupe comprenant des phototrophes (comme les diatomées) et des hétérotrophes souvent parasites tels que les oomycètes et les bicosoecida. Les phaeophycées, les chrysophycées, les synurophycées et les xanthophycées sont des groupes majoritairement photosynthétiques, présents dans tous les milieux terrestres mais principalement en milieu aquatique. Au sein des Stramenopiles, plusieurs groupes sont composés d'unicellulaires de petite taille (< 3 µm) tels que les bolidophycées, les eustigmatophycées, les pelagophycées ou encore les pinguiophycées. Plus de 10 000 espèces de **diatomées** (Bacillariophyta) ont été décrites et

sont caractérisées par la présence d'une frustule en silice leur conférant une grande protection. Ces algues unicellulaires sont abondantes dans la plupart des habitats aquatiques et ont un rôle important dans la production primaire à l'échelle planétaire. Elles sont largement utilisées comme bioindicateurs de l'état trophique des systèmes aquatiques (Battarbee *et al.* 2001, Smol 2008, Pawlowski *et al.* 2016). Les **chrysophycées** sont retrouvées dans des sols humides et les eaux douces. Un grand nombre de chrysophycées sont connus pour être à la fois phagotrophes et phototrophes. D'autres groupes comme les actinophytra sont des prédateurs d'unicellulaires notamment dans le plancton d'eau douce alors qu'une partie des labyrinthulobiontes et des oomycètes sont connus pour être des parasites dans les milieux aquatiques et terrestres.

#### (h) Autres groupes

La phylogénie de certains groupes reste encore aujourd'hui incertaine notamment les cryptophytes et les haptophytes (Pawlowski 2013, Burki 2014). Environ 500 espèces d'**haptophytes** sont décrites, principalement réparti au sein de deux groupes : les pavlophytes et les prymnésiophytes. La synapomorphie des haptophytes se caractérise par la structure de son appareil cinétique composé de 2 flagelles et d'un haptonème (constitué de 5-8 microtubules). Cet appareil permet la détection d'obstacles, la fixation et la capture de la proie. Bien que la majorité des haptophytes soient photosynthétiques, un certain nombre sont mixotrophes. Ces algues peuvent être à l'origine de blooms notamment dans le plancton marin et d'eau douce (exemple de l'algue *Prymnesium* sp. dans le lac Texoma; Jones *et al.* 2013). Les **cryptophytes** sont tous unicellulaires et ont acquis la photosynthèse par endosymbiose secondaire (avec une Rhodophyte). Ces biflagellés possèdent des mastigonèmes tubulaires dipartites sur leurs 2 flagelles. La plupart des 200 espèces décrites sont photosynthétiques et sont retrouvées dans le plancton marin ou d'eau douce et dans l'eau interstitielle des sols. D'autres groupes d'eucaryotes unicellulaires ont également une position phylogénétique floue dans l'arbre des eucaryotes notamment les raphidomonadés, les katablepharidés, les telonemidés, les centrohelidés et les picrozoaires.

### I.1.3.3. Rôle des micro-eucaryotes dans les réseaux trophiques pélagiques

Les microorganismes sont des acteurs clés du fonctionnement des écosystèmes aquatiques et des flux en éléments et énergie inter-écosystèmes et sont donc largement impliqués dans les cycles biogéochimiques (Falkowski *et al.* 2008). Les micro-eucaryotes aquatiques constituent une partie importante de la diversité biologique des réseaux trophiques pélagiques (Pomeroy 1974, Azam & Graf 1983, Amblard *et al.* 1998, Bertrand *et al.* 2011, Figure I.9). Ils sont notamment connus pour leur diversité fonctionnelle, couvrant une vaste palette trophique, de la phototrophie pure à l'hétérotrophie pure (Figure I.10). Ils peuvent agir comme prédateurs ou parasites d'autres unicellulaires ou des métazoaires (Caron *et al.* 1999, Sherr & Sherr 2002, 2007, Caron *et al.* 2009a, 2009b, Pawlowski *et al.* 2012, de Vargas *et al.* 2015).

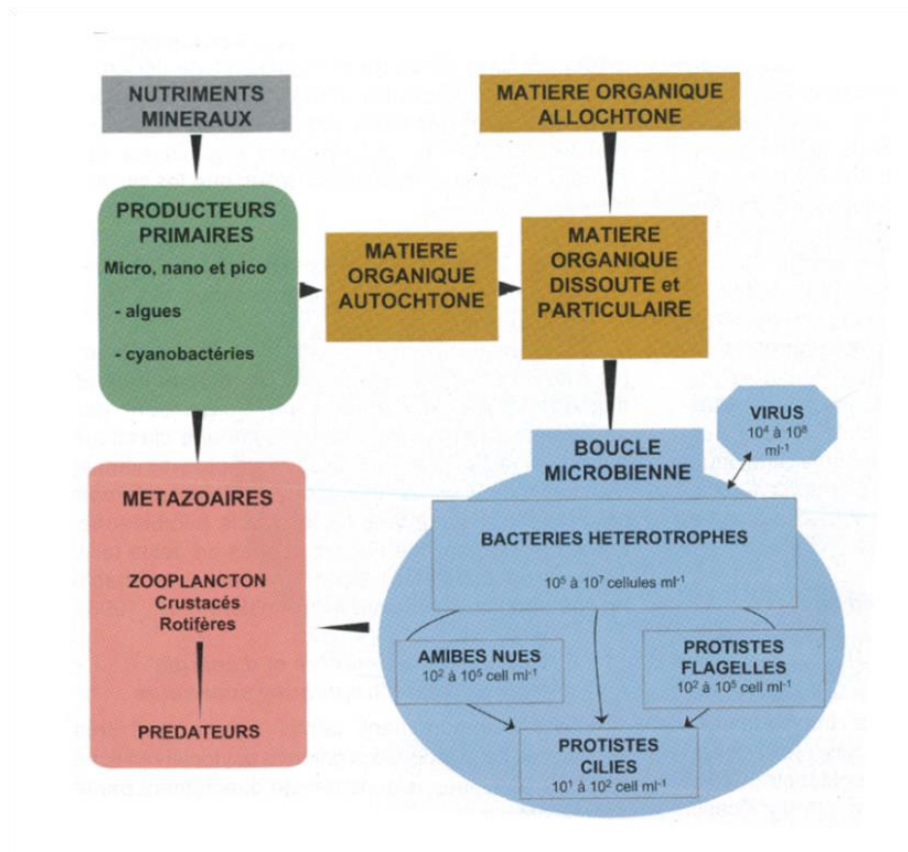


Figure I.9. Schématisation des réseaux trophiques pélagiques lacustres. L'abondance estimée des différents compartiments de la boucle microbienne est indiquée. Cette figure est issue de Amblard *et al.* (1998) redessinée par Bertrand *et al.* (2011).



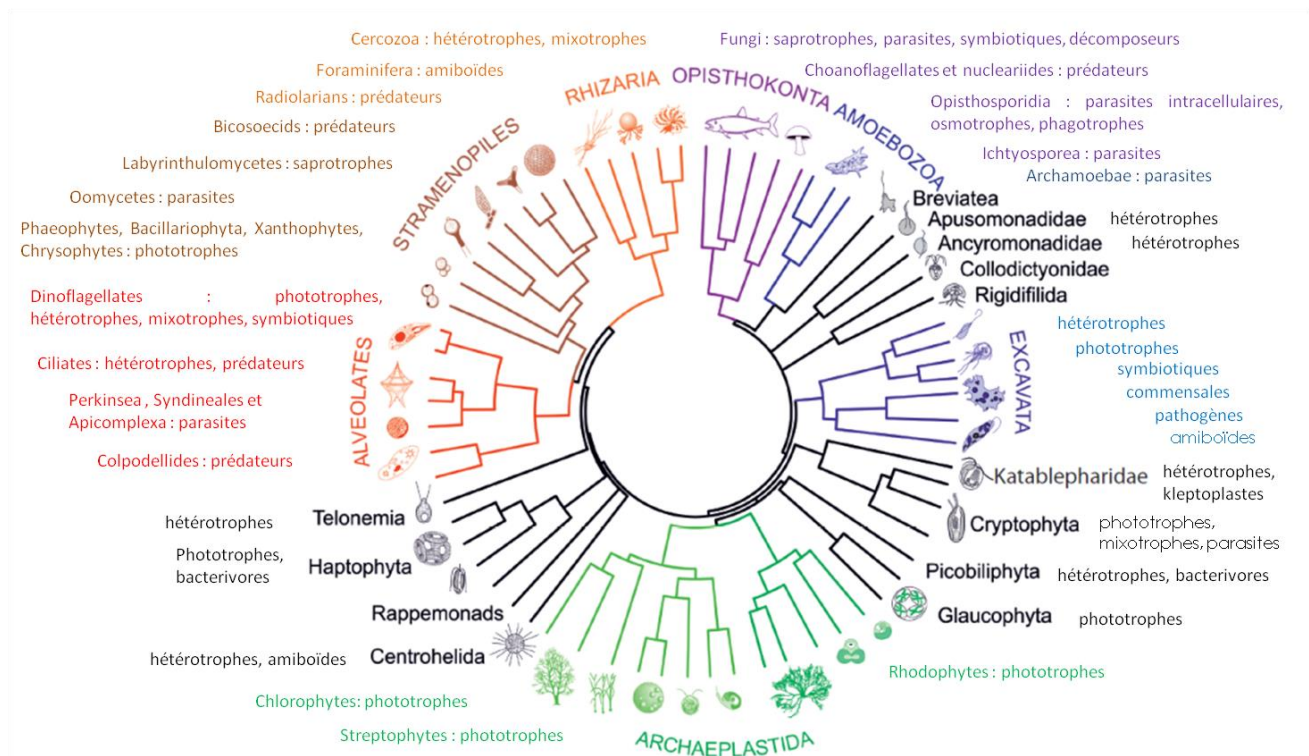


Figure I.10. Assignment de rôles fonctionnels putatifs aux différents groupes de micro-eucaryotes (modifiée à partir de Pawlowski *et al.* 2012).

Les microbes procaryotes et eucaryotes assurent la quasi-totalité de la production de biomasse dans la zone pélagique des milieux aquatiques (Bertrand *et al.* 2011). Les espèces photosynthétiques (cyanobactéries et micro-algues eucaryotes) ont ainsi la capacité de produire de l’oxygène utilisé par les organismes des réseaux trophiques notamment par les organismes composant boucle microbienne (caractérisée par la prédation des bactéries par les amibes, les flagellés et les ciliés). Des interactions hôte-parasite sont également décrites au sein du réseau trophique microbien. Ainsi, le parasitisme du phytoplancton par des Chytrides a été largement étudié au cours des dernières années (Ibelings *et al.* 2007, Kagami *et al.* 2007, Rasconi *et al.* 2012). Au sein des micro-eucaryotes, les perkinsés, les bicosoecida, les apicomplexés et les oomycetes sont également connus pour leur capacité de parasitisme (Lefèvre *et al.* 2008, Mangot *et al.* 2011, Bertrand *et al.* 2011). Le broutage du phytoplancton par les protistes hétérotrophes (dont les ciliés) a été peu étudié en comparaison avec la prédation par le zooplancton (daphnies, copépodes). Cette interaction est pourtant connue depuis longtemps et était déjà intégrée dans le modèle PEG (Plankton Ecology Group) il y a 30 ans par Sommer *et al.* (1986). Cette interaction a depuis été démontrée notamment chez les ciliés dans le cas de blooms phytoplanctoniques en eau douce (Müller 1989, Müller *et al.*

1991, Posch *et al.* 2015, Simek *et al.* 2014). La compréhension du rôle des micro-eucaryotes au sein des réseaux trophiques lacustres bénéficie et bénéficiera des études portant sur la diversité génétique des communautés. Notre étude bibliométrique met en évidence un nombre relativement restreint d'études décrivant la diversité génétique des micro-eucaryotes lacustres en comparaison avec les autres écosystèmes (majoritairement portant sur la santé humaine, animale et les écosystèmes marins, Figure I.11). Les travaux menés, dont une synthèse non exhaustive est présentée dans la Figure I.12, ont permis de mettre en évidence (i) une grande diversité génétique (intra et inter espèces) des micro-eucaryotes des lacs, (ii) la dynamique temporelle des communautés à différentes échelles (journalière, mensuelle, saisonnière), (iii) les activités biologiques de ces microbes, (iv) la dispersion des microbes inter-écosystèmes ainsi que (v) l'effet des forçages environnementaux externes (ajout de nutriments) et internes (prédation, parasitisme) sur la diversité et la composition de ces communautés. Ainsi, la dynamique des communautés de micro-eucaryotes est régulée par différents forçages environnementaux au sein du réseau trophique (variations cycliques et perturbations d'origine anthropique et climatique). Ces perturbations sont classiquement réparties entre des pressions descendantes (ou « top-down », notamment via la prédation par le zooplancton) et des pressions ascendantes (ou « bottom-up », via la modification temporelle des ressources énergétiques et nutritives bio-disponibles). L'action humaine peut exacerber ces deux pressions de plusieurs façons : par une forte pression de pêche (impactant par cascade trophique la prédation zooplanctonique des micro-eucaryotes), la modification des teneurs en nutriments dans le lac engendrant l'eutrophisation et ses effets sur l'augmentation du réchauffement climatique.

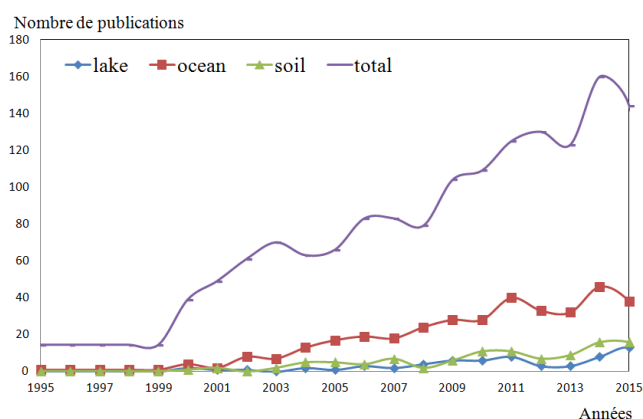


Figure I.11. Recherche bibliographique. Nombre de publications listées dans ISI-Web avec les combinaisons de mots clés suivantes: « protists » ou « microbial eukaryotes » ET « DNA » ET « lake » ou « ocean/marine » ou « soil » (ou rien pour la rubrique total ; par année entre 2000 et 2015 et moyenne sur 20 ans entre 1980 et 1999)

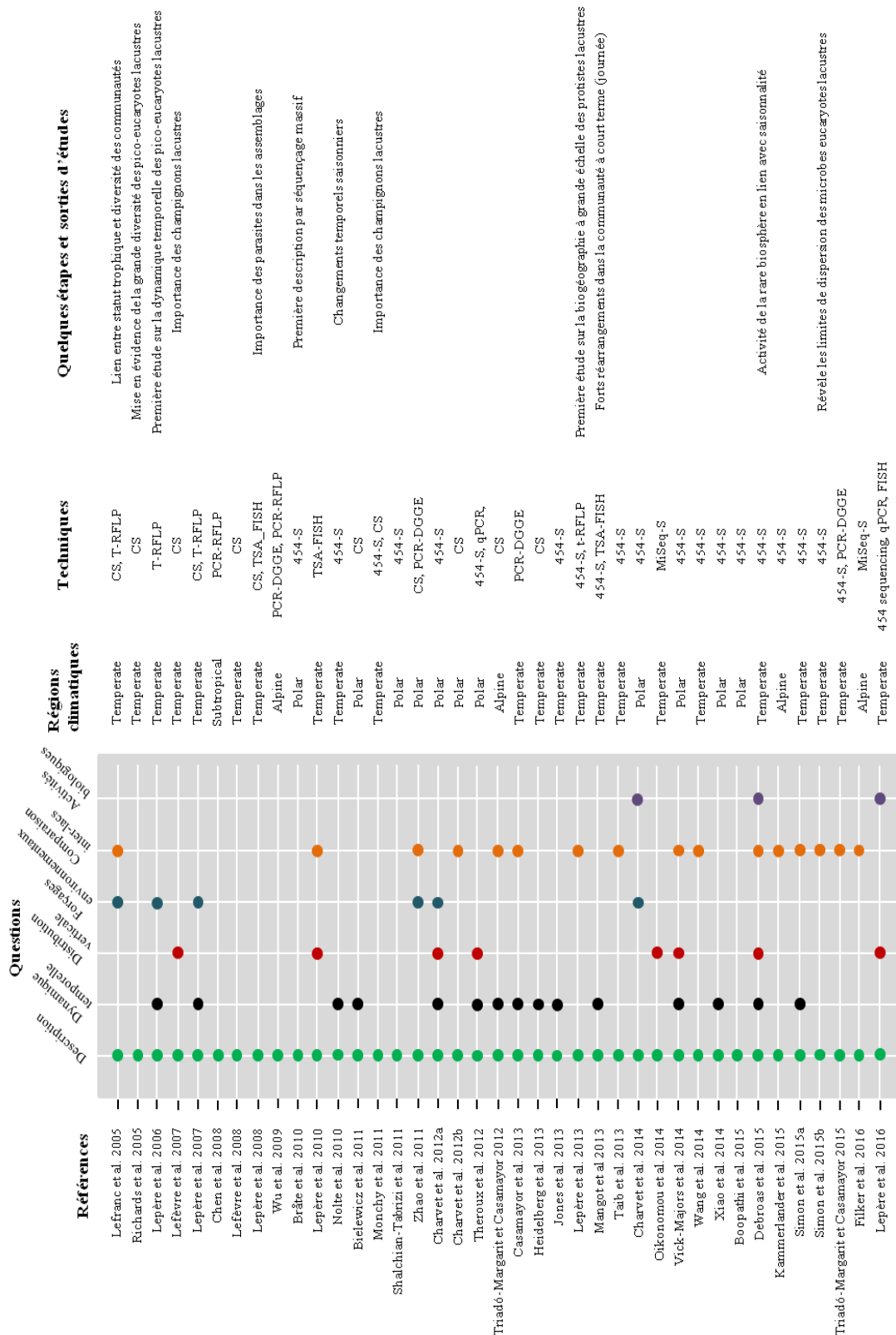


Figure I.12. Liste non exhaustive des principales publications portant sur l'étude des communautés microbiennes eucaryotes lacustres utilisant des analyses de biologie moléculaire, les questions d'intérêt, les régions climatiques étudiées, les techniques moléculaires utilisées et l'apport de ces publications à la compréhension des assemblages de micro-eucaryotes.

#### *I.1.3.4. Etat des connaissances sur la réponse des micro-eucaryotes à l'eutrophisation et aux changements climatiques*

Bien que l'eutrophisation et le réchauffement climatique soient deux facteurs connus pour impacter les communautés biologiques aquatiques et plus généralement le fonctionnement du lac, l'effet de ces forçages sur les assemblages de micro-eucaryotes est encore peu connu.

La plupart des études portant sur les effets de l'eutrophisation des lacs se sont focalisées sur les micro-algues (Jeppesen *et al.* 2000, Kong & Gao 2005, Paerl *et al.* 2011). Certains travaux révèlent également la réponse des communautés bactériennes marines à l'enrichissement en nutriments (Haukka *et al.* 2006, Baltar *et al.* 2015) notamment marquée par l'augmentation de l'abondance et de la production bactérienne (Nogales *et al.* 2011). Cependant les études relatives aux micro-eucaryotes dans leur ensemble (hétérotrophes, phototrophes ou mixotrophes) sont plus rares. Ainsi, l'état trophique des lacs a rarement été mis en relation avec la diversité au sein d'assemblages microbiens eucaryotes (Lefranc *et al.* 2005, Zhao *et al.* 2011) mais jamais sans prendre en compte les changements temporels de la communauté. De plus, bien que certains groupes de protistes sont connus pour être abondants dans les milieux aquatiques en conditions eutrophes tels que les chlorophytes (Naselli-Flores & Barone 2000, Moss *et al.* 2003, Soares *et al.* 2013), les diatomées (Battarbee 2000, Battarbee *et al.* 2002, Smol 2008, Davidson & Jeppesen 2013) ou les ciliés (Beaver & Crisman 1989, Kisand & Zingel 2000, Pawlowski *et al.* 2016), la réponse d'une grande partie des micro-eucaryotes à l'eutrophisation est encore aujourd'hui totalement inconnue.

Les changements climatiques (température, précipitation...) sont également connus pour impacter les communautés biologiques des milieux lacustres provoquant notamment l'apparition de pics de biomasse planctonique, la dominance de cyanobactéries et de zooplanctons de petites tailles ainsi que des activités bactériennes plus élevées (Moss *et al.* 2003, Mooij *et al.* 2005, Jeppesen *et al.* 2005, de Senerpont Domis *et al.* 2013). Mais, à l'instar des études sur l'eutrophisation, les modifications des assemblages concernent quelques groupes taxonomiques (i.e. micro-algues). Les micro-eucaryotes pourraient également être impactés par les effets directs (l'élimination des niches écologique, des processus hétérotrophes plus importants que les processus phototrophes, la sensibilité aux UV) et indirects (via des pressions exercées par le climat sur leurs proies et prédateurs) du changement climatique (Figure **I.13**, Caron & Hutchins 2013). Rose *et al.* (2009) ont notamment montré que l'augmentation de la température de la masse d'eau modifie l'abondance et la composition taxonomique du micro-zooplancton (notamment les ciliés et

dinoflagellés hétérotrophes), ces changements pouvant s'expliquer par des modifications au sein du phytoplancton et donc par de multiples interactions au sein du réseau trophique (Rose *et al.* 2009).

La difficulté à déconvoluer les effets de multiples forçages sur les communautés de micro-eucaryotes explique en grande partie le manque de connaissances concernant l'impact du climat et de l'eutrophisation sur ce compartiment d'autant plus que ces effets peuvent être synergiques ou antagonistes en fonction des groupes/espèces ciblées (Anneville *et al.* 2004, Gallina *et al.* 2013, Rigosi *et al.* 2014, de Senerpont Domis *et al.* 2014). De façon générale, l'augmentation des températures dans un système eutrophisé semble exacerber les effets de l'eutrophisation (Paerl & Huisman 2008). Dans le cas d'un système en phase de re-oligotrophisation, le réchauffement climatique pourrait freiner les effets de la re-oligotrophisation empêchant le retour du réseau trophique à son état de pré-eutrophisation.

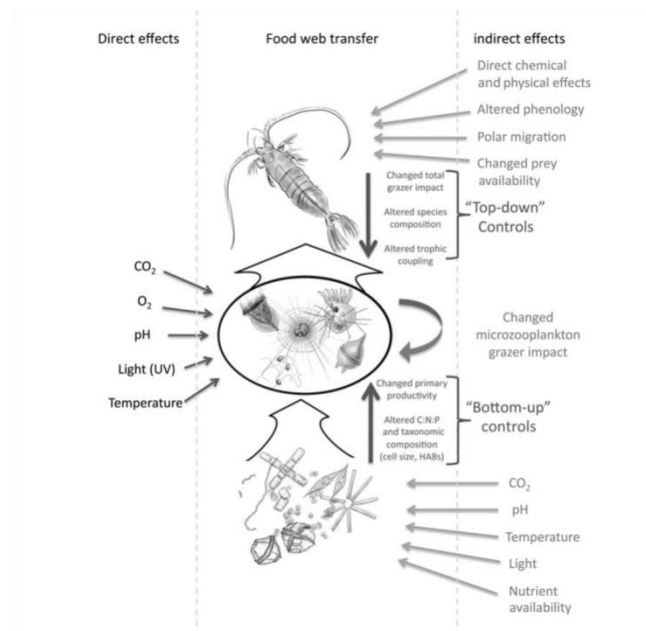


Figure I.13. Schéma général de la position trophique du nano et micro-zooplancton au sein du réseau trophique planctonique et des effets potentiels directs et indirects des changements climatiques sur ces assemblages (issu de Caron & Hutchins 2013).

## **I.2. Objectifs et stratégie des travaux de thèse**

### **I.2.1. Objectifs de la thèse**

L'objectif principal de cette thèse est d'analyser la réponse des communautés microbiennes eucaryotes aux forçages environnementaux. L'originalité de ces travaux vient du cadre temporel utilisé (plusieurs dizaines d'années à plusieurs millénaires) rendu accessible grâce à l'approche de paléoécologie moléculaire. La stratégie appliquée est basée sur le séquençage à haut débit de l'ADN sédimentaire pour retracer les changements dans les communautés microbiennes eucaryotes sur des pas de temps relativement longs. L'application de ce type d'approche (paléolimnologie moléculaire) est encore dans une phase de développement, c'est pourquoi cette thèse inclue certains travaux qui relèvent de la calibration du signal ADN sédimentaire. Deux principaux axes ont été développés au cours de ces travaux de thèse.

#### *I.2.1.1. Axe 1 : Calibration/validation de l'utilisation de l'ADN sédimentaire*

Une problématique importante dans les études de paléoécologie moléculaire est la pertinence de l'utilisation d'ADN sédimentaire comme le reflet des communautés planctoniques et benthiques de l'écosystème. De la sédimentation des cellules mortes jusqu'à son enfouissement, plusieurs événements peuvent empêcher un archivage efficace de l'ADN planctonique dans les sédiments, notamment la dégradation par les microorganismes dans la colonne d'eau, la forme sous laquelle l'ADN atteint le fond du lac (intra ou extracellulaire) et les processus diagénétiques pouvant dégrader l'ADN au cours de l'enfouissement dans les sédiments. Alors que plusieurs études portent sur les limites et les conditions relatives à l'archivage de l'ADN des communautés de la masse d'eau vers les sédiments marins (Corinaldesi *et al.* 2008, Boere *et al.* 2011a, Torti *et al.* 2015) et lacustres (Coolen & Gibson 2009), la qualité de l'archivage en fonction des groupes taxonomiques reste encore à étudier. Les travaux présentés dans les chapitres II et III apportent de nouveaux éclairages concernant la pertinence de l'utilisation de l'ADN sédimentaire lacustre comme reflet des communautés planctoniques.

**Chapitre II** : Efficacité d'archivage de l'ADN des microorganismes eucaryotes de la masse d'eau vers les sédiments du lac du Bourget

Les travaux présentés dans le chapitre II ont pour objectif de vérifier l'efficacité d'archivage de l'ADN de micro-eucaryotes planctoniques de la masse d'eau vers les sédiments. La stratégie d'étude de ce premier volet est de comparer les inventaires moléculaires (séquençage à haut débit) d'ADN extrait de la masse d'eau du lac du Bourget (à 2 m et 130 m) et de sédiments récents (archives des 60 dernières années). Les échantillons de la masse d'eau du lac du Bourget proviennent d'un suivi moléculaire de la communauté microbienne eucaryote (par séquençage à haut débit 454 Roche) au cours de deux ans (prélèvements mensuels). Ce suivi s'intègre dans le suivi de l'état écologique des lacs mené par l'Observatoire des lacs alpins (OLA; <https://www6.inra.fr/soere-ola>). Dans ce projet, nous avons tiré profit de la disponibilité de ces données pour les comparer aux inventaires moléculaires provenant des couches sédimentaires du lac du Bourget. En complément, une étude de transcriptomique, basée sur l'amplification PCR d'ADNc (rétro-transcrits à partir d'ARN) permet de vérifier l'absence/présence d'activité biologique de micro-eucaryotes dans les sédiments. L'efficacité d'archivage de l'ADN de taxons planctoniques peut ainsi être mise en évidence pour de nombreux groupes taxonomiques et fonctionnels de micro-eucaryotes.

**Questions spécifiques** : quels taxons présents dans la masse d'eau à 2 m de profondeur sont également retrouvés à 130 m et par conséquent quels microbes sont les plus susceptibles d'avoir été consommés/lysés au cours de la sédimentation dans la colonne d'eau ? Quelle est la proportion de taxons planctoniques retrouvés dans les archives sédimentaires ? Certains groupes microbiens planctoniques sont-ils sous-représentés dans les archives sédimentaires ? Y a-t-il des taxons actifs dans les sédiments, et si oui : sont-ils dominants dans les inventaires moléculaires?

### **Chapitre III** : Préservation de l'ADN sédimentaire au cours de l'enfouissement récent

L'objectif des travaux présentés dans le chapitre III est de tester l'hypothèse que l'enfouissement dans les sédiments a un impact sur la quantité et la qualité de l'ADN au cours des premières années d'archivage. Cette hypothèse résulte de la mise en évidence par Gälman *et al.* (2008) de la chute des teneurs en carbone et azote dans des sédiments du lac Nylandssjön au cours de 5 années d'archivage (Gälman *et al.* 2008, Rydberg & Martinez-Cortizas 2014). L'approche consiste à comparer les caractéristiques physico-chimiques de strates sédimentaires de carottes prélevées à différentes années (Figure I.14). Dans ces travaux, nous avons voulu tester cette hypothèse concernant la qualité du signal ADN (impact sur la richesse et la composition des communautés microbiennes eucaryotes). Une collaboration avec Christian Bigler et ses collègues (Université d'Umeå, équipe de paléolimnologie) a permis l'application d'analyses moléculaires à une collection de carottes sédimentaires unique au monde (carottes du lac Nylandssjön prélevées et conservées depuis plusieurs années à  $-18\text{ }^{\circ}\text{C}$ ). La stratégie de ces travaux consiste à comparer les inventaires moléculaires de 12 strates sédimentaires (datées entre 1973 et 2006 ; issus de séquençage Illumina MiSeq) provenant de deux carottes prélevées en 2007 et 2013 et conservées par congélation jusqu'aux analyses moléculaires (en 2014-2015).

**Questions spécifiques** : l'enfouissement au cours des premières années a-t-il un impact sur la quantité d'ADN archivé, la richesse, la composition et la structure des communautés microbiennes eucaryotes? Certains groupes de micro-eucaryotes sont-ils sujets à plus de modifications que d'autres ?

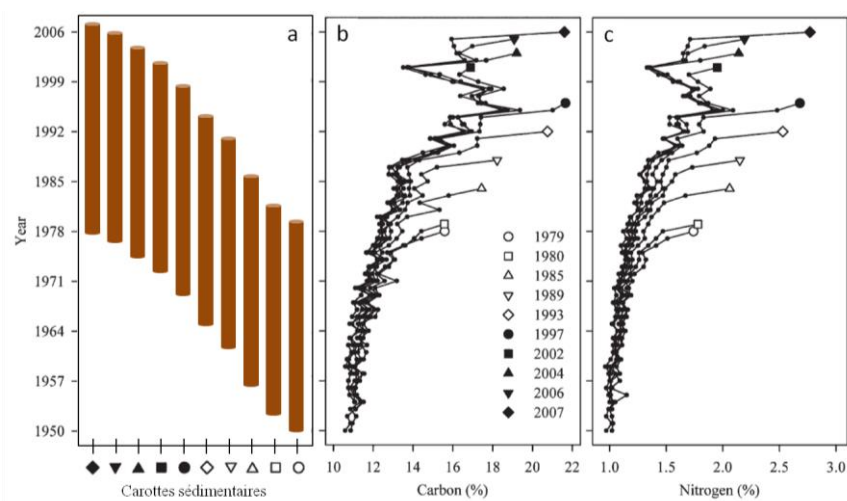


Figure I.14. Schéma illustrant la période temporelle couverte par chacune des 10 carottes sédimentaire prélevées entre 1979 et 2007 (a) et les teneurs en carbone (b) et azote (c) (pourcentage par rapport à la masse de sédiments secs) pour chaque varve individuelle des 10 carottes (modifiée à partir de Gälman *et al.* 2008).



Bien qu'il reste des aspects méthodologiques à valider concernant la préservation de l'ADN dans les sédiments, le niveau d'aboutissement atteint concernant l'application du séquençage à haut débit à de l'ADN sédimentaire permet de déployer cette méthode pour explorer les changements temporels dans les communautés microbiennes eucaryotes.

#### *I.2.1.2. Axe 2 : Dynamique temporelle des communautés microbiennes en réponse aux forçages environnementaux*

Alors que plusieurs études moléculaires ont mis en évidence la réponse de groupes spécifiques microbiens aux forçages environnementaux à partir d'archives sédimentaires lacustres (Coolen *et al.* 2004, Domaizon *et al.* 2013, Hou *et al.* 2014, Randlett *et al.* 2014), il existe un manque de connaissances concernant l'impact des forçages au niveau d'un réseau trophique/écologique complexe. Au sein d'un réseau trophique lacustre, le compartiment microbien eucaryote est intéressant car il intègre de nombreux liens trophiques (proie et prédateur, hôte et parasite, producteurs et consommateurs de matière organique). Grâce aux avancées des techniques de biologie moléculaire, il est aujourd'hui possible de caractériser les changements temporels de richesse et composition de ces assemblages (Mangot *et al.* 2013, Debrias *et al.* 2015) même sur des pas de temps très longs (Coolen *et al.* 2013). Les travaux du second axe s'inscrivent dans cette démarche de compréhension de la dynamique temporelle des assemblages biologiques et de la réponse de ces communautés aux forçages environnementaux. Dans ces travaux, la dynamique temporelle des communautés a été étudiée dans deux lacs au cours des 2200 dernières années (lacs du Bourget et Igaliku, chapitre IV) et, de façon plus résolutive, dans les lacs d'Annecy et du Bourget au cours du siècle dernier (chapitre V) permettant de révéler les effets des forçages climatiques et anthropiques sur des pas de temps contrastés. Le premier projet a pour but d'étudier l'impact des fluctuations climatiques au cours des derniers 2200 ans, notamment les larges oscillations de température marquées par des périodes chaudes (période médiévale du 9 au 13<sup>ème</sup> siècle et période actuelle) et des périodes froides comme le petit âge glaciaire (au cours du 16<sup>ème</sup> siècle) (Büntgen *et al.* 2011). Le second projet vise à révéler les variations temporelles dans les communautés microbiennes eucaryotes sur un pas de temps plus court (100 ans) en lien avec l'accroissement du réchauffement climatique (surtout au cours des 30 dernières années, Auer *et al.* 2007) et la modification des teneurs en phosphore dans deux lacs ayant des états trophiques contrastés. Le chapitre V apporte, par rapport au chapitre IV, un échantillonnage

plus exhaustif réalisé sur des carottes sédimentaires relativement courtes correspondant aux 100 dernières années.

**Chapitre IV** : Impact des forçages environnementaux au cours des 2200 dernières années sur les communautés microbiennes eucaryotes des lacs du Bourget (France) et Igaliku (Groenland)

La stratégie appliquée dans les travaux présentés dans le chapitre IV a été d'étudier l'impact au cours des 2200 dernières années des fluctuations climatiques et des activités humaines sur la communauté biologique de deux lacs de typologies et d'histoires humaines contrastées. Alors que le lac du Bourget (grand, profond et chaud, France) a été caractérisé par la présence d'activités humaines durant toute la période étudiée (notamment marquée par une eutrophisation dans les années 1940s), le lac Igaliku (petit, peu profond et froid, Groenland) a été faiblement impacté par les activités humaines (principalement entre le 9 et le 13<sup>ème</sup> siècle après J.C. et au cours du siècle dernier). Les données acquises sont exploratoires car bien que plusieurs études aient montré que l'ADN de la masse d'eau du lac du Bourget peut être analysé dans les archives sédimentaires récentes (100 ans ; Domaizon *et al.* 2013, Savichtcheva *et al.* 2015), sa présence et sa persistance au cours de plusieurs millénaires n'avaient jamais été étudiées pour ce lac et encore moins pour le lac Igaliku qui a pourtant bénéficié de plusieurs études paléocéologiques s'appuyant sur des proxy classiques (Gauthier *et al.* 2010, Perren *et al.* 2012, Millet *et al.* 2014). Les questions que nous avons abordées dans ce chapitre sont listées ci-dessous.

**Questions spécifiques** : Quelle est l'importance relative des effets des fluctuations climatiques et des activités humaines sur la dynamique temporelle des communautés de micro-eucaryotes lacustres au cours des dernières 2200 années ? Peut-on détecter des points de basculement dans les communautés en lien avec ces forçages ? Existe-t-il une réponse commune des communautés des deux lacs étudiés aux fluctuations climatiques ? Est-ce que l'eutrophisation du lac du Bourget dans les années 1940s a un impact fort sur la structure de la communauté biologique de ce lac ? Quels groupes fonctionnels de micro-eucaryotes répondent à ces forçages ?

**Chapitre V** : Impact des forçages environnementaux au cours des 100 dernières années sur les communautés microbiennes eucaryotes des lacs d'Annecy et du Bourget (France)

Les travaux réalisés dans le chapitre IV ont montré que l'eutrophisation du lac du Bourget à la moitié du 20<sup>ème</sup> siècle a été le forçage le plus fort au cours des deux derniers millénaires pour la communauté microbienne eucaryote. Afin d'étudier plus précisément la réponse de cette communauté à cette pollution des eaux et aux fluctuations climatiques récentes, les travaux réalisés dans le chapitre V sont basés sur l'étude de la dynamique temporelle au cours du siècle dernier de la communauté microbienne eucaryote du lac du Bourget (eutrophisation marquée) et du lac Annecy (peu eutrophisé). La concentration en phosphore a été reconstruite à partir de restes de daphnies retrouvés dans les sédiments lacustres (Berthon *et al.* 2014). Concernant les variations climatiques, alors que les précipitations ont été décrites comme non significativement variables dans cette écorégion, les anomalies de température de l'air (Auer *et al.* 2007) ont été utilisées comme indicateur de forçages climatiques lors de ces travaux.

**Questions spécifiques** : Quels sont les changements temporels dans la richesse, la composition et la structure des communautés au cours des 100 dernières années ? Quel est l'impact des forçages environnementaux (climat et eutrophisation) sur ces communautés ? Quels taxons contribuent aux réarrangements dans les communautés en lien avec les forçages ? Peut-on détecter des réponses aux forçages au niveau des réseaux écologiques (réseaux de cooccurrence) ? Peut-on mettre en évidence des réponses communes aux micro-eucaryotes des deux lacs ?

## I.2.2. Approche méthodologique

Ces travaux de thèse se basent sur une approche de paléocéologie moléculaire, consistant à utiliser l'ADN préservé dans les sédiments lacustres pour retracer les changements temporels dans les communautés biologiques du système. Une partie de l'ADN, extrait à partir des sédiments lacustres, provient des communautés pélagiques et benthiques du lac et peut être utilisée pour retracer les changements dans la richesse, l'abondance et la composition de la communauté microbienne lacustre. Les outils de biologie moléculaire appliqués à l'ADN sédimentaire au cours des 10 dernières années ont été majoritairement utilisés pour cibler des groupes spécifiques de protistes et de cyanobactéries et rarement via le séquençage à haut débit de l'ADN (principalement qPCR, PCR-DGGE ; Coolen *et al.* 2004, Stoof-Leichsenring *et al.* 2012, Domaizon *et al.* 2013, Savichtcheva *et al.* 2015). Bien que de plus en plus utilisé dans les études de diversité microbienne, notamment concernant les protistes et les champignons (Taïb *et al.* 2013, Logares *et al.* 2014, Adl *et al.* 2014, de Vargas *et al.* 2015, Filker *et al.* 2016), le séquençage à haut débit de l'ADN n'a été pour l'instant appliqué qu'une seule fois sur l'ADN sédimentaire pour étudier le plancton eucaryote via une approche rétrospective (Coolen *et al.* 2013). Compte tenu du potentiel du séquençage à haut débit appliqué à l'ADN sédimentaire pour retracer la dynamique temporelle des communautés microbiennes eucaryotes des écosystèmes aquatiques, les travaux présentés dans ce manuscrit sont basés principalement sur l'utilisation de cette méthode. Des carottes sédimentaires provenant de plusieurs lacs (en France, en Suède et au Groenland) et couvrant des cadres temporels contrastés (de 40 ans à 2000 ans), ont été utilisées pour répondre à des questions portant sur des aspects méthodologiques et écologiques (Tableau I.2). Les méthodes et protocoles mis en œuvre sont détaillés dans chaque chapitre correspondant aux principaux résultats acquis. L'approche utilisée est décrite ci-dessous et dans la Figure I.15.

Tableau I.2. Caractéristiques des carottes sédimentaires utilisées dans ces travaux de thèse.

Nom des carottes	Lacs	Année de prélèvement	Période temporelle (années)	Questions
LDB13-I	Bourget (France)	2013	2000	écologie
Iga-2013	Igaliku (Groenland)	2013	2000	écologie
LDA_09_P2_03	Annecy (France)	2009	100	écologie
LDB_09_P2_02	Bourget (France)	2009	100	écologie/méthodologie
Nylandssjön-2007	Nylandssjön (Suède)	2007	40	méthodologie
Nylandssjön-2013	Nylandssjön (Suède)	2013	40	méthodologie

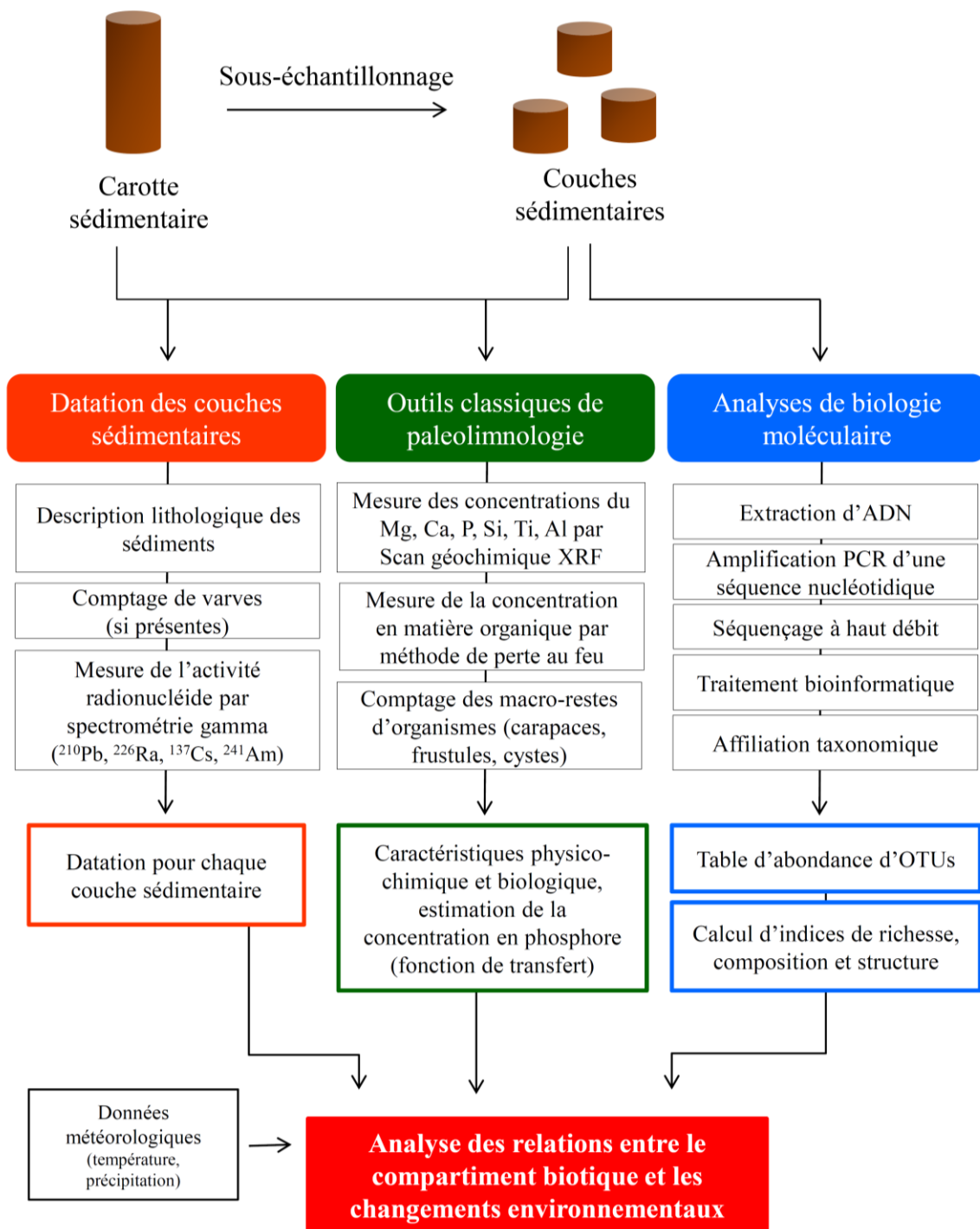


Figure I.15. Approche méthodologique appliquée dans ces travaux de thèse.

Ce travail s'appuie sur une démarche pluridisciplinaire mobilisant des compétences complémentaires notamment pour les aspects concernant les analyses sédimentologiques réalisées sur des carottes de références (mesure des concentrations en éléments minéraux et matière organique, reconstruction des teneurs en phosphore basée sur le comptage de restes de diatomées ou de daphnies) et les datations (par détection de marqueurs lithologiques, comptage des varves quand possible et mesure des activités radionucléides). Ces analyses ont été réalisées au cours de la période 2009-2014 par plusieurs scientifiques des laboratoires EDYTEM (Bourget-du-lac), CARRTEL, Chrono-environnement (Besançon) et de l'équipe de paléolimnologie de l'Université d'Umeå (en Suède).

Le cœur de ces travaux de thèse repose sur les analyses moléculaires à partir des archives sédimentaires en complément du corpus d'études réalisées à l'aide d'autres outils de paléoécologie (comptages des restes d'organismes, lipides, voir Perga *et al.* 2015 pour synthèse). Chacune des carottes sédimentaires dédiée aux analyses moléculaires est découpée longitudinalement et des couches sédimentaires sont prélevées (épaisseur de 0.5 ou 1 cm en fonction des taux de sédimentation connus et des datations). Seul le centre de chaque couche sédimentaire est échantillonné pour les analyses moléculaires afin d'éviter toute contamination externe (pendant le prélèvement, le stockage de la carotte à 4 °C ou la prise en main de la carotte). L'ADN présent dans chaque couche sédimentaire est extrait à partir de 0.5 g de sédiments à l'aide d'une méthode permettant l'**extraction d'ADN** intra et extra cellulaire (kit Macherey-Nagel, Düren, Germany). Afin de cibler les micro-eucaryotes (protistes et champignons), l'**amplification par PCR** est réalisée sur des portions géniques retrouvées dans le génome de tous les eucaryotes et dont l'analyse permet la discrimination et l'affiliation taxonomique. Dans ces travaux de thèse, les analyses du chapitre II sont basées sur le **séquençage à haut débit** (méthode de pyroséquençage 454) d'une portion génique de 350 pb (région V4 du gène codant pour l'ARNr 18S) car les données provenant de la masse d'eau ont été acquises dans le cadre de l'observatoire des lacs pour le suivi mensuel de la communauté microbienne eucaryote. Tous les autres travaux présentés dans cette thèse (chapitres III, IV et V) ont été réalisés par le séquençage haut débit (MiSeq Illumina) d'une partie de la région V7 du gène codant l'ARNr 18S (260 pb). Le fragment ciblé est plus court ce qui permet de retracer de l'ADN plus ancien, prompt à la fragmentation au cours du temps par des processus diagénétiques dans les sédiments (voir description partie I.1.1.4). L'amplification de l'ADN extrait de sédiments du lac du Bourget (carotte LDB13-1) a été testée sur des fragments V4 et V7 ; alors que la portion génique correspondant au fragment

court (260 pb) a été amplifiée à partir de sédiments correspondant à toute la période ciblée (environ 2000 ans), après 1000 ans, aucune amplification n'a été détectée pour les fragments longs. Pour les travaux portant sur l'écologie des lacs (chapitres IV et V) et la question de l'enfouissement de l'ADN au cours du temps dans les échantillons du lac Nylandssjön (chapitre III), nous avons choisi d'utiliser le fragment court (région V7, 260 pb) pour s'assurer d'obtenir une image plus représentative de la communauté biologique du lac. Tout au long de la procédure permettant d'obtenir des tables d'abondance d'OTUs à partir de sédiments, des précautions sont mises en place pour s'assurer de la fiabilité des inventaires moléculaires obtenus notamment via l'utilisation de blouses et gants stériles. De plus, le sous échantillonnage des couches sédimentaires, l'extraction de l'ADN et l'amplification PCR sont des étapes réalisées dans 3 pièces différentes dédiées à chaque tâche et sous hôte à flux laminaire. A chaque étape, des contrôles négatifs (des tubes Eppendorf de 1.5 mL contenant 500 µL d'eau ultra pure) permettent de vérifier la présence de contamination des échantillons. En complément, deux inventaires moléculaires sont séquencés pour chaque couche sédimentaire correspondant à deux extractions d'ADN. Seuls les OTU présents dans les deux inventaires sont gardés pour les travaux portant sur les analyses écologiques afin d'éviter un potentiel biais lié à l'hétérogénéité de la répartition de l'ADN au sein des couches sédimentaires (chapitres IV et V) et des biais d'amplification par PCR. Par la suite, le traitement bioinformatique des données issues du séquençage à haut débit permet d'obtenir pour chaque couche sédimentaire des tables d'abondance d'OTUs. Cette étape, détaillée dans les chapitres correspondant aux travaux, a été réalisée par Didier Debroas au sein du laboratoire LMGE (Aubière).

Dans les domaines de l'écologie microbienne, l'écologie temporelle et la paléoécologie, plusieurs outils/méthodes analytiques et statistiques sont utilisées pour étudier la dynamique des communautés et leurs réponses aux forçages environnementaux (Ramette 2007, Legendre & Legendre 2012, Dornelas *et al.* 2012, Buttigieg & Ramette 2014). Les outils utilisés dans ces travaux, principalement appliqués via le logiciel R, permettent d'étudier :

(i) les variations temporelles et points de basculements des indices de diversité de la communauté (estimateur de richesse Chao1, indice de diversité de Shannon et indice de Simpson, indice d'équitabilité de Pielou). Ces indices ont été transformés pour être exprimés sous la forme des nombres de Hill dans les travaux du chapitre IV (Jost 2007, Gotelli & Chao 2013, Chao *et al.* 2014, Hsieh *et al.* 2016) qui permettent une appréhension plus intuitive de la notion de biodiversité (nombres équivalent d'espèces effectives). La présence de points de

changements dans la dynamique temporelle de ces indices est étudiée via l'analyse *bayesian change point* (bcp ; Erdman & Emerson 2007). L'analyse *bcp* calcule une valeur de probabilité de changements pour chaque paire de dates consécutives (valeurs entre 0 et 1) permettant d'estimer la présence de points de changements dans la série temporelle.

(ii) la modification temporelle dans la structure de la communauté : dans ces travaux de thèse, deux méthodes sont utilisées. La première méthode n'implique pas de contrainte temporelle comme variable et se base sur le calcul de l'indice de dissimilarité Bray-Curtis à partir des données d'abondance des OTUs dans chaque échantillon (ou d'indice de dissimilarité de Jacard basé sur les présences/absences des OTUs). Les matrices de dissimilarité obtenues peuvent être visualiser sous forme de graphique à 2 axes (analyse nMDS) ou sous forme d'arbre hiérarchique clustérisant les échantillons (*hierarchical clustering tree*) (Oksanen *et al.* 2015). Des groupes d'échantillons (suivant une trajectoire temporelle ou non) peuvent ainsi être déterminés (et confirmés par analyse statistique de type ANOSIM) et mettre en évidence les changements temporels abrupts dans la structure de la communauté. La deuxième méthode dite *multivariate tree regression* (MRT; De'ath 2002) inclue le temps comme variable et permet de détecter des points de basculements dans la structure des communautés. Elle est basée sur le partitionnement des données de chaque échantillon en fonction des données d'abondance et sur la cross-validation des résultats (Borcard *et al.* 2011).

(iii) les réarrangements au sein de la communauté : grâce à la méthode SIMPER, il est possible de mettre en évidence la contribution des changements d'abondance des OTUs dans les réarrangements temporels de la communauté. Les réarrangements au sein de la rare biosphère, suspectée d'avoir un rôle important dans la réponse de la communauté aux perturbations, sont étudiés dans ces travaux de thèse par l'étude d'OTUs caractérisées par de larges oscillations au cours du temps entre des états rares et abondants (*oscillating OTUs*).

(iv) les effets de forçages environnementaux sur les communautés au cours du temps via l'analyse GAM (*generalized additive models*; Hastie & Tibshirani 1986, Simpson & Anderson 2009, Wood 2015). Le GAM est une méthode de régression non linéaire permettant de montrer quand le prédicteur (composé d'une somme de fonctions pas forcément paramétriques), affecte une variable réponse. Cette méthode permet de mettre en évidence la contribution de plusieurs facteurs de forçage à la variation temporelle dans les paramètres de la communauté (richesse, diversité, structure de la communauté totale ou de groupes spécifiques ou même de modules constituant les réseaux de cooccurrence). La figure ci-



dessous présente le graphique illustrant la contribution de deux facteurs de forçages (predicteurs) à une variable réponse (par exemple la richesse d'une communauté). Les 4 types de contributions mis en évidence par cette figure montre notamment qu'il est possible de hiérarchiser l'importance relative de chaque forçage aux changements dans la réponse variable (Figure I.16).

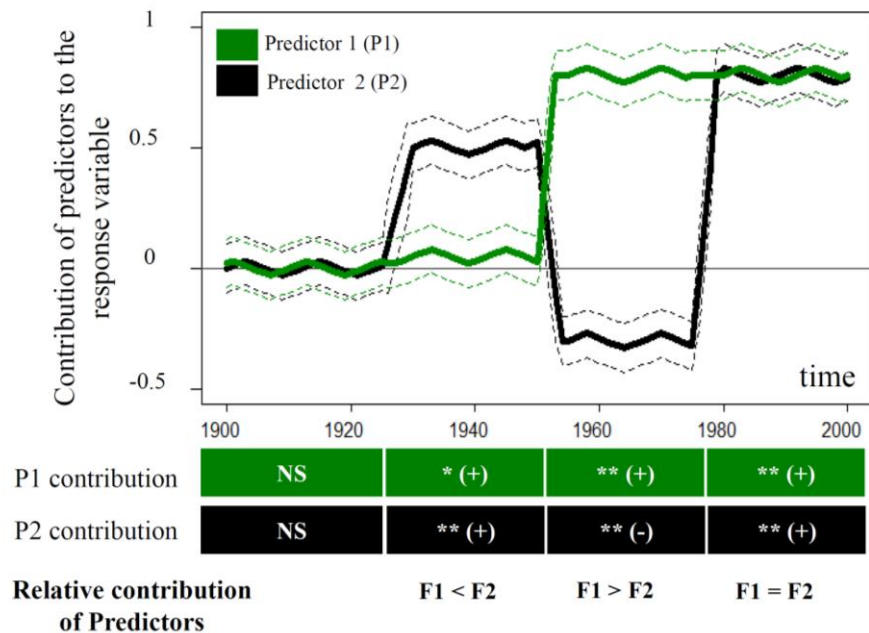


Figure I.16. Analyses GAMs. Figure illustrant la contribution relative de deux prédicteurs aux changements dans une variable réponse.

(v) la structuration des assemblages en réseaux de cooccurrence. Ces analyses permettent de mettre en évidence des groupes d'OTUs interconnectés (modules) qui semblent être corrélés au cours du temps (synchronie dans leur abondance relative), de rechercher des modules communs inter-écosystèmes ou encore de mettre en évidence les liens entre ces modules et les modifications environnementales. Plusieurs méthodes sont utilisées par les écologues microbiens pour créer des réseaux de cooccurrence à partir de données d'abondance d'OTUs (voir synthèse de Faust & Raes 2012 et Faust *et al.* 2015). Dans ces travaux de thèse, la méthode dite WGCNA (Langfelder & Horvath 2008), est utilisée pour créer des réseaux de cooccurrence à partir de données temporelles comme précédemment appliquées dans plusieurs études d'écologie microbienne basées sur des données spatiales (Aylward *et al.* 2015, Guidi *et al.* 2016, Bryant *et al.* 2016).

### I.2.3. Structure de la thèse

En complément des chapitres d'introduction et de discussion (chapitres I et VI), le manuscrit de thèse est composé de 4 chapitres correspondant chacun à une publication scientifique acceptée, soumise ou en préparation. Pour les chapitres IV et V, des analyses non incorporées aux travaux soumis pour publication ont été ajoutées dans ce manuscrit de thèse. En complément, un article de synthèse portant sur l'application des méthodes moléculaires en paléolimnologie, de nouvelles opportunités pour étudier la dynamique à long terme de la biodiversité lacustre à été ajouté en Annexe. Ce papier est actuellement en révision dans le *Journal of Paleolimnology* (Domaizon et al. *in rev*).

- le **chapitre II** : Efficacité d'archivage de l'ADN des microorganismes eucaryotes de la masse d'eau vers les sédiments du lac du Bourget ; Papier accepté en 2015 dans *Microbial Ecology*

- le **chapitre III** : Préservation de l'ADN sédimentaire au cours de l'enfouissement ; Papier en préparation ; Papier en préparation pour soumission dans *Journal of Paleolimnology*

- le **chapitre IV** : Impact des forçages environnementaux au cours des 2200 dernières années sur les communautés microbiennes eucaryotes des lacs Bourget (France) et Igaliku (Groenland) ; Papier accepté dans *Molecular Ecology*

- le **chapitre V** : Impact des forçages environnementaux au cours des 100 dernières années sur les communautés microbiennes eucaryotes des lacs d'Annecy et Bourget (France) ; Papier en préparation pour soumission dans *Environmental Microbiology* au cours du mois d'octobre 2016



## II. Efficacité d'archivage de l'ADN des microorganismes eucaryotes de la masse d'eau vers les sédiments du lac du Bourget

### Is Planktonic Diversity Well Recorded in Sedimentary DNA? Toward the Reconstruction of Past Protistan Diversity

Publié dans *Microbial Ecology* (2015) 70:865-875

Eric Capo<sup>1,2</sup>, Didier Debroas<sup>3</sup>, Fabien Arnaud<sup>4</sup> & Isabelle Domaizon<sup>1</sup> \*

<sup>1</sup> INRA, UMR 42 CARRTEL, 74203 Thonon-les-Bains Cedex, France

<sup>2</sup> Université Savoie Mont Blanc, UMR 42 CARRTEL, 73379 Le Bourget du Lac, France

<sup>3</sup> Clermont Université, Université Blaise Pascal, CNRS, UMR 6023 LMGE, 63171 Aubière, France

<sup>4</sup> CNRS, UMR 5204 EDYTEM, Université Savoie Mont Blanc, 73379 Le Bourget du Lac Cedex, France

#### Contribution des auteurs :

- Conception de l'étude: ID, DD, EC
- Acquisition des données: ID, DD, EC (plateau d'analyses du CARRTEL et LMGE)
- Traitements bioinformatiques: DD
- Analyses et interprétations des données: EC, ID, DD, FA
- Ecriture du papier: EC, ID, DD, FA

#### Resumé

Les travaux présentés dans ce chapitre sont basés sur la comparaison d'inventaires moléculaires provenant de la masse d'eau (2 m et 130 m) et des sédiments récents (âgés de moins de 60 ans) du lac du Bourget. Ces inventaires moléculaires, obtenus par le séquençage 454 de la région V4 du gène codant l'ARNr 18S, ont permis de révéler la bonne efficacité d'archivage de l'ADN planctonique dans les sédiments pour la majorité des groupes microbiens eucaryotes. En effet, plus de 70 % des taxons présents dans la masse d'eau ont également été retrouvés dans les sédiments, notamment au sein de groupes photosynthétiques (Chlorophytes, Dinophycées, Chrysophycées), parasites (Apicomplexés, Chytrides) et saprotrophes (Cryptomycota). Seuls les Haptophytes et les Cryptophytes ont été identifiés comme des groupes moins bien archivés dans les sédiments. En complément des analyses ADN, une analyse transcriptomique (basée sur l'ARNr) a été menée ; aucune amplification PCR de rétro-transcrits d'ARN n'a été obtenue dans les sédiments au-delà des deux premiers cm suggérant que le signal ADN archivé dans les sédiments n'est pas biaisé par la présence d'une biosphère active eucaryote dans les sédiments du lac du Bourget.

## II.1. Abstract

Studies based on the coupling of a paleolimnological approach and molecular tools (e.g., sequencing of sedimentary DNA) present a promising opportunity to obtain long-term data on past lacustrine biodiversity. However, certain validations are still required, such as the evaluation of DNA preservation in sediments for various planktonic taxa that do not leave any morphological diagnostic features. In this study, we focused on the diversity of planktonic unicellular eukaryotes and verified the presence of their DNA in sediment archives. We compared the molecular inventories (high-throughput sequencing of 18S ribosomal DNA) obtained from monitoring the water column with those obtained for DNA archived in the first 30 cm of sediment. Seventy-one percent of taxonomic units found in the water samples were detected in sediment samples, including pigmented taxa, such as Chlorophyta, Dinophyceae, and Chrysophyceae, phagotrophic taxa, such as Ciliophora, parasitic taxa, such as Apicomplexa and Chytridiomycota, and saprotrophs, such as Cryptomycota. Parallel analysis of 18S ribosomal RNA (rRNA) transcripts revealed the presence of living eukaryotic taxa only in the top 2 cm of sediment; although some limits exist in using RNA/DNA ratio as indicator of microbial activity, these results suggested that the sedimentary DNA mostly represented DNA from past and inactive communities. Only the diversity of a few groups, such as Cryptophyta and Haptophyta, seemed to be poorly preserved in sediments. Our overall results showed that the application of sequencing techniques to sedimentary DNA could be used to reconstruct past diversity for numerous planktonic eukaryotic groups.

## II.2. Introduction

Unicellular eukaryotes (i.e., protists) are key components of aquatic systems, mainly as contributors to plankton biomass and nutrient cycling (Caron *et al.* 2009b). Over the last decade, several studies showed that unicellular eukaryotes represent a diverse and abundant fraction of lake plankton and cover all functional roles (photosynthetic, saprotrophic, parasitic, and phagotrophic) (Richards *et al.* 2005, Lepère *et al.* 2008, Mangot *et al.* 2009). Next-generation sequencing technologies (NGS) now offer the opportunity to deeply investigate protist diversity in marine and freshwater environments (Amaral-Zettler *et al.* 2009, Lepère *et al.* 2013, Mangot *et al.* 2013, Logares *et al.* 2014, Debroas *et al.* 2015). However, field surveys based on NGS tools are still rare, and when they exist, they are generally used for short-term studies. Such DNA surveys should be used to acquire long-term data to consider the dynamics of planktonic diversity on time scales that are consistent with changes in environmental conditions, e.g., changes in lake trophic state that occur on the pluri-annual/decennial time scale.

One opportunity for studying the dynamics of lacustrine diversity over long-time scales comes from paleolimnological approaches. Morphological remains or cysts/spores preserved in sediments are used as classical biomarkers to track changes in diatoms, chironomids, or cladocerans assemblages (Battarbee *et al.* 2001, Millet *et al.* 2009, Alric & Perga 2011). However, these classical proxies reflect only a small part of total planktonic diversity. To complement these classical proxies, sedimentary DNA can be used to potentially obtain information about all lacustrine taxa, including those with no visible diagnostic features preserved in sediment. Some recent paleolimnological approaches showed the potential of DNA studies to generate new insights on the dynamics and diversity of various microbial taxa, including protistan groups (Coolen *et al.* 2004, Boere *et al.* 2009, Corinaldelsi *et al.* 2011, Stoof-Leichsenring *et al.* 2012), cyanobacteria (Fernandez-Carazo *et al.* 2013, Savichtcheva *et al.* 2015), and bacteria (Coolen *et al.* 2008), even enabling the evaluation of the relative importance of global warming and eutrophication as drivers of changes in microbial community composition (Domaizon *et al.* 2013, Savichtcheva *et al.* 2015).

Until now, most of these molecular paleoecological studies focused on specific taxonomic groups. The challenge is now to investigate a wider scope of diversity, such as protistan diversity, which includes a large variety of taxa and functions (Pawlowski *et al.* 2012). In

2013, Coolen *et al.* (2013) applied high-throughput sequencing to sedimentary DNA (454-sequencing on 18S ribosomal RNA (rRNA) gene) to reconstruct the eukaryotic community structure in the Black Sea over the last ca. 11,400 years. They showed that diverse eukaryotic groups can be retrieved from all sediment strata, and they identified tipping points in eukaryotic diversity related to the freshwater-marine transition. Despite promising results, more validations must be done to better understand the advantages and limitations of this approach of identifying past species from sedimentary DNA. One of the main issues is that little has been done to quantify how well sedimentary DNA reflects the biological community present in the water column. It is indeed hard to determine if, when a taxon is not detected in a sedimentary DNA sample, the DNA has not been efficiently archived in the sediment or the taxon was not present in the water column. The outcomes of the DNA-paleolimnological studies are likely to be impaired by variable DNA preservation among different species, as shown by Boere *et al.* (2011b) for three marine microbial groups (dinoflagellates, diatoms, and green-sulfur bacteria). Thus, some validations and calibrations are necessary to investigate the congruence between sedimentary DNA and planktonic DNA signals, especially since modifications or even degradation could affect planktonic DNA during sedimentation along the water column and/or during burial in sediments due to diagenetic processes (Dell'Anno *et al.* 2002, Corinaldelsi *et al.* 2008, Boere *et al.* 2011a). Such validations require specific experimentations and in situ surveys, e.g., the use of sediment traps to follow the DNA signatures of planktonic taxa that are sedimenting.

Alternatively, a basic approach could be to verify the presence of planktonic taxa DNA in recent sedimentary deposits by comparison of molecular inventories (water and sediment). We performed this comparison between protistan diversity detected in the water column (epilimnion and hypolimnion) during a 1 ½ year survey and protistan diversity retrieved from sediments (uppermost 30 cm of the core). This work was conducted in a deep peri-alpine lake, Lake Bourget. Both planktonic and sedimentary DNA were subjected to 454 pyrosequencing of the 18S rRNA gene (V4 region). We also accounted for the potential presence of active protists in the sediment core through RNA analyses.

The molecular inventories were produced both in terms of operational taxonomic units (OTUs) and phylogenetic units (PUs). OTUs are classically used as proxies of ecological species (Sogin *et al.* 2006, Shade *et al.* 2013 and Logares *et al.* 2014); however, the thresholds used to delineate OTUs are still under debate, and the presence of a nucleotide polymorphism

that could artificially increase the number of OTUs cannot be excluded. An alternative is to use phylogenetic trees to define PUs instead of OTUs (Taïb *et al.* 2013, Dunthorn *et al.* 2014); this can be performed even using NGS data (Debroas *et al.* 2015, Vergin *et al.* 2013). PUs are monophyletic units gathering OTUs that are assigned to the same nearest neighbor reference and that clustered as a monophyletic branch of the tree. The use of PUs, defined phylogenetically, limits the overestimation of richness that could be linked to the arbitrary cutoff thresholds of OTUs. This phylogenetic approach is particularly well adapted to our paleoecological study because sedimentary DNA is subjected to potential modifications (during sedimentation and burial) that may induce an artificial increase of the number of OTUs. Therefore, here we focused on the PU level instead of the classical OTU level.

In the present study, we investigated whether the same PU could be retrieved from both water and recent sedimentary deposits. The goal was to obtain initial knowledge on how the sedimentary DNA signal may reflect the diversity of planktonic protists present in the water column. We further investigated if the importance (i.e., recurrence or number of DNA reads within molecular inventories) of taxa in the water column tends to favor the archiving of DNA in sediments. The present study presents original results regarding the potential utility of DNA to analyze long-term protistan diversity from sedimentary archives in lakes.

## **II.3. Materials and Methods**

### **III.3.1. Study site**

Lake Bourget (45°44' N; 5°51' E), situated at an altitude of 231 m, has a maximum depth of 147 m and is currently classified as an oligo-mesotrophic lake (Long-term Observatory of Alpine Lakes, INRA Thonon; <http://www6.inra.fr/soere-ola>). This lake is considered a meromictic lake, but mixing events can occur throughout the whole water column. This lake presents hypoxic to anoxic deep water layers during the stratification period (~June to December).

### **II.3.2. Sampling**

Water samples were collected monthly using a Van Dorn bottle for 1.5 years (15 dates from March 2011 to September 2012). To consider the potential heterogeneity of the vertical



distribution of protistan taxa, both the epilimnion (2 m depth) and the deep hypolimnion (130 m depth) were sampled. Moreover, the analysis of these two contrasted depths offers the possibility to detect the potential loss of DNA signal of epilimnetic taxa that may occur during the sedimentation along the water column. A subsample of water (2 L) was collected on 0.2- $\mu\text{m}$  pore polycarbonate filters (Millipore, Merck Millipore, Billerica, MA, USA), following the protocol described by Mangot *et al.* (2013); the filters were stored at  $-80\text{ }^{\circ}\text{C}$  until nucleic acid extraction.

The sediment core was sampled in 2010 (same location as water sampling). Sediment strata were sampled along the first 30 cm of the sediment core. The sampling interval was chosen to fit two complete lamina triplets, corresponding to 2 years of sedimentation (Domaizon *et al.* 2013, Jenny *et al.* 2013). Only the center of each slice of sediment was kept for molecular analysis because it was preserved from any contamination due to contact with the plastic liner at the periphery. For each sediment stratum, an aliquot of wet sediment was dried at  $60\text{ }^{\circ}\text{C}$  and weighed to estimate the dry weight; the rest of the sample was divided in two subsamples subjected to molecular analysis (DNA and RNA extractions). No specific treatment was applied to the DNA samples, while samples dedicated to RNA analysis were preserved in RNAlater solution (Qiagen, Hilden, Germany) for immediate RNA stabilization and protection. All samples dedicated to molecular analysis (DNA and RNA) were kept at  $-80\text{ }^{\circ}\text{C}$  until nucleic acid extraction was performed.

The detailed protocol for dating is available in Domaizon *et al.* (2013) and Jenny *et al.* (2013). The analyzed sediment core was precisely dated, both by counting visible annual layers and by careful correlation with a reference core that had been dated using radionuclide-based chronologies ( $^{210}\text{Pb}$  and  $^{137}\text{Cs}$ ) and reported flood events (1910, 1919, 1924, 1944, and 1952). For the studied time period we provided an accurate and precise chronology; the dating uncertainty is close to zero when annual lamina are present (*i.e.*, since 1933 AD) (Jenny *et al.* 2013, Giguet-Covex *et al.* 2010, Savichtcheva *et al.* 2015).

### **II.3.3. Molecular Analysis**

#### *DNA Extraction, PCR, and Pyrosequencing*

For water samples ( $n=30$ ), the nucleic acid extraction method was modified from Hugoni *et al.* (2013) using a combination of mechanical and enzymatic cell lysis followed by extraction using the AllPrep DNA/RNA kit (Qiagen, Valencia, CA, USA). For each sediment layer

(n=30), DNA was extracted from three subsamples (1 g of wet sediment each) to avoid bias from any heterogeneity of the sediment, and the resulting DNA extracts were pooled. DNA was extracted using an UltraClean Soil DNA Isolation Kit according to the manufacturer instructions (Mobio, Carlsbad, CA, USA).

All 30 water samples (covering a period of 1 ½ years) were submitted to PCR amplification. For sediment, six strata (each one representing two years of sedimentary deposit) were selected from the first 30 cm of sediment. Two sediment strata corresponded to oligo-mesotrophic conditions in the lake, in recent and older periods (i.e., 2008–2009 and 1951–1952); four other sediment layers were selected in between those two oligo-mesotrophic periods and corresponded to contrasted conditions in terms of trophic status in the lake (i.e., eutrophic conditions for the strata 1972–1973 and 1969–1970, mesotrophic to meso-eutrophic conditions for the strata 2003–2004 and 2000–2001; see Table **II.1**).

PCR amplification of the V4 region of 18S rDNA gene (350 bp) was performed using the general eukaryotic primers NSF563 (CGCGGTAATTCCAGCTCCA) and NSR951 (TTGGYRAATGCTTTCGC) from Mangot et al. (2013). Each DNA sample was individually PCR amplified, in duplicate 25- $\mu$ L reactions, by initial denaturation at 95 °C for 5 min followed by 25 cycles of 40 s at 95 °C, 40 s at 53 °C, and 60 s at 72 °C. At the end of the amplification, the amplicons were subjected to a final 7-min extension at 72 °C. The PCR products were purified (UltraClean®PCR Clean-Up Kit, Mobio, Carlsbad, CA, USA) and quantified using PicoGreen (Promega, Sunnyvale, CA, USA). Finally, amplicons of all of the samples were pooled at equimolar concentrations and pyrosequenced using a Roche 454 GS-FLX system (Titanium Chemistry) by GATC (Konstanz, Germany).

### *Bioinformatics Processing Applied for Pyrosequencing Data*

Pyrosequencing data obtained for DNA samples (water and sediment) represented 1,062,828 raw sequences. These sequences were first submitted to stringent cleaning procedures: all of the sequences were checked against the following quality criteria: (i) no undefined bases (Ns), (ii) quality score  $\geq 23$  according to the PANGEA process (Giongo *et al.*, 2010), (iii) a minimum sequence length of 200 bp, and (iv) no sequencing error in the forward primer. Putative chimeras were detected by UCHIME (Edgar *et al.* 2011). After the cleaning step, 448,287 reads remained.

Table II.1. The upper part of the table provides a description of the analyzed samples regarding the period covered by each sample (periods), the depth of sampling in the water column and in the sediment core (depth), the phosphorous concentration in the lake for each period ([P] in water in  $\mu\text{g/L}$ ), and the number of samples analyzed. The concentration of phosphorous in lake water is provided as an indicator of the trophic state of the lake: long-term changes in lake total phosphorus concentrations were reconstructed from a diatom-inferred transfer function (Berthon *et al.*, 2013) based on the identification and quantification of diatoms' frustules preserved in sediment (all values with asterisks); total phosphorus concentrations for recent periods were also verified from direct measurement on lake water (chemical survey performed by the observatory on alpine lakes; all values with crosses). The lower part of the table present the number of PUs detected from massive sequencing in water and sediment samples, as well as percentages of PUs from water retrieved in sediment samples.

Periods	Water	Sediment						
	2011–2012	2008–2009	2003–2004	2001–2002	1972–1973	1969–1970	1951–1952	All sed. samples
Depth (m for water, cm for sediment)	2 and 130	0.5–1.2	1.8–2.5	2.5–3.5	17.5–18.5	19–19.9	25.5–26.5	
[P] in water ( $\mu\text{g/L}$ )	14 <sup>†</sup>	14 <sup>†</sup> *	17 <sup>†</sup> *	23 <sup>†</sup> *	71*	50*	14*	
Number of samples	15 and 15	1	1	1	1	1	1	
DNA extracted in 3 replicates for each sediment layer								
Number of PUs	954	476	426	349	355	413	478	798
% of “water PUs” retrieved in sediment		47 %	43 %	37 %	36 %	41 %	48 %	71 %

After the cleaning step, clustering and taxonomic affiliations were performed. The DNA reads were clustered at a 95 % similarity threshold with UCLUST (Edgar 2010) in order to define OTUs. By using FastTree (Price *et al.* 2010), the representative OTUs were used to build phylogenetic trees for the main taxonomic groups with reference sequences. These references, included in PANAM database, were extracted to the SSURef SILVA database (Pruesse *et al.* 2007) according to the following criteria: length >1200 bp, quality score > 75 %, and a pintail value > 50. In addition, this reference database was modified to account for the taxonomy of typical previously defined freshwater lineages. The taxonomic annotation was conducted using the nearest neighbor (NN) and the last common ancestor (LCA) methods. Finally, PUs were defined; each PU represented a monophyletic group that gathered close OTUs affiliated to a common reference (the nearest neighbor among the reference sequences in the tree; see supplemental material Figure **II.S1**). This bioinformatics process was implemented in the PANAM pipeline (<http://code.google.com/p/panam-phylogeneticannotation/downloads/list>) and described in greater detail in the related paper (Taïb *et al.*, 2013).

#### *RNA Preparation of DNA-Free RNA from Sediments, cDNA Synthesis, and Sequencing*

For each sediment layer (n=30), RNA was extracted using TRI Reagent Solution (Ambion), according to the manufacturer's instructions. All extracted RNA samples were treated with 1 U DNase (Invitrogen) to eliminate any traces of DNA. Before PCR, DNA-free RNA samples were randomly transcribed into complementary DNA (cDNA) using the SuperScript III First-Strand Synthesis System (Invitrogen) and random hexamers, following the protocol described by the manufacturer. The reverse transcriptase (RT) reactions were performed in a TProfessional Basic Gradient Thermocycler (Biometra) and comprised incubation for 10 min at 25 °C, 50 min at 50 °C, and 5 min at 85 °C, followed by an RNA-removal step with addition of 1 µL of RNase H and incubation for 20 min at 37 °C. The resulting cDNAs served as PCR templates to study the metabolically active eukaryotes in the lake sediments. PCR amplifications were performed using the same protocol used for DNA (described above). The amplification products from these PCRs were used for the construction of clone libraries using a TOPO TA cloning kit (Invitrogen, Carlsbad, CA) with PCR vector 2.1 according to the manufacturer's instructions. One hundred clones were sequenced (Beckman Coulter Genomics) for each of the sediment samples for which cDNA amplicons were obtained.

Sanger sequences obtained from cDNA were aligned using the CLUSTALW procedure (Thompson *et al.* 1994) with the BioEdit Sequence Alignment Editor (version 7.0.5.3). The 18S rRNA alignments were analyzed using DOTUR software (Schloss & Handelsman 2005) (<http://schloss.micro.umass.edu/software>) to determine the operational taxonomic units with a cutoff equal to 95 %. The taxonomic affiliation was obtained by comparison with sequences available in databases using the alignment search tool BLASTN (Altschul *et al.* 1990) from the National Center for Biotechnology Information; moreover, cDNA sequences were compared with DNA sequences (pyrosequencing data) with similar taxonomic affiliation.

#### **II.3.4. Data Accessibility**

Data obtained for DNA samples by pyrosequencing were deposited in the Dryad Digital Repository: doi:10.5061/dryad.5mv53. The nucleotide sequences obtained for RNA (cDNA analyzed by Sanger sequencing) were submitted to the GenBank database. The assigned accession numbers are: KP059150 to KP059246.

### II.3.5. Data Analysis and Statistical Tests

To estimate the effect of PU recurrence, which varies from 0 to 15 based on the presence/absence of the PU for the 15 sampling dates, and the effect of PU relative abundance (number of reads for a given PU/total number of reads) on their presence/absence in sediment, ANOVA tests were performed. All statistical analyses were carried out in the R environment ([www.r-project.org](http://www.r-project.org)).

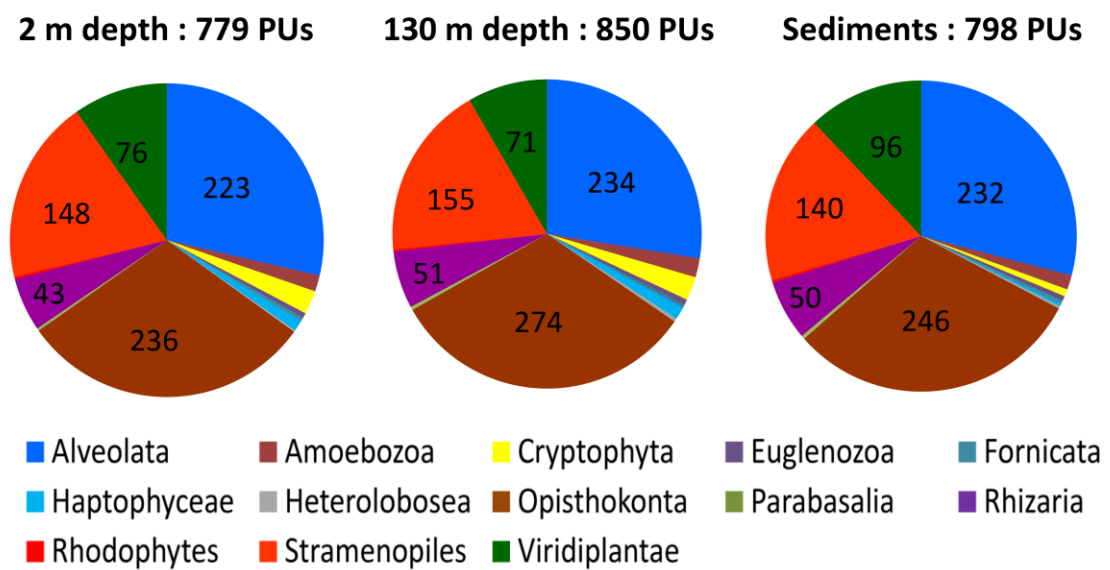


Figure II.1. Richness (number of PUs) distribution within the different phyla in water samples at 2 and 130 m depth and in sediment samples

Table II.2. Taxonomic composition of active eukaryotes detected in the first two centimeters of sediment, expressed as the relative number of reads affiliated to the main phylogenetic groups.

	Phylum	cDNA % reads
unicellular eukaryotes	Alveolata	2.0
	Ciliophora	66.2
	Perkinsea	1.4
	Fungi	4.0
	Rhizaria	0.7
	Cercozoa	2.0
	Biliphyte	0.7
	Chlorophyta	4.8
	Eukaryota	2.0
	metazoans	Copepod

## II.4. Results

### II.4.1. Richness and Taxonomic Composition of Unicellular Eukaryotes in Water and Sediments

DNA extracted from the thirty water samples and six sediment strata was successfully amplified. After DNA-sequence processing, 34,138 OTUs were obtained (95 % similarity cutoff), and 11,945 OTUs were kept after removing OTUs with only one read (singletons). Based on phylogenetic trees, 1361 phylogenetic units (PUs) were defined (185 to 478 PUs per sample) (Table II.S1). We removed 284 PUs because they constituted only one OTU, and finally 1077 PUs were kept for further analysis.

Five phyla predominated, in terms of richness in both the water and sediment samples (Figure II.1): Alveolata, Opisthokonta (represented mainly by Fungi), Rhizaria, Stramenopiles, and Viridiplantae (mainly Chlorophyta). These five phyla represented 94 % of the total richness and accounted for 90 % of the total DNA reads (Table II.S2). The most diverse groups within these phyla were Dikarya (316 PUs within Fungi) and Ciliophora (132 PUs within Alveolata). The richness distribution within the main phylogenetic groups was quite similar, regardless of the type of sample (water or sediment). The Chao1 indices varied from 1112 to 11,762 for water samples and from 5441 to 8177 for sediment samples, showing the high diversity that could be retrieved from sediment (Table II.S1). However, the number of PUs retrieved from sediment, 798 PUs (4020 OTUs), was the lowest compared with the total richness found in water samples, 954 PUs (9972 OTUs) (Table II.1, Table II.S1).

Concerning RNA analysis, 18S rRNA transcripts were only recovered from the top 2 cm of sediment. The obtained sequences belonged mainly to Ciliophora (66 % of the sequences). Some reads belonging to Fungi, Chlorophyta, Perkinsea, and others were also retrieved (Table II.2). Thirty seven percent of the sequences affiliated with protistan taxa were represented by metopid ciliates. These dominant RNA sequences matched DNA sequences found only in the uppermost sediment layers at low abundances (i.e., less than 0.1 % of DNA reads obtained for these two strata). Other RNA sequences were assigned to ciliates of the genera *Stokesia* and *Perisincirra* and were also detected in water and sediment DNA samples at low abundances in terms of reads number (< 0.15 % of DNA reads obtained for these two sediment layers). Some pigmented taxa were identified within RNA sequences, in particular a Scenedesmaceae taxa matching one PU detected in water and sedimentary DNA (PU\_473 Scenedesmaceae sp.).

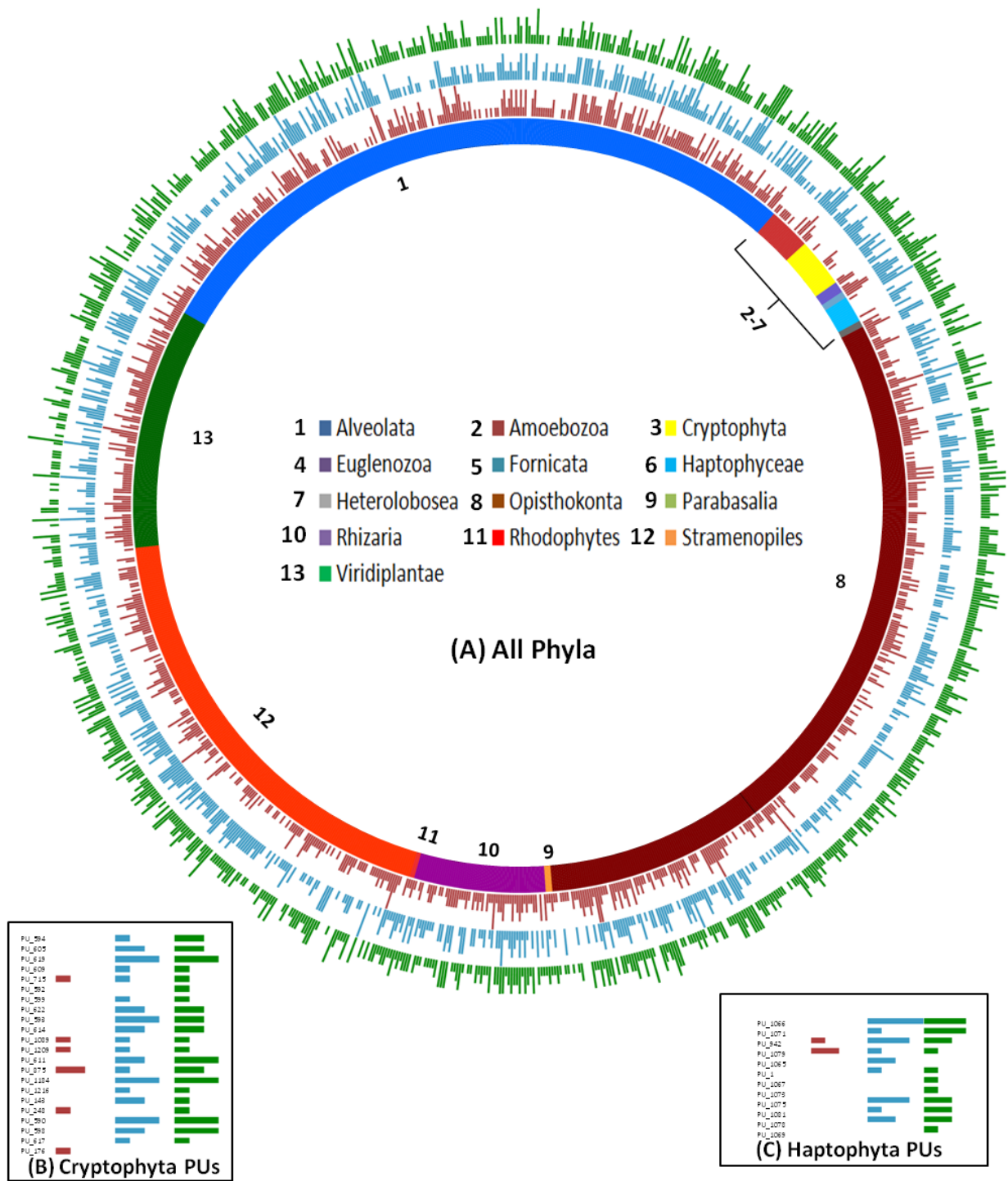


Figure II.2. (A) Composition of the protistan community in water samples (at 2 and 130 m depth) and in sediment samples, expressed as the number of DNA reads obtained for each PU within the main phylogenetic groups; (B) focus on Cryptophyta PUs; (C) focus on Haptophyta PUs. The graph was prepared with iTOL (Letunic & Bork 2011). Five ranges of abundance of reads are represented according to the bar height: 1–10, 11–100, 101–1000, 1001–10,000, and 10,000–100,001 reads. The colored PU bars indicate sediment samples in brown–red, and water samples in blue (2 m depth) and green (130 m depth).

#### **II.4.2. Phylogenetic Diversity Archived in Sediment**

##### *Number of Planktonic PUs Archived*

Seventy-one percent of the PUs detected in the water column was found in sedimentary DNA (Table **II.1**). Our results suggest that a major part of the PUs detected in the sediment originated from taxa living in the water column. When considering each sediment layer separately, the percentage of common PUs between water and sediment varied from 36 to 48 % (Table **II.1**). A higher percentage of shared PUs was observed for both the upper sediment layer and the deepest of the six analyzed sediment strata (corresponding to sediment deposited in 2008–2009 and 1951–1952, respectively). Only 123 PUs found in the analyzed sediment layers were not detected in water samples; the taxa retrieved specifically from sediment belonged to various protistan groups, including photosynthetic phyla such as Viridiplantae (mainly represented by Chlorophyta) (Figure **II.2**).

##### *Taxonomic Groups Archived Well in Sediment*

The planktonic PUs retrieved from the sediment belonged to 71 sub-phyla. We observed high archiving for 50 of these subphyla because at least 60 % of their constitutive PUs was retrieved from the sediment. Some constituted a moderate number of PUs (less than 25 PUs in the water column), such as Fungi groups (Cryptomycota and Glomeromycota), Oomycetes, Perkinsea, and Streptophyta, and appeared to be particularly well archived in sediment (more than 80 % of the PUs were in sediment and the water column) (Figure **II.2**, Table **II.S3**). Others were major contributors to the overall diversity in the water column (i.e., each subphyla comprising more than 25 PUs) and were also well preserved in sediment. This is the case for nine subphyla, which represented 479 PUs out of the 675 PUs shared between water and sediment (Table **II.S3**). For instance, for Cercozoa, Chlorophyta, and Dinophyceae, we observed that 80 % of their constitutive PUs was retrieved from sediment. The percentages for Apicomplexa, Bacillariophyta, Chrysophyceae, Ciliophora, and Fungi groups (Chytridiomycota and Dikarya) were slightly lower, with 60 to 80 % of the PUs shared between water and sediment (Table **II.S3**).



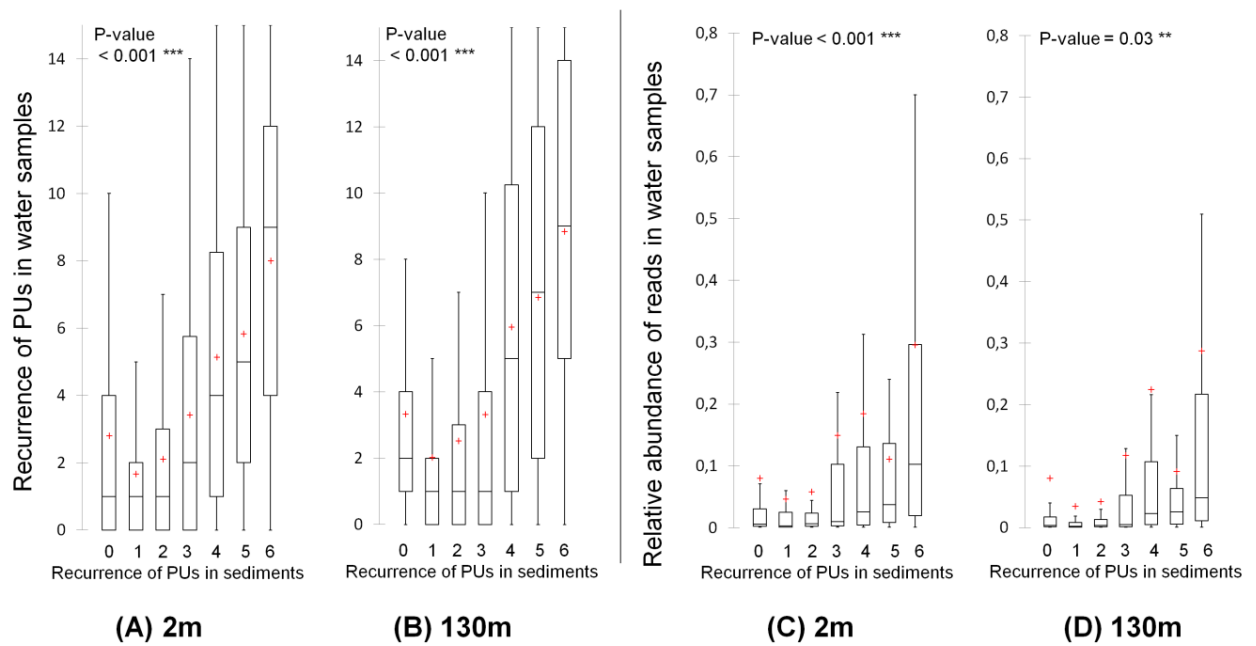


Figure II.3. Boxplots and p value obtained for ANOVA tests performed to assess the relationships between PU recurrence in water at 2 m (a) and 130 m (b) depths and their recurrence in sediments, and, the relationships between the relative number of reads associated with PUs at 2 m (c) and 130 m (d) depth and their recurrence in the sediment

#### *Cryptophyta and Haptophyta: Two Groups Whose Diversity Is Incompletely Retrieved from Sediments*

Some relatively diverse groups (more than 12 PUs found in the water column) had a low number of PUs retrieved from sediment. For example, less than 25 % of the Cryptophyta and Haptophyta PUs from the water was detected in sediment samples. Of the 21 Cryptophyta PUs detected in the water column (~1.30 % of the total reads obtained for water samples), only 5 PUs were retrieved from sediment samples (representing 0.1 % of total reads from sediment samples). Among the 16 non archived PUs, 11 were well represented in terms of recurrence and number of reads in the water samples analyzed; for instance, PU\_590 Cryptophyta\_4 was found in 22 water samples (with 887 reads) but was not retrieved from any of the sediment samples (Figure II.2b, Table II.S3). Only 2 of the 12 Haptophyta PUs found in water samples were retrieved from the analyzed sediment layers (Haptophyta PUs represented 0.04 and 2.28% of the total reads from sediment and water samples, respectively). Some Haptophyta were found with high recurrence or number of reads in water samples but were not detected in sediment, as for instance the Prymnesiophyte PU\_1066 represented by 3798 reads in the water samples (i.e., 1.73 % of the total reads) (Figure II.2c, Table II.S3).

### **II.4.3. Potential Links Between Recurrence and Abundance of PUs in Water and Their Archiving in Sediment**

High recurrence over the year and predominance, in terms of number of reads, in the water column are factors that could favor the archiving of PUs in sediment. We considered the recurrence of each PU at 2 and 130 m depth. In both cases, we observed a significantly positive effect of the recurrence of PUs in water on the archiving rate of PUs (i.e., presence in sediment and recurrence in the 6 sediment samples) (Figure II.3). We also detected a significant effect of the relative abundance of reads in water layers on their presence and recurrence in the sediment samples, especially for PUs originating from the 2-m water layer (Figure II. 3).

## **II.5. Discussion**

The present work does not attempt to precisely study sedimentation processes but rather to provide original results regarding common taxa (PUs) between water layers and sediment with the aim of providing some clues about the preservation of protistan DNA in sedimentary archives. For the first time, high-throughput sequencing (454-sequencing) was used to investigate how protistan DNA detected in the water column is preserved in sediment and, in particular, which planktonic PUs are represented in the sediment.

### **II.5.1. Overall Diversity of Protistan Taxa in Water and Sediment**

The richness distribution (numbers of PUs) in the various phyla highlights the large diversity of unicellular eukaryotes in analyzed samples (both water and sediment). Over the last decade, several papers showed that unicellular eukaryotes represent a diverse and abundant fraction of lacustrine plankton (Richards *et al.* 2005, Mangot *et al.* 2009, Lepère *et al.* 2013, Taïb *et al.* 2013). For instance, Taïb *et al.* (2013) showed that the Fungi, Alveolata, and Stramenopiles taxa were the most diverse phylogenetic groups of planktonic protists in the epilimnion of eight lakes (including Lake Bourget). However, the available studies have generally used unique sampling, while we performed repeated sampling over 15 months. The distribution patterns of protistan richness observed from this temporal survey were quite similar to those obtained by Taïb *et al.* (2013), including the same dominant phylogenetic groups for the 1077 identified PUs. In addition, our results showed that the taxonomic diversity found from sedimentary DNA was globally similar to those found in the water

samples, with a large part of the diversity associated with Alveolata, Fungi, Stramenopiles, Viridiplantae, and Rhizaria.

A major part of the diversity is composed by rare PUs (778 PUs, representing less than 0.1% of the total reads in the water column). Debates about the true nature of this rare biosphere are still numerous (effects of PCR artifacts, dormant cells, etc.), but recent studies showed that the rare OTUs and PUs actually represent a significant part of the active taxa in freshwater (Wilhelm *et al.* 2014, Debroas *et al.* 2015) and marine ecosystems (Logares *et al.* 2014). There is no doubt that additional NGS molecular inventories based on DNA and RNA are still needed to provide an in-depth analysis of complex protistan assemblages, and multiple temporal-scale studies are crucial to obtaining an accurate picture of the dynamics of protistan diversity. Characterizations of the microbial community are typically performed at short or moderate time scales (months or at most years) within the context of lake monitoring (Mangot *et al.* 2013, Nolte *et al.* 2010). Therefore, no direct evidence exists for changes in protistan assemblages occurring over longer time scales (decades, centuries). The coupling between paleolimnology and high-throughput sequencing should provide a key advancement in describing the long-term temporal dynamics of protistan diversity.

However, some protistan taxa may live in sediments (Coolen *et al.* 2009), therefore the potential presence of an active eukaryotic biosphere in the sediment has to be considered. The risk of overrepresentation of DNA assigned to active taxa has to be assessed more particularly, because this may induce a bias in the “DNA paleosignal”.

### **II.5.2. Active Protists Do Not Bias the Paleosignal**

We performed a parallel RNA-based analysis to verify the presence of active eukaryotes in our sediment samples, as proposed by Stoeck *et al.* (2007). The limitations of using RNA as an indicator of microbial activity were recently discussed by Blazewicz *et al.* (2013); however, in the absence of other methodologies to identify active microbial groups in molecular inventories, the use of rRNA is the best solution (e.g., Logares *et al.* 2014). From the rare available data, the presence of active eukaryotic taxa in sediment was mostly attributed to fungal taxa, although a few sequences from Viridiplantae, Bacillariophyta, and Alveolata phyla were also detected in deep-sea subsurface sediments down to 37 m below the sea floor (Edgcomb *et al.* 2011, Orsi *et al.* 2013, Nolte *et al.* 2010). Here, we showed that active protistan taxa were retrieved only in the uppermost strata of the sediment core. While active bacterial and archaeal rRNA were amplified and sequenced for all the sediment samples in the first 30 cm (data not shown), PCR amplification of eukaryotic DNA was

unsuccessful in the deeper strata, suggesting that there was no or negligible eukaryotic activity. From our results, few eukaryotic PUs correspond to active taxa; those that did were mainly from Ciliophora while some reads were attributed to the Fungi, Chlorophyta, Cercozoa, and Perkinsea taxa. Recently, Chambouvet *et al.* (2014) also reported the presence of ribosomally active Perkinsea in superficial marine sediments, suggesting that these taxa either exhibited parasitic activity in marine sediments or were present as a “seed bank”.

In our study, the RNA signal associated with putative parasites was low, and among the dominant active taxa, we detected the presence of metopid ciliates, which are known to live in oxygen-depleted sediments (Fenchel & Finlay 1991). Other ciliates (*Stokesia* and *Perisincirra*) and pigmented taxa, such as Scenedesmaceae, for which RNA reads were detected are likely planktonic and were probably in sediment as dormant cells still producing RNA (Blazewicz *et al.* 2013). Orsi *et al.* (2013) suggested that rRNA associated with photosynthetic groups can be protected within dehydrated cellular forms, such as spores, cysts, and pollen, which are dormant cellular forms in lake sediments. For instance, viable resting cysts of dinoflagellates have been found in one-century-old sediments (Miyazono *et al.* 2012).

From our data, we verified that the active taxa were not overrepresented in NGS data obtained from sedimentary DNA, confirming that these active taxa will not mask past inactive taxa in sedimentary DNA inventories.

### **II.5.3. Efficiency of Planktonic DNA Archiving in Recent Sedimentary Deposits**

We found that 71 % of the PUs from the water column were detected in the sediments, showing great potential for the application of molecular tools to DNA archived in lake sediments for the reconstruction of the past protistan diversity.

The highest similarity between the water and sediment DNA signatures was obtained from the oldest sediment samples analyzed (deposits from 1950 to 1951). This is consistent with what we know about the trophic status of the lake; the lake is currently oligo-mesotrophic, as it was in the early 1950s (Berthon *et al.* 2013, Savichtcheva *et al.* 2015). However, not all of the PUs detected in the water column were retrieved in the sediment. Indeed, we observed that the diversity of Cryptophyta and Haptophyta was poorly preserved. Cryptophyta were relatively diverse (21 PUs) and recurrent in water layers, while only 5 of those PUs were retrieved from sediment. Similarly, out of the 12 PUs of Haptophyta recorded in water samples, only 2 PUs were archived in sediment.

The following two main factors can explain the low efficiency of preservation: (i) degradation that occurs during sedimentation along the water column, due to grazing or cells' lyses (released DNA is more easily fragmented and the short fragments were not efficiently amplified in our PCR reactions) and (ii) sedimentary DNA is degraded during burial in the sediment (Dell'Anno *et al.* 2002, Corinaldelsi *et al.* 2008, 2011, Boere *et al.* 2011a).

For both Cryptophyta and Haptophyta, we verified that the loss of diversity in archived DNA did not occur during sedimentation along the water column. Indeed, all Cryptophyta and Haptophyta PUs occurring at 2 m depth were also retrieved at 130 m depth, showing that the loss of diversity primarily occurred at the sediment interface and during burial (after deposition on sediment).

The way DNA arrives on sediments directly influences the efficiency of preservation during burial; DNA can be protected in intact cells or, in contrast, can be rapidly released as extracellular DNA, which is more prone to fragmentation/degradation (Corinaldelsi *et al.* 2011). A key parameter of the preservation of cells and DNA during burial in sediment is linked to the morphological features of the taxa (i.e. cells with simple membranes or resistant cells, such as cysts and spores). Selective preservation of DNA in cysts has, for example, been shown in Holocene sediment from an Antarctic fjord (Boere *et al.* 2009).

From our data, we could not efficiently study the link between preservation rates in sediment and the morphological features of the PUs because a large number of PUs were not precisely characterized (morphology, ability to produce cysts, etc.). We can, however, mention that Cryptophyta cells are known to be quite fragile because their simple membranes may be easily disrupted, and this may influence the level of preservation of their DNA in sediment. As for Haptophyta, since DNA sequences were recovered in very ancient sediments (Lake Ace from Antarctica, up to 10,450 years old; Lake Van from Turkey, over 270,000 years) (Coolen *et al.* 2004, Randlett *et al.* 2014), it is rather clear that Haptophyta DNA can be preserved in the paleosignal, as already shown by Coolen *et al.* (2013); however, our data suggest that some taxa within Haptophyta may be less well preserved, inducing a potential bias in the estimation of their taxonomic richness.

The conclusions regarding these missing taxa must be tempered in view of the depth of the sequencing effort we applied. Our observations suggest that high recurrence and abundance of protistan taxa in the water column could represent favorable factors for archiving in sediment. We detected a positive link between the recurrence and abundance of PUs in the water and their presence in sedimentary DNA, implying that rare taxa are not easily found in sedimentary DNA; this is probably particularly true when the sequencing effort is moderate.

We assume that a larger proportion of the rare taxa (rare in water) could have been detected in the sediments if the sequencing depth was higher; this includes Haptophyta and Cryptophyta, which were not dominant in the water column. The 71 % recovery of PUs may have been higher if ultra high-throughput sequencing (Illumina MiSeq, for instance) had been used instead of 454 pyrosequencing. In addition the limited amount of sediment used in our study (3 replicates of 1 g for each sediment layer, with pooled DNA extracts) could also explain that the rarest taxa were incompletely detected, we cannot exclude that a larger amount of sediment or separated analyses of replicated DNA extractions may have improved the detection of rarest taxa as suggested by Billard *et al.* (2015).

This work provides some clues regarding how well sedimentary DNA reflects the huge diversity of planktonic protists and confirms that the application of sequencing techniques to sedimentary DNA can be used to reconstruct past diversity for numerous planktonic groups. While previous studies have focused on specific groups (e.g., Bacillariophyta or Cyanobacteria), the possibility of studying complex assemblages involving various functional groups (e.g., bacterivores such as Ciliophora, photoautotrophs such as Chlorophyta, parasites such as Chytridiomycota, etc.) opens up new avenues of research in paleolimnology. A view of such large biodiversity offers, for instance, opportunities to analyze ecological networks as it was recently done for modern microbial assemblages (Barberán *et al.* 2012, Chow *et al.* 2014), or to verify the existence of tipping points within the community structure due to environmental forcing factors (Coolen *et al.* 2013).

## II.6. Supporting information

Figure II.S1. Supporting information presenting the flow chart for bioinformatic analyses using PANAM pipeline (Taïb *et al.*, 2013), figure modified from Debroas *et al.* (2015). 1- Cleaning procedure as described in material and methods section (reference for PANGAEA: Giongo *et al.* 2010). 2- Clustering reads into OTUs with a cut-off of 95 %. 3- OTUs are compared against the reference database. The query OTUs are sorted according to the taxonomy of their best hits among eukaryotes. Several files are generated, containing the reads and their 5 best hits, assigned to one of the 25 specific monophyletic groups among the Eukaryota. 4- After reads have been assigned to phyletic groups, they are aligned to the reference sequences of the corresponding profile alignment for that group using hmmlalign from the HMMER package (Edgar 2010). 5- Using FastTree (Price *et al.* 2010), a bootstrapped phylogenetic tree (100 iterations) is built for each phyletic profile, including OTUs associated with their 5 best hits and the reference sequences in the profile. 6- Two methods for taxonomy assessment are implemented: lowest common ancestor (LCA) and nearest neighbor (NN) (Paradis *et al.* 2004; Kembel *et al.* 2010). 7- Simplified representation of a PU: In this example, 3 OTUs are gathered in PU\_1 with a bootstrap equal to 99%. The nearest neighbor among the reference sequences in the tree, AJ535144, is affiliated to the *Rossia* sp. (Bacillariophyceae) and this taxonomy is propagated to all the OTUs of this PU.

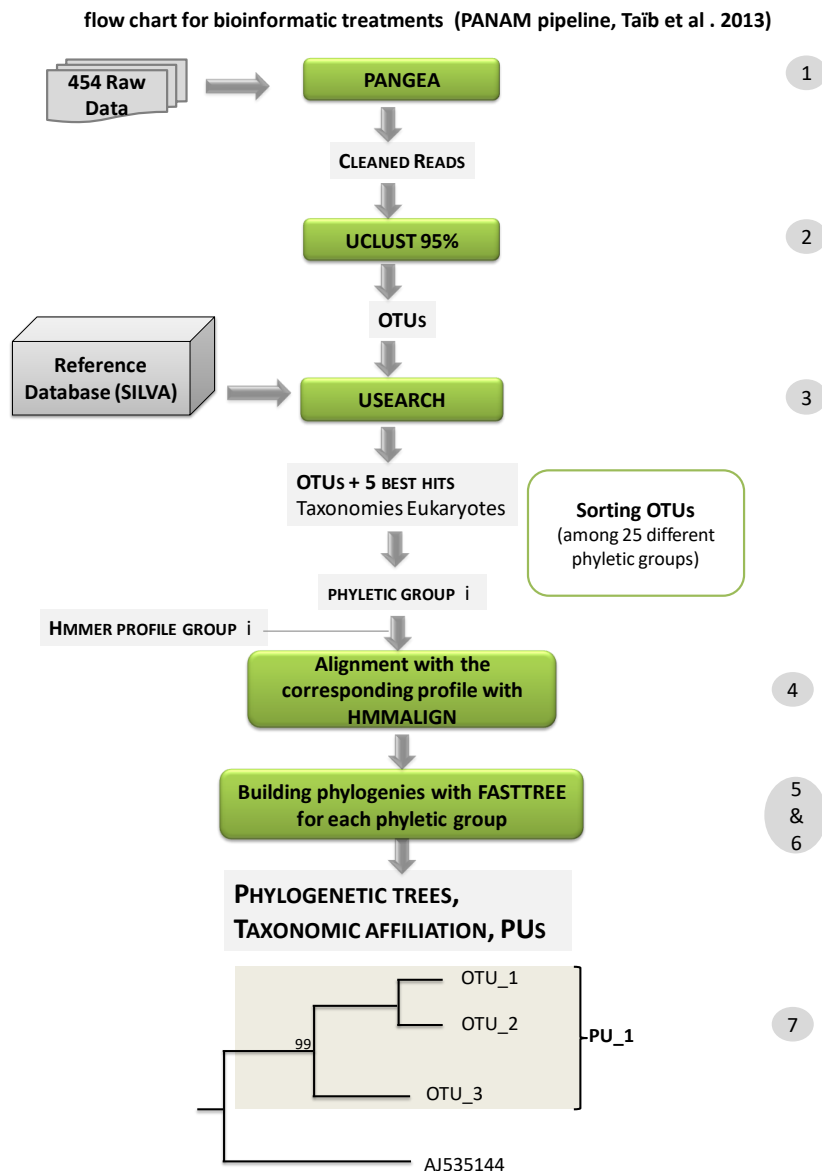


Table II.S1. Number of sequences, OTUs, PUs, and values of Chao index associated to each sample.

Water samples	Sampling Date	Depth (m)	#Sequences	#OTUs	#PUs	Chao1
A2m	March 2011	2	7457	993	253	2139
B2m	May 2011	2	3069	570	185	1249
C2m	June 2011	2	4470	710	223	1544
D2m	July 2011	2	7270	1221	277	2688
E2m	August 2011	2	4508	737	248	1675
F2m	September 2011	2	5387	605	199	1113
G2m	October 2011	2	5483	908	245	1908
H2m	November 2011	2	5288	968	255	2225
I2m	January 2012	2	5149	1074	308	3112
J2m	February 2012	2	5482	947	297	2846
K2m	April 2012	2	9360	1630	314	4376
L2m	May 2012	2	7288	1157	262	3006
M2m	June 2012	2	13045	1626	251	4882
N2m	July 2012	2	11979	1428	274	3780
O2m	September 2012	2	12149	1571	337	3534
A130m	March 2011	130	7483	914	239	2004
B130m	May 2011	130	9963	1398	274	3085
C130m	June 2011	130	11521	1538	315	3297
D130m	July 2011	130	8404	866	234	1748
E130m	August 2011	130	9927	1383	319	3098
F130m	September 2011	130	5620	1022	227	2139
G130m	October 2011	130	7708	1442	275	3085
H130m	November 2011	130	5409	968	222	1880
I130m	January 2012	130	14929	1985	355	5517
J130m	February 2012	130	24572	3521	475	11762
K130m	April 2012	130	13257	1451	274	4913
L130m	May 2012	130	7870	1157	290	3919
M130m	June 2012	130	9857	1482	309	4685
N130m	July 2012	130	10800	1586	351	4997
O130m	September 2012	130	10299	1575	329	4996

Sediment samples	Dated strata	Depth (cm)	#Sequences	#OTUs	#PUs	Chao1
S2009	2008-2009	1	15903	2536	476	7035
S2003	2003-2004	3	16141	2509	426	7738
S2000	2000-2001	4	12188	1754	349	5441
S1972	1972-1973	20	18001	2493	355	8110
S1969	1969-1970	21	19132	2465	413	7892
S1951	1951-1952	27	22226	2801	478	8177



Table II.S2. Number of reads obtained for the different phyla in water at 2 m and 130 m depth and in sediment samples.

Phylum	Water		Sediments
	2m depth	130m depth	
Alveolata	35841	68668	34639
Amoebozoa	1753	15020	246
Cryptophyta	1085	1776	70
Euglenozoa	420	121	67
Fornicata	690	1523	502
Haptophyceae	3858	1143	29
Heterolobosea	17	23	17
Opisthokonta	14295	18661	15545
Parabasalia	38	40	29
Rhizaria	5281	5127	4117
Rhodophytes	16	4	2
Stramenopiles	14559	12168	6947
Viridiplantae	12580	4847	11149
<i>Sum</i>	<i>90433</i>	<i>129121</i>	<i>73359</i>

Table II.S3. (simplified): Number of PUs and percentage of shared PUs between water and sediment samples: (A) number of PUs in all samples, (B) number of PUs in sediment samples, (C) number of PUs in all water samples and percentage of these PUs retrieved in sediments, (D- E) for each water layer (i.e. 2 m and 130 m depth) the table presents the total number of PUs (# PUs), the number of PUs found specifically during mixing or stratification (# specific PUs), and the percentage of these ‘period-specific’ PUs that is retrieved in sediment. Uncult: uncultured ES: environmental samples

Taxonomic groups	(A) All samples #PUs	(B) Sediments all #PUs	(C) Water all		(D) 2m depth		(E) 130m depth			
			#PUs	% PUs retrieved in sediments	#PUs	% PUs specific #PUs	#PUs	% PUs specific #PUs	% PUs retrieved in sediments	
Alveolata	303	232	270	74	223	36	69	234	47	53
Amoebozoa	19	12	17	59	13	0	Ø	17	4	25
Cryptophyta	22	6	21	24	19	0	Ø	21	2	50
Euglenozoa	5	4	5	80	4	0	Ø	5	1	0
Fornicata	3	3	3	100	3	0	Ø	3	0	Ø
Haptophyceae	12	2	12	17	9	1	0	11	3	0
Heterolobosea	3	2	3	67	1	0	Ø	3	2	50
Opisthokonta	337	246	302	70	236	55	79	274	127	56
Parabasalia	3	3	3	100	2	0	Ø	3	1	100
Rhizaria	59	50	53	83	43	2	50	51	10	50
Rhodophytes	2	2	2	100	2	0	Ø	2	0	Ø
Stramenopiles	201	140	174	65	148	19	58	155	26	50
Viridiplantae	108	96	89	87	76	17	65	71	13	92

# III. Préservation de l'ADN sédimentaire au cours de l'enfouissement

**To what extent is the DNA of microbial eukaryotes modified during burying into lake sediments? A repeat-coring approach on annually laminated sediments.**

Papier en préparation pour soumission dans *Journal of Paleolimnology*

Eric Capo<sup>1,2</sup>, Christian Bigler<sup>3</sup>, Didier Debroas<sup>4</sup>, Isabelle Domaizon<sup>1,2</sup> \*

<sup>1</sup> INRA, UMR 42 CARRTEL, 74203 Thonon-les-bains Cedex, France

<sup>2</sup> Université Savoie Mont Blanc, UMR 42 CARRTEL, 73379 Le Bourget du Lac, France

<sup>3</sup> Department of Ecology and Environmental Science, Umeå University, Umeå, Sweden

<sup>4</sup> Clermont Université, Université Blaise Pascal, CNRS, UMR 6023 LMGE, 63171 Aubière, France

## **Contribution des auteurs :**

- Conception de l'étude: ID, CB, EC
- Acquisition des données: EC, ID, CB
- Traitements bioinformatiques: DD
- Analyses et interprétations des données: ED, ID, CB, DD
- Ecriture du papier: ID, EC, CB, DD

## **Résumé**

L'objectif de ces travaux est d'estimer l'impact de la diagénèse précoce sur la quantité et la qualité du signal ADN et plus particulièrement sur la richesse, la composition et la structure des communautés microbiennes eucaryotes. La stratégie d'étude a été de comparer les inventaires moléculaires de 12 strates sédimentaires (datées entre 1973 et 2006) provenant de deux carottes annuellement varvées qui ont été prélevées dans le lac Nylandssjön à 6 années d'intervalle (2007 et 2013). Les principaux résultats issus du séquençage de la région V7 du gène codant l'ARNr 18S ont montré que la quantité d'ADN extrait des varves sédimentaires, la richesse totale et la diversité (indices de Shannon, Simpson, équitabilité) de la communauté microbienne eucaryote n'ont pas été significativement modifiées au cours des 6 années d'enfouissement. Par ailleurs, les dynamiques temporelles des communautés de micro-eucaryotes obtenues à partir des inventaires moléculaires des 2 carottes sédimentaires sont congruentes. L'analyse menée sur les différents de microbes eucaryotes a montré que seul le groupe des Champignons présente des différences significatives entre les 2 carottes, avec une modification significative à la fois de leur proportion en nombre de séquences ADN et de leur

richesse. Ces variations peuvent être liées à des processus diagénétiques, la variabilité spatiale entre les deux carottes et également à la présence d'une biosphère fongique active dans ces sédiments superficiels. Les Champignons représentant une part mineure de la communauté eucaryote (en proportion relative et richesse), ces variations sont peu impactantes au niveau de la communauté eucaryote globale. Enfin, l'analyse de dissimilarité basée sur les proportions relatives des OTUs montre que le signal ADN obtenu pour les deux carottes est très comparable et semble stabilisé pour toutes les strates âgées de plus de 15 ans ; toutefois on détecte une variation dans les proportions relatives de certains OTUs (Chlorophyta, unclassified Stramenopiles et Alveolata, Champignons) pour les strates les plus récentes. Ce résultat suggère un potentiel effet des processus diagénétiques dans les 7 premiers cm de ces sédiments.

### **III.1. Abstract**

Paleogenetics provides a powerful framework to reconstruct the long-term temporal dynamics of various biological groups from aquatic sediments. However, validations are still required to ensure the authenticity of the molecular signal obtained from sedimentary DNA. In this study, we investigated the effects of early diagenesis between cores on the DNA signal from micro-eukaryotes preserved in sediments as well as the effects of spatial heterogeneity in between cores of this DNA signal. We used two cores sampled in 2007 and 2013. We detected no decline in DNA quantity due to the 6 supplementary years of archiving. No significant difference was detected between the molecular inventories obtained for the two cores both for total richness (number of OTUs, Chao estimator), diversity indices (Shannon and Simpson indices) and evenness values (Pielou index). Moreover, the temporal dynamics of communities (as revealed by nMDS analysis) obtained for the two cores were congruent, showing no distorted picture of the community from one core to another. Though most of the eukaryotic groups showed no significant difference in terms of richness and relative proportion according to the core, only fungi were found to differ with an increase in richness and proportion of fungal OTUs in the core sampled in 2013. These differences could be due to diagenetic processes, spatial heterogeneity between cores but also to the presence of active fungi on these sediment layers. Considering the OTUs level, and based on dissimilarity indices, our results showed that, for the older strata (age > 15 years, depth > 7 cm), the composition and structure of communities were very similar for the two cores (exception for fungi) and the DNA signal was considered stable. However, for the uppermost strata (age < 15 years; depth < 7cm), changes were detected in the relative abundance of few OTUs (Chlorophyta, Unclassified Stramenopiles and Alveolata, Fungi). This suggests that diagenetic processes might affect molecular inventories more particularly in the uppermost sediment layers.

### **III.2. Introduction**

Paleogenetics studies have recently revealed the potential of the DNA preserved in lacustrine sediments to reconstruct the temporal dynamics of biological communities at long time scales (centennial to millennial) (e.g. Domaizon *et al.* 2013, Savichtcheva *et al.* 2014, Epp *et al.* 2011, 2015, Stoof-Leichsenring *et al.* 2012, 2015, Kyle *et al.* 2015a, 2015b, Capo *et al.* 2015,

in press). The expansion of paleolimnology to include DNA-based approaches on a more routine basis has the potential to challenge existing limits within this scientific field. In particular the ability to reconstruct a large variety of past eukaryotic groups such as Chlorophyta, Haptophyceae, Ciliophora, Cercozoa that were not previously considered in lacustrine sediment record analyses, opens up many new opportunities for the reconstruction of taxa that do not leave distinct morphological fossils. For instance more than 10 thousands of OTUs distributed within 50 phylogenetic groups were found from dated sediment of lakes Bourget (France) and Igaliku (Greenland) (Capo *et al.* in press). These results allowed illustrating long-term temporal dynamics in the communities of microbial eukaryotes (i.e. protists and Fungi) which are key components of lacustrine systems playing roles as primary producers, phagotrophs, saprotrophs, and parasites.

Although these approaches are promising for further revolutionizing our understanding of long-term ecological dynamics, careful calibration studies are still needed, as with any paleolimnological proxy. Some knowledge exists about the extent to which DNA is preserved in sediment and the factors and the conditions that control its preservation (e.g. Dell'Anno *et al.* 2002). Anoxic conditions, absence of bioturbation, and low temperature are among the most crucial factors promoting DNA preservation (Coolen & Gibson 2009, Boere *et al.* 2011a). Despite these advances in understanding, there is no systematic study that has quantified the degree to which DNA preservation is optimized in sediment (either in the field or in the lab) especially considering the diverse origins of sedimentary DNA (different biological groups, different types of DNA i.e. extracellular DNA or DNA protected in cells may be differently affected by degradation).

Modifications of the DNA from the lake community can occur (i) during sedimentation in the water column, (ii) after deposition at the sediment interface and during burying in the sediments (Dell'Anno *et al.* 2002, Dell'Anno & Corinaldesi 2004, Corinaldesi *et al.* 2008, 2011, Torti *et al.* 2015). Most of studies about the DNA preservation in sediments were focused on the degradation of extracellular DNA, these studies showed that the decays of DNA in sediments (degradation of nucleotidic acids and fragmentation of DNA sequences) occurred at long –term time scale (Torti *et al.* 2015). A high fragmentation of microbial eukaryotic DNA was for instance detected in 2700 years-old marine sediments Boere *et al.* (2011a) and Kirkpatrick *et al.* (2016) reported that chloroplastic DNA declines after 100 000 years in marine sediments.

However, there is still a lack of knowledge about the effect of early diagenetic processes (first years of archiving) on the DNA signal, in particular regarding the potential differential preservation of the large diversity of eukaryotic taxa and the subsequent impact on the relevance of molecular inventories produced from sedimentary DNA. Early sedimentary diagenesis has been studied in laboratory incubation experiments and *in situ* using sediment traps, but never targeting DNA. Laboratory experiments applied to estimate the decay of algal (eukaryotic) material have reported high variability in organic matter decay rates according to the algal material, incubation temperatures and the decay models used (Lehmann *et al.* 2002).

A few studies have also compared lake sediment cores taken on different dates, in this case between-core differences were used to estimate declines in carbon and nitrogen rates during the first years of burying in the sediment of shallow lakes (Nylandssjön, Sweden; Gälman *et al.* 2008, Rydberg & Martinez-Cortizas 2014). This “repeat-coring” approach demonstrated its potential to provide knowledge on how diagenesis affects the sediment record over decadal time scales; varved lake sediment indeed provided a robust framework to compare material deposited during exactly the same time period (Gälman *et al.* 2006, 2008, 2009a, 2009b, Rydberg *et al.* 2008, Maier *et al.* 2013). These studies demonstrated that the C concentration of the sediment decreased by 20 % and N decreased by 30 % within the first five years after deposition suggesting critical modifications due to early diagenesis (Nylandssjön, Sweden, Gälman *et al.* 2008). In the same lake, Maier *et al.* (2013) confirmed that compaction rate was greatest during the first few years after deposition, with varve thickness decreasing by 60 % after 5 years (sediment depth 5 cm). Diagenetic organic matter transformations at the oxic/anoxic interface in aquatic sediments are indeed known to occur within the first cm of sediment where organic compounds and associated nutrients are mineralized through heterotrophic decomposition. Depending on the environmental characteristics of each lake, the active zone of microbial decomposition varied (notably in relation with oxygen penetration depths in sediments, from millimeters in productive hypoxic systems to cm or dm scales in oligotrophic oxygenated systems, Kristensen 2000) and have serious implications for the rates of organic matter diagenesis in sediments.

In the present study, we used 2 cores collected from Nylandssjön, which has distinctly annually laminated sediments (Renberg 1986). The cores were sampled with a time lag of 6 years, in 2007 and 2013, and stored intact in a freezer room from coring until they were sub-sampled and analysed. The collection cores from this lake was previously used to test the stability of various elements during sediment diagenesis: total carbon and nitrogen and their

stable isotopes (Gälman et al. 2008, 2009a), iron and sulfur (Gälman et al. 2009b), mercury and methyl-mercury (Rydberg et al. 2008), and sediment compaction (Maier et al. 2013). The aim was here to study the effect of early diagenesis on DNA and how this impacts the molecular inventory obtained for microbial eukaryotes. The influence of diagenetic processes, occurring during this 6 years time lag, were followed on 12 sediment strata which were selected in a time window of 40 years. The comparison between the 2 cores allowed us to determine (i) if the DNA quantity of individual varves changed with time (6 years apart) (ii) if the richness, the composition and the structure of microbial eukaryotes communities varied under the effect of early diagenesis (iii) if the response to diagenesis was group dependent, i.e. differential preservation according to the taxonomic groups.

### **III.3. Materials & Methods**

#### **III.3.1. Study site and sediment sampling**

Nylandssjön (62\_570N, 18\_170E) is a circum-neutral, mesotrophic, and dimictic lake that preserves distinct annual sediment laminations, or varves (Renberg 1986, Gälman *et al.* 2006, Maier *et al.* 2013). It is a headwater lake located 34 m above sea level, with an area of 0.28 km<sup>2</sup> and a catchment of 0.95 km<sup>2</sup>. At present, about 16 % of the catchment consists of agricultural land, whereas the rest is managed boreal forest. In the deepest part of the lake (17.5 m), two sediment cores were sampled during late winter in 2007 (core C) and 2013 (core C+6) using a freeze corer (Renberg 1981). Following sampling, the freeze cores were stored at - 18 °C until further analyzed. Because core sampling was done during late winter, the topmost varve in each core was deposited mostly during the year prior to sampling. Unlike common lake sediments, which are mixed and require dating by radiometric methods that involve a degree of uncertainty, the varves provide an accurate timescale. The two sediment cores were manually sub-sectioned into annual samples along the black winter layer using a steel scalpel and a magnifying lamp (Renberg 1981) and following the required methodological precautions to ensure the non-contamination by modern DNA (sterile disposable labware, at -18°C). For each core (C and C+6), 12 varves dated in 1973, 1976, 1979, 1982, 1985, 1988, 2000, 2003 and 2006 were then selected for further analyzes.

### III.3.2. Molecular data

For each sample (12 from core C and 12 from core C+6), two DNA extractions were performed on 0.5 g of wet sediment by using the NucleoSpin® Soil kit according to the manufacturer instructions (Macherey-Nagel, Düren, Germany). The concentration of bulk DNA was estimated using a Nanodrop ND-1000 Spectrophotometer (Thermo Scientific, Wilmington, DE, USA). The V7 region of 18S rRNA gene was PCR amplified by targeting a 260 bp long fragment, from c.a. 25 ng of environmental DNA extract for each sample and using the general eukaryotic primers 960f (5'-GGCTTAATTTGACTCAACRCG-3') (Gast *et al.* 2004) and NSR1438 (5'-GGGCATCACAGACCTGTTAT-3') (Van de Peer *et al.* 2000). Each PCR was performed in duplicate in a total volume of 25 µl containing 3µl of 10X NH<sub>4</sub> reaction buffer, 1.2 µl of 50 mM MgCl<sub>2</sub> solution, 0.25 µl of BioTaq (Bioline), 0.24 µl of 10 mM dNTP, 0.36 µl of 0.5 µg/µl BSA and 1 µl of each primer. The amplification conditions consisted of an initial denaturation at 94 °C for 10 min followed by 35 cycles of 1 min at 94 °C, 1 min at 55 °C and 1 min 30 s at 72 °C. The amplicons were then subjected to a final 10-min extension at 72 °C. The two PCR products obtained for a given DNA extract were pooled and purified using Illustra™ GFX™ PCR DNA and Gel Band Purification Kit (GE Healthcare Life Sciences, Velizy Villacoublay, France) and quantified using Quant-it PicoGreen kit (Invitrogen, Carlsbad, CA, USA) on a Fluoroskan Ascent™ FL (ThermoLabsystems, Beverly, MA, USA). Finally, the 48 purified amplicons (12 samples per core, each treated in duplicates) were pooled at equimolar concentrations and sent to Fasteris SA (Geneva, Switzerland) for library preparation and paired-end (2 x 250 bp) sequencing on a MiSeq Illumina instrument (San Diego, CA, USA).

All laboratory procedures (sub-sampling, DNA extraction, PCR amplifications) were carried out in a dedicated laboratory, following strict precautions to ensure the authenticity of the results including the work in sterile conditions the use of blanks during the sub-sampling step (opened tube containing ultrapure water in order to detect any aerial contamination, DNA extraction and PCR amplification (tube with ultrapure water).

### III.3.3. Bioinformatics and statistical analysis

The paired-end reads were merged together using UPARSE tools (option -fastq\_mergepairs with a minimal overlap equal to 150 and no mismatch allowed) (Edgar, 2013) allowing to obtain 4 496 749 raw sequences. These sequences were then submitted to three cleaning procedures: (i) no undefined bases (Ns), (ii) a minimum sequence length of 200 bp, and (iii)



no sequencing error in the forward and reverse primers. Putative chimeras were detected by UCHIME (Edgar *et al.* 2011). The combining of these cleaning procedures with a sequencing depth (2 x 250 bp) covering all the amplicon length (~ 220 bp without the primers) has allowed to sequence twice each base and to minimize drastically the sequencing errors. After this cleaning step and demultiplexing process, the remaining sequences were clustered at a 95 % similarity threshold with UPARSE 7.0 (Edgar 2013) to obtain the seed OTUs (option -cluster\_fast). The delineating of the similarity threshold was chosen following the recommendations from Mangot *et al.* (2013) who estimated the errors during PCR and sequencing by introducing an internal 18S standard (a unique alien DNA sequence) and demonstrated that 95 % was a relevant cut-off threshold for OTUs delineation allowing to limit the artificial inflation of OTUs number. In addition, this threshold was verified by sequencing a mock community in the same experimental conditions (primers, sequencing platform and bioinformatics pipelines); as reported in Capo *et al.* (in press), we obtained 16 OTUs from the 16 species of the mock community when applying the 95% threshold. The taxonomic affiliation was performed by BLAST against the SSURef SILVA database (Pruesse *et al.* 2007) after application of the following selection criteria: length > 1200 bp, quality score > 75 % and a p-intail value > 50; SILVA database was enriched by lacustrine DNA sequences originating from various studies on lacustrine systems (Lefranc *et al.* 2005, Lepère *et al.* 2006, 2008, Mangot *et al.* 2013). The taxonomy of an OTU corresponded to the best hit given by the similarity search. If an OTU was associated with several best hits (hits having the same identity), then the taxonomy was the common taxonomy of these hits named also the lowest common ancestor. In addition, a specific taxonomy based on previous publications (Lefranc *et al.* 2005, Lepère *et al.* 2006, 2008, Mangot *et al.* 2013) allows to identify typical freshwater clades (e.g. Cryptophyta\_1, 2, Perkinsea\_1). OTUs affiliated to metazoans and embryophytes were removed from these analyzes in order to consider only microbial eukaryotes (protists and Fungi at the basis of trophic networks). After these different steps (cleaning, clustering and taxonomic affiliation), 2 119 423 DNA sequences were clustered in 17 428 OTUs classified as microbial eukaryotes. For the further analyzes, the 48 molecular inventories were normalized at 5,497 reads. The data (demultiplexed sequences) will be deposited in the Dryad Digital Repository (<http://datadryad.org/>).

### III.3.4. Data treatment and statistical analyses

Richness (Chao1), diversity (Shannon, Simpson) and evenness (Pielou) indices were computed from the OTUs table separately for each sediment layer using the R package “vegan” (Oksanen *et al.* 2015). ANOVA analysis and t-test were used to test the difference between the two cores in terms of DNA quantity, total richness and diversity indices; statistical analysis we performed using XLSTAT software.

Comparisons between the molecular inventories obtained for each sample were also based (i) on the richness and relative abundances of each taxonomic groups (ii) and at the OTU level, on the values of dissimilarity obtained for Bray-Curtis index (considering the relative abundance of each OTU) and for Jaccard index (based on the presence/absence of OTUs).

A hierarchical clustering tree was realized based on the dissimilarity matrices using the *hclust* function from the R package “vegan” (Oksanen *et al.*, 2015). To identify the OTUs that contributed to the dissimilarity between the molecular inventories obtained for the two cores (dissimilarity between ‘C’ samples and ‘C+6’ samples, i.e. inter-core comparison), SIMPER analyzes were conducted using PAST v.2.11 (Hammer *et al.* 2001).

Finally the community structures obtained respectively for the cores C and C+6 were compared using the Procrustean analysis (PROTEST software package - PROcrustean randomization test). Procrustean analyses superimpose, scale and rotate one data matrix upon the other until an optimal fit is found (Jackson & Harvey 1993, Peres-Neto & Jackson 2001). The PROTEST permutation procedure (1 000 permutations) was used to assess the statistical significance of the Procrustean fit between the two matrices (Peres-Neto & Jackson 2001).

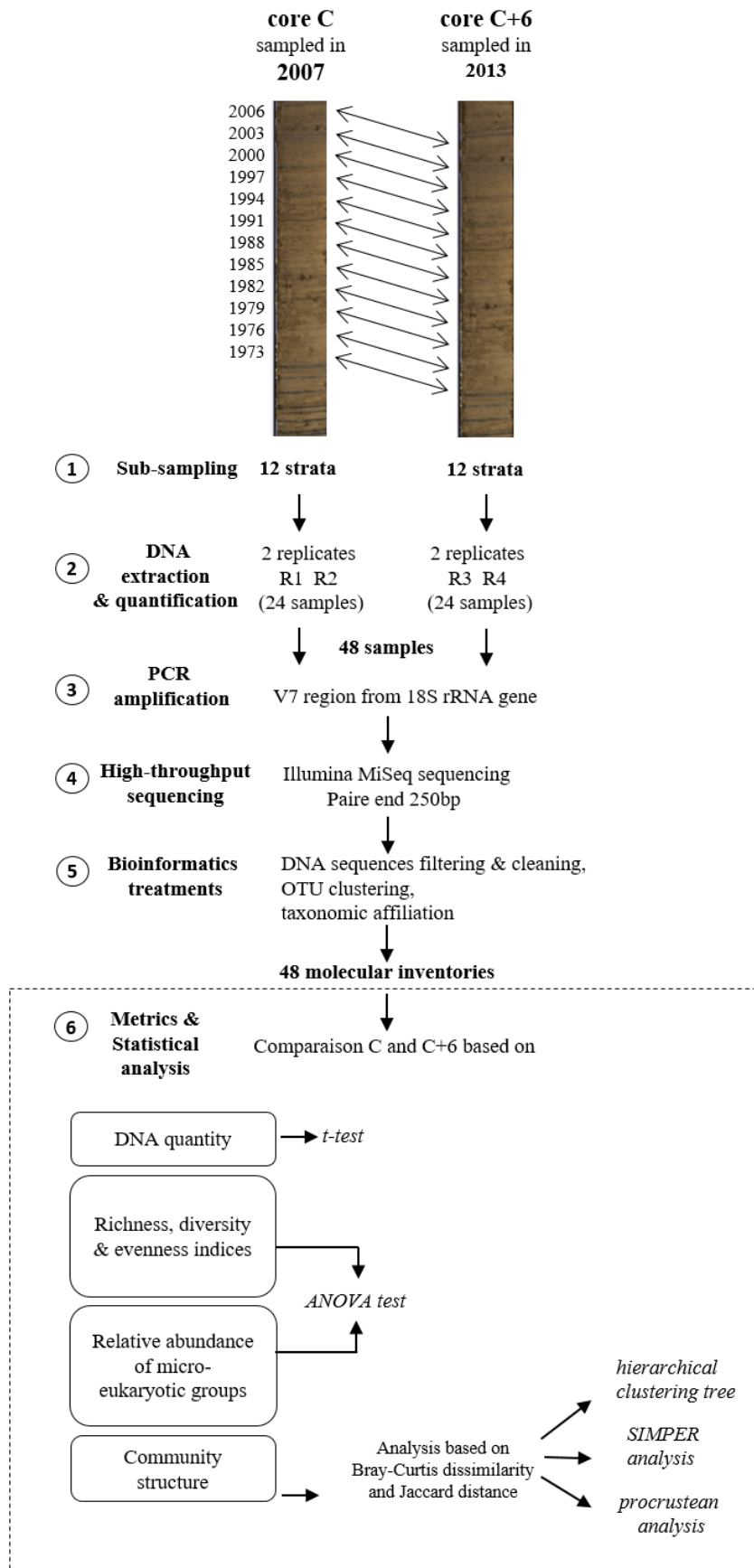


Figure III.1. Experimental procedure used in the present study.

### III.4. Results

Two DNA extractions were performed from each sediment sample (the replicates are named R1 and R2 for the core C, and R3 and R4 for the core C+6). Consequently, 48 molecular inventories were obtained (2 replicates \* 12 varves \* 2 cores; Figure III.1). Replicates from each sediment layer were analysed as separated samples to estimate the intra-core variability in comparison to inter-core variability (C versus C+6).

#### III.4.1. DNA quantities

DNA varied from 21 to 82  $\mu\text{g/g}$  wet sediment (Figure III.2, Figure III.S1). No significant difference was detected comparing the samples from the core C and C+6 (t-test  $p = 0.368$ ).

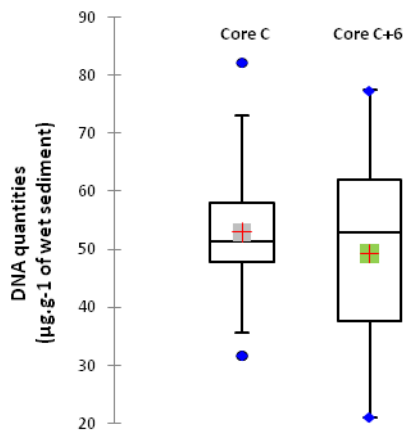


Figure III.2. Boxplots of DNA quantities obtained from the 12 varves in the cores C and C+6. ANOVA test showed no significant differences between C and c+6 ( $p = 0.368$ )

#### III.4.2. Richness and diversity

The 5717 OTUs detected here were distributed within 46 eukaryotic phylogenetic groups (Table III.S1). Most of those OTUs belonged to Chlorophyta (2455 OTUs), Fungi, Chrysophyceae, Ciliophora and Dinophyceae (more than 250 OTUs for each group). A similar number of OTUs was retrieved from the two cores: in total 4078 and 4476 OTUs in the molecular inventories obtained for core C and core C+6, respectively. The number of OTUs retrieved in each sediment layer varied from 425 to 735 OTUs per sample (mean  $581 \pm 96$ ) and from 426 to 990 OTUs ( $653 \pm 155$ ), in C and C+6 cores respectively.

Table III.1. The table presents (i) richness and relative abundance of the main microbial eukaryotic groups (14 most diverse and abundant groups) and (ii) the significance of the comparison test (ANOVA) performed to compare these metrics for C and C+6; differences were considered significant when p-values were inferior or equal to 0.01 (\*: significant p value ; NS: not significant p value).

Taxonomic groups	Richness		Relative abundance	
	#OTUs	p	% reads	p
Apicomplexa	46	NS	833	NS
Bacillariophyta	162	NS	2835	NS
Cercozoa	115	NS	868	NS
Chlorophyta	2455	NS	126763	NS
Chrysophyceae	405	NS	6082	NS
Ciliophora	275	NS	12452	NS
Dinophyceae	342	NS	5295	NS
Fungi	610	0.01 *	6719	0.02
Labyrinthulomycetes	129	NS	6020	NS
Perkinsea	35	NS	305	NS
Synurophyceae	142	NS	3833	NS
unclassified_Alveolata	169	NS	21170	NS
unclassified_Rhizaria	56	NS	739	NS
unclassified_Stramenopiles	557	NS	67597	NS
Other_protists	219	NS	2345	NS
Total	5717		263856	

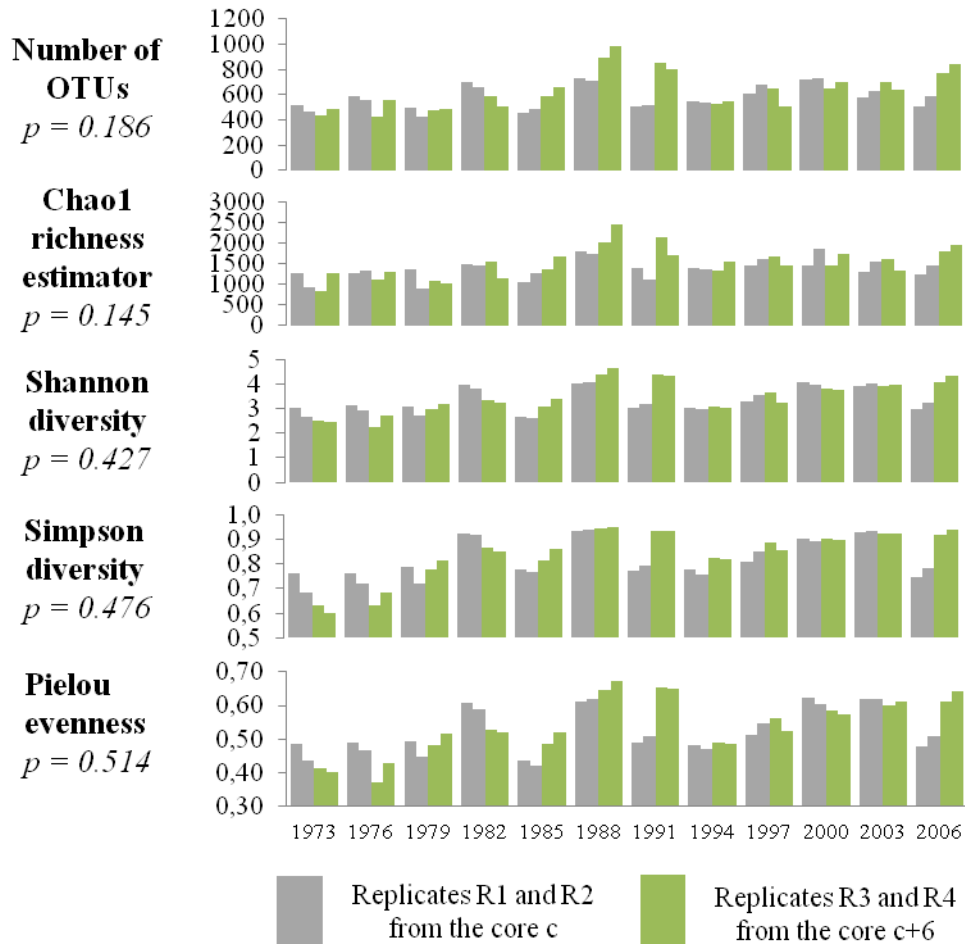


Figure III.3. Barplots of diversity indices: number of OTUs, Chao1 estimator, Shannon diversity, Simpson diversity, Pielou evenness obtained for the 48 samples. ANOVA p-values were calculated for each metric by comparing the values from the cores C and C+6.

Considering the whole dataset, no significant difference was detected between the molecular inventories obtained for the two cores (C and C+6) both for total richness (number of OTUS, Chao estimator), diversity indices (Shannon and Simpson indices) and evenness values (Pielou index) ( $p > 0.1$ ; Figure III.3). This was also verified considering separately the richness of the different taxonomic groups. The comparison of the molecular inventories from the two cores showed no significant difference as regards the richness of the 14 major microbial groups (i.e., the most abundant and diverse groups). Only Fungi, for which a higher richness and relative abundance was detected in the core C+6, differed significantly between the two cores (Table III.1, Figure III.S2). The highest richness of Fungi in core C+6 was associated to the richness increase of several fungal groups including Chytridiomycota, Cryptomycota and Glomeromycota (Figure III.S2).

### III.4.3. Community structure

Bray-Curtis dissimilarity (based relative abundance of OTUs) and Jaccard distance (based on presence/absence of OTUs) were used to compare the intra-core variability against the inter-core variability. As illustrated by the hierarchical clustering trees, the highest similarities were obtained for the two replicates originating from the same sediment layer (R1 & R2 in core C and R3 & R4 in core C+6) (Figure III.4). For samples originating from to the 6 most ancient deposits (strata 1973, 1976, 1979, 1982, 1985 and 1988) the inter-core variability was in the same range as intra-core, revealing that the molecular inventories obtained for C and C+6 were very similar. In contrast, for the more recent deposits (1991, 1994, 1997, 2000, 2003 and 2006), a higher dissimilarity was observed between molecular inventories obtained for C and C+6 (Figure III.4). The highest dissimilarities between C and C+6 molecular inventories were detected from Bray-Curtis indices, while the values obtained for Jaccard indices were less contrasted (between C and C+6). This indicates that when discrepancies were detected between the 2 molecular inventories, they were more particularly due to changes in the relative abundance of OTUs rather than the OTUs composition (i.e., presence/absence of OTUs).

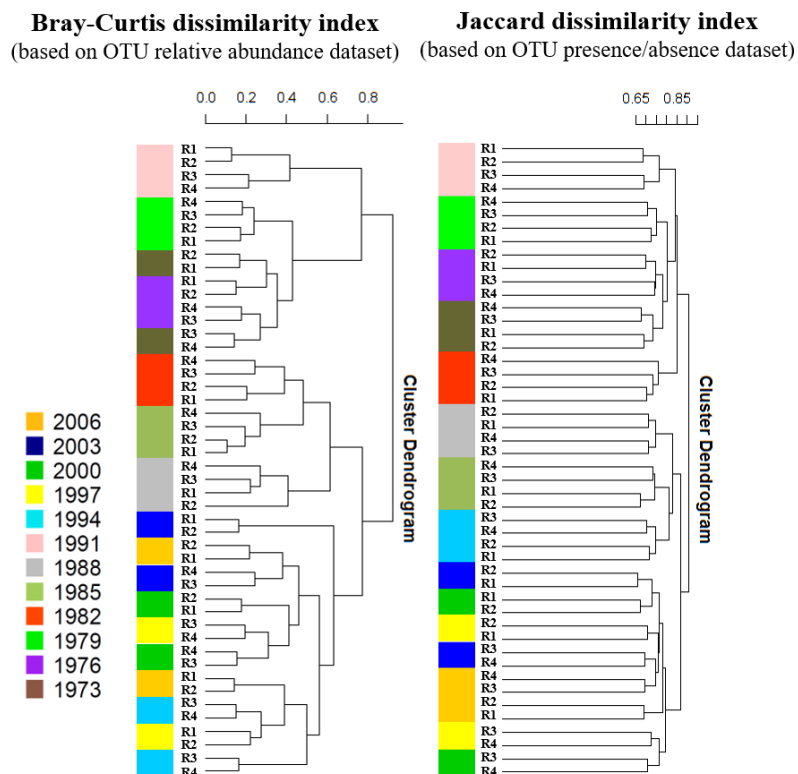


Figure III.4. Hierarchical clustering tree computed using Bray-Curtis index (calculated from the OTU relative abundance dataset) and Jaccard index (calculated from presence/absence dataset) from the 48 molecular inventories.

ANOVA test confirmed the significant difference of Bray-Curtis dissimilarity values obtained for the 6 most recent deposits (C versus C+6 comparison:  $p < 0.0001$ ) (Figure III.5). The modifications in the relative proportions of OTUs thus occurred during the first 15 years of burying (i.e. strata 1991 was aged of 15 years when the core C+6 was sampled in 2013).

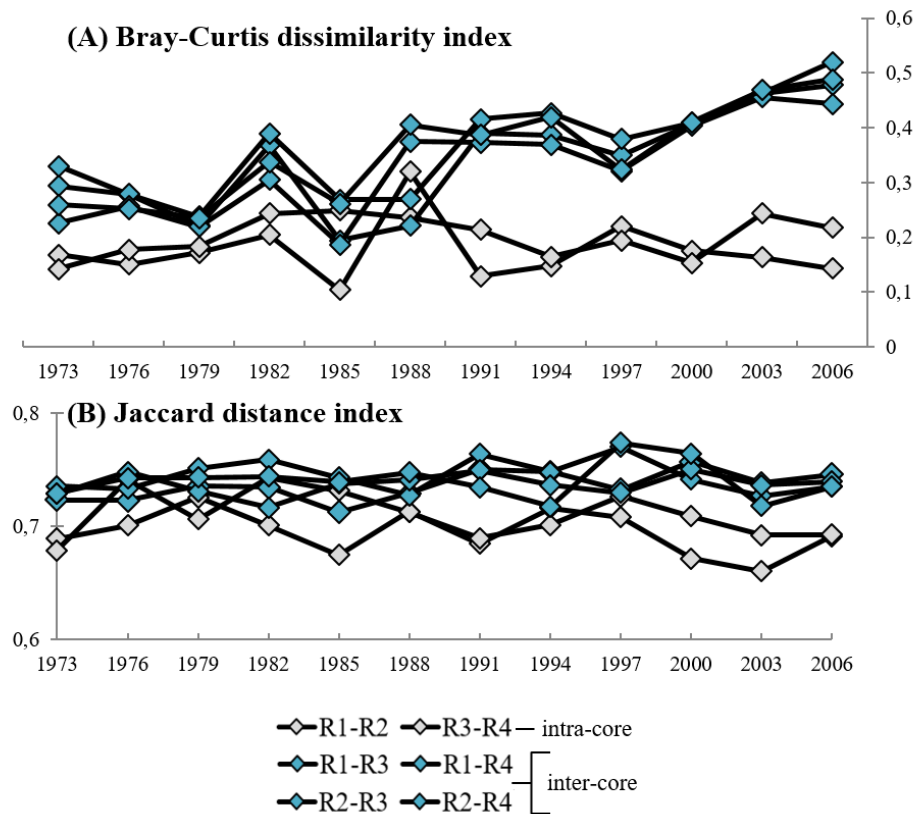


Figure III.5. Values of Bray-Curtis dissimilarity (A) and Jaccard distance (B) obtained for each couple of replicates for the 12 analysed strata. Intra-core comparison is represented by the dissimilarity of R1-R2 (for C core) and R3-R4 (for C+6 core); inter-core comparison is represented by the dissimilarity of R1-R3, R1-R4, R2-R3, R2-R4).



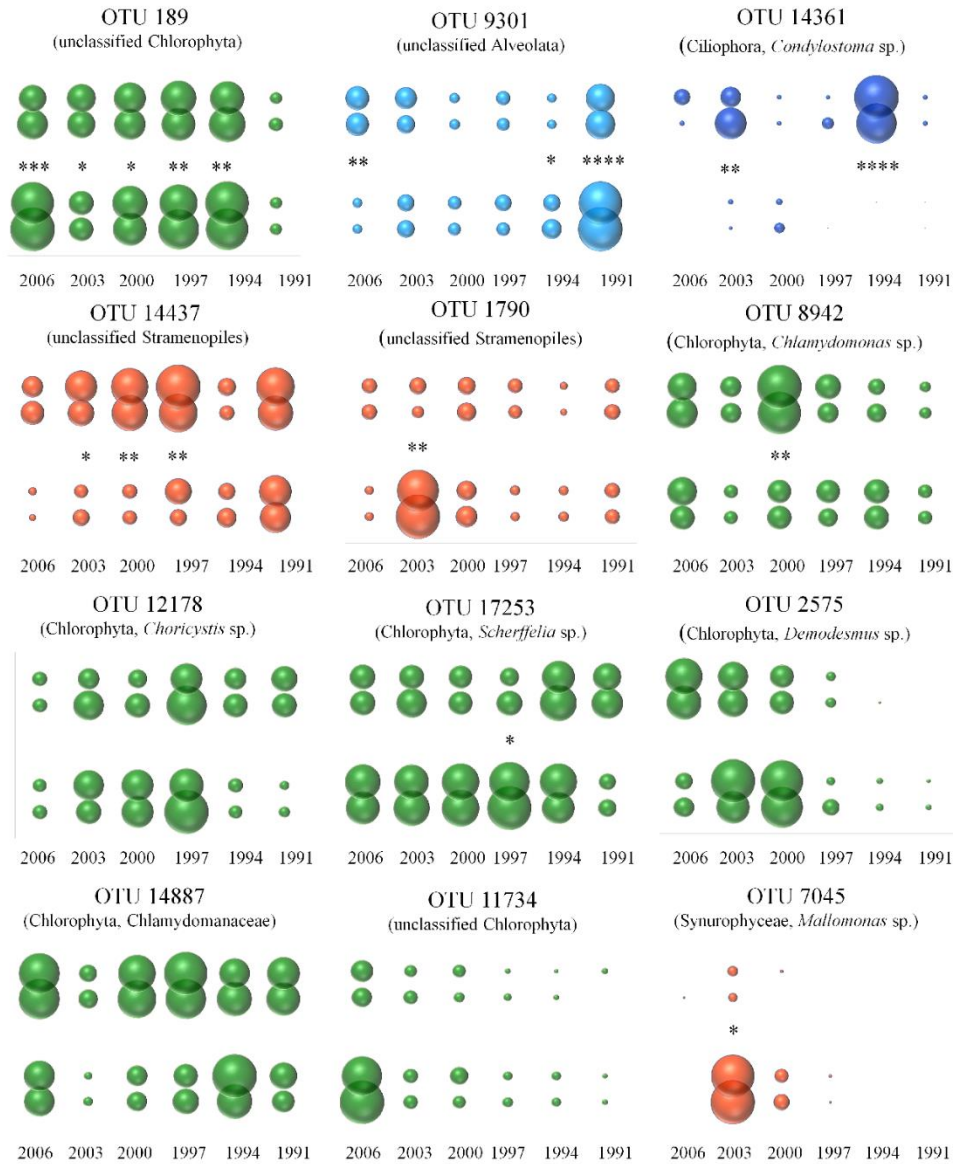


Figure III.6. Relative abundance (number of DNA sequences) of the 12 OTUs that contributed to more than 1 % to the dissimilarity between the molecular inventories obtained for the cores C and C+6, in each of the 12 analysed varves. Percentages of contribution superior to 5 % were illustrated by the following legend: more than 5 %: \*, 10 % \*\*, 20 % \*\*\*, 30 % \*\*\*\*. The color assigned to each OTU was related to their taxonomic affiliation: green for chlorophyтан OTUs, orange for Stramenopiles, blue for unclassified Alveolata and dark blue for the Ciliophora OTU.

In order to detect which OTUs were responsible for the dissimilarity observed between the structure of microbial eukaryotic communities in C and C+6 for these 6 most recent strata, SIMPER analyses were used and showed that only 39 OTUs (out of the 5717 OTUs) explained more ~70 % of the dissimilarity between the composition of microbial eukaryotes communities from the 2 cores (Table **III.S2**). Among these 39 OTUs, the major contributors to dissimilarity belonged to the groups of Chlorophyta, Alveolata (Ciliophora and unclassified OTUs) and Stramenopiles (Table **III.S2**, Figure **III.6**).

The unclassified Stramenopiles OTU (OTU14437) was found in lower amounts (in terms of % of DNA reads) in the core C+6 for the 6 most recent deposits, while for all the other contributor OTUs no clear trend could be observed since they either increased or decreased according to the strata. This was for instance the case for the highest contributor OTUs, i.e. several chlorophytan OTUs, an unclassified Alveolata OTU and ciliophoran OTU (*Condylostoma* sp.) (Figure **III.6**).

These changes detected at the OTUs level did not affect significantly the relative abundance of their respective phylogenetic groups (Figure **III.7**, Table **III.1**) which were not significantly different in the cores C and C+6. Only for Fungi, the p value was close to the significance level ( $p = 0.02$ ) and could reveal a non-negligible increase of their relative abundance in the core C+6 (Table **III.1**, Figure **III.S2**).

Finally, despite these variations in the relative abundance of contributors OTUs, the procrustean analysis showed significant levels of concordance between community structures obtained from the two cores (Figure **III.8**,  $P = 0.001$ ), showing that both cores provided congruent DNA signal as regards the global structure of the eukaryotic community.

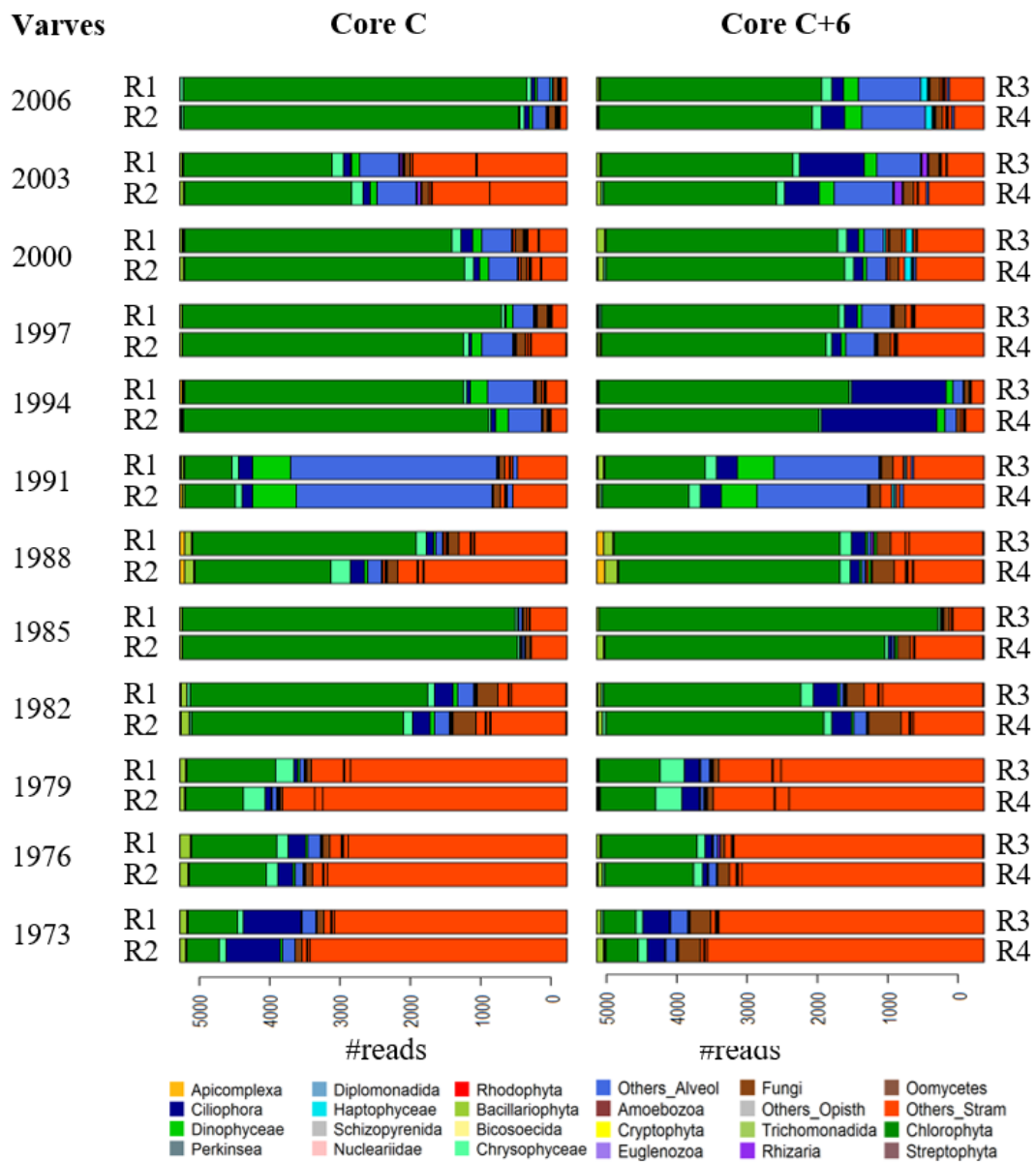


Figure III.7. Distribution of DNA sequences in the molecular inventories with the main microbial eukaryotes groups.

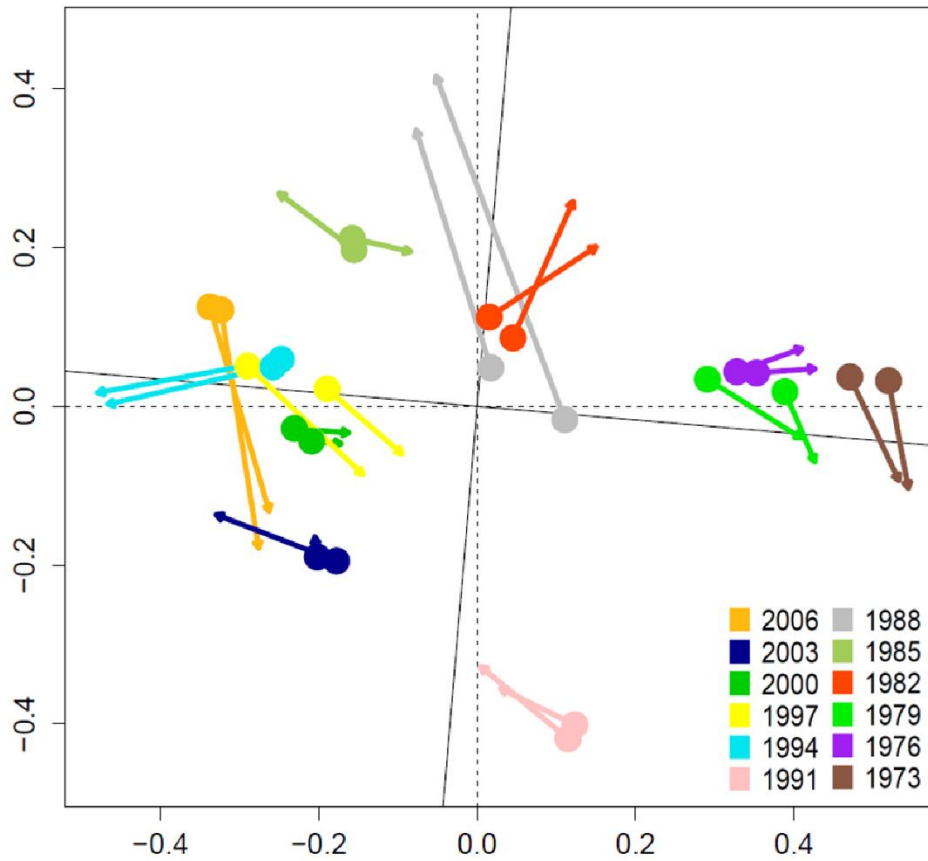


Figure III.8. Plot of procrustean rotation analysis illustrated the potential differences in community structure for molecular inventories from core C and core C+6. Statistics revealed significant levels of concordance between community structures obtained from the two cores.

## III.5. Discussion

Over the last decade, DNA from sediments has been successfully employed to study the succession of various lacustrine and marine photosynthetic microorganisms at long time scales (from 100 years to 270 k years) (Coolen *et al.* 2004, 2013, Boere *et al.* 2011a, Domaizon *et al.* 2013, Hou *et al.* 2014, Randlett *et al.* 2014, Epp *et al.* 2015, Li *et al.* 2016). Although the understanding of archiving processes is obviously crucial to properly reconstruct temporal changes in biological communities from their sedimentary DNA records, this issue has not been sufficiently considered yet. While, as reported by (Behrensmeyer *et al.* 2000), notable advances have been obtained in the domain of taphonomy (for instance as regards microbial and biogeochemical controls on the preservation of different biological tissue types), to date few studies have quantified the influence of environmental factors in regulating final composition of total sedimentary DNA assemblage (Vuillemin *et al.* 2016).

The potential limits of ancient DNA related to the known modifications of the DNA signal when aging (nucleotidic mutations, DNA fragmentation, ...) are described especially for the case of very ancient DNA (> thousand years old) for applications in archaeology, anthropology, paleontology, evolutionary biology and conservation (Willerslev & Cooper 2005, Hebsgaard *et al.* 2005, Rizzi *et al.* 2012, Hagelberg *et al.* 2015). Here we considered more particularly the potential problem of degradation processes acting at the sediment–water interface and during the early burial. The early diagenesis occurring when organic material (among which DNA) is buried in sediments is indeed known to be a critical step for archiving as it was illustrated for some microfossils (e.g. foraminifers; Berkeley *et al.* 2007), pigments (Leavitt & Hodgson 2002) or for carbon and nitrogen amounts (e.g. Gälman *et al.* 2008) which are modified during the first years of archiving. Moreover, as previously reported by Behrensmeyer *et al.* (2000), one of taphonomy's most promising new frontiers is comparative analysis applied to different taxonomic groups within biological assemblages. Here we simultaneously explore the effect of early diagenesis on DNA of more than 40 micro-eukaryotic groups.

### III.5.1. Overall, stability of the DNA signal of micro-eukaryotes

We observed no significant effect of early diagenesis at the level of the entire micro-eukaryotic community, and the comparison of the two sets of data (C and C+6) based on

procrustean analyses showed no significant difference showing that a comparable temporal dynamics was obtained for each dataset (based on the analysis of 12 strata covering a 40 yrs time period). However, some variations between the molecular inventories obtained for C and C+6 cores were detected at a finer taxonomic level.

We first verified that DNA quantities did not change after 6 additional years of archiving. This result was not really expected since we could hypothesize that, especially for extracellular DNA, exposure to bacterial degradation processes should have resulted in a decrease of DNA quantities after 6 years of burying, as previously observed for C and N losses rates in the same lake (Gälman *et al.* 2008). Stability was also observed in the total micro-eukaryotic richness; no significant differences were indeed detected between the molecular inventories obtained for the two cores (C and C+6) in terms of richness, diversity and evenness of the whole microbial eukaryotic diversity.

As mentioned before, the structure of microbial eukaryotes community provided very similar patterns for the 2 cores as illustrated by the procrustean analysis. This result about unchanged global structure of the community under early diagenesis differs from the results reported for prokaryotic communities in the sediments of the Laguna Potrok (Argentina) (Vuillemin *et al.* 2015). In this oxic lake, the authors indeed showed, based on the comparison of sedimentary DNA composition and fossil pigments, that post-depositional diagenesis resulted in a rapid change in the initial nucleic acid composition. However, because prokaryotic sedimentary DNA partly originates from active taxa living in the sediments (with vertical living distribution that may change temporally, between seasons for instance), the comparison with eukaryotes is not totally relevant. Indeed very few eukaryotes are known to be able to be active in lacustrine sediments. Capo *et al.* (2015) reported that eukaryotic sedimentary DNA mostly represented DNA from past and inactive communities in Lake Bourget sediments; based on RNA analyses of sedimentary DNA. The authors showed that active eukaryotes were detected only in the top 2 cm of sediment, and were mainly represented by one benthic ciliate and to a lesser extent by fungi and chlorophyta. The rare available literature shows that the presence of active eukaryotic taxa in lake sediment was mostly attributed to fungal taxa. Wurzbacher *et al.* (2010) suggested for instance a possible heterotrophic fungal activity in the upper sediment layers, while Fungi are generally thought to be aerophilic, several species from all fungal groups can withstand or remain active under anoxic conditions (Scupham *et al.* 2006). If bacterial and archaeal communities from the sediment of various aquatic environmental systems have been widely examined, in contrast, sediment fungal community

structure and diversity are not well examined although fungal species may participate to transformation of organic matter decomposition at the water–sediment interface (Zhang *et al.* 2015).

### **III.5.2. Only one taxonomic group varied according to the core: focus on Fungi**

Interestingly, while for most of the taxonomic groups, the richness did not differ between the 2 cores suggesting that the early diagenesis did not affect significantly the number of taxa detected in 12 sedimentary varves analysed here, the only significant change detected in terms of richness and relative abundance was the increase of fungal OTUs in the core C+6 in comparison to the core C. We first took into consideration that a potential contamination during sampling, storing and molecular analysis could have induced this change; however this seems to be unlikely because of the methodological precautions used during these steps following the known recommendations (e.g. Rizzi *et al.* 2012) and above all because blanks used to verify the absence of contaminant DNA during sampling, DNA extraction and PCR amplifications, were found to be negative, showing the absence of amplifiable contaminant DNA.

We hypothesize that the presence of active Fungi, including Chytridiomycota, Cryptomycota and Glomeromycota taxa which may comprise saprotrophic taxa, partly explains the difference in terms of abundance and diversity observed between the two cores. The reason why one of the 2 cores was more favorable to the activity of these saprotrophic Fungi remains unclear, and might be related to both spatial and temporal variability in the conditions at water/sediment interface (duration of reoxygenation of the bottom of the lake for instance) when sampling was realized. RNA-based analyses could be used to verify the presence and identity of active Fungi in Nylandssjön sediments. Among the fungal taxa we pointed out in this study, most of them could be either saprotrophs or parasites and we cannot here definitely conclude about their ecological function due to a limited taxonomic resolution and above all due to current knowledge of fungal ecology in lakes which is still rudimentary (Wurzbacher *et al.* 2010). For instance, *Rhizophidium* sp. which largely contributed to the increase of Fungi relative abundance in the core C+6, is known to harbor both saprotrophs (on cellulose, chitin, keratin; Sparrow 2013) and parasite of various algae (Jones *et al.* 2014), with more than 220 described species (Zhang *et al.* 2015). The presence of putative fungal parasites have already been reported in lakes and marine systems (Coolen & Shtereva 2009, Wurzbacher *et al.* 2010, Edgcomb *et al.* 2011, Orsi *et al.* 2013, Chambouvet *et al.* 2014); fungal parasites may reach

sediment with their respective hosts (in an intracellular or extracellular state), their activity in the sediment is uncertain.

Another explanation to the increase of Fungi richness in C+6 core could be that diagenetic process induced modifications and genetic polymorphism in DNA sequences that therefore led to an increase in the number of false OTUs. In this case one can wonder why only Fungi are affected, since no other group was found to present the same trend. In any case, our results suggest that, for reconstruction of past eukaryotic diversity based on sedimentary DNA sequencing, the interpretation of Fungi data have to be considered carefully in the absence of systematic studies allowing to clarify their diversity, activity and role in the upper sediments.

### **III.5.3. Variations in relative abundance of some OTUs in the uppermost strata**

Considering the OTUs level, inter-core variability was detected for molecular inventories from the 6 most recent strata (ages: 7 to 15 years), revealing changes in the relative abundance of few specific OTUs. These OTUs belonged to Chlorophyta, Ciliophora and related unclassified Alveolata and unclassified Stramenopile. One hypothesis regarding the DNA preservation during burying into sediments is that extracellular could be less protected than intracellular DNA (physically protected within resistant cells), though adsorption to mineral and organic matrices is known to partly protect extracellular DNA from nuclease action. Here, no link could be detected between variations of the relative abundance of OTUs (between C and C+6) and their putative level of DNA protection by cells architecture (including the capacity of taxa to form resistant cells (spores, cysts ...)).

In addition, it is noticeable that, with the exception of one OTU (unclassified stramenopiles), changes of the relative abundances of these OTUs between C and C+6 followed no clear trend that could reveal degradation under the effect of early diagenesis. According to the strata, they could indeed either increase or decrease after 6 years of burying.

Based on the comparison of quantity decline with time of DNA originating from various groups (dinophyceae diatom and green sulfur bacteria), Boere et al. (2011a) previously showed that the DNA preservation is not only related to the cell architecture. However, this issue is still to be further explored; among the tests to perform to gain knowledge in this topic, we assume that the analysis of extracellular DNA could be of interest. When working with standardized commercial DNA extraction kits (as we and many other authors, including Boere et al (2011a), did), it can be questioned whether amplified DNA came mainly from DNA



originating from resting cells or from extracellular genetic material (both types of DNA are theoretically extracted the commercial kits we used). Whatever its origin, extracted DNA could be exploitable to study past communities from Sed-DNA sequencing, however resulting molecular inventories may vary. Methods allowing to extract specifically extracellular DNA or to remove extracellular DNA in order to keep only intracellular DNA are available (Corinaldesi *et al.* 2005; Taberlet *et al.* 2012, Lever *et al.* 2015). Comparative investigations could be performed to assess the extent to which molecular inventories vary according to DNA extraction methods used; the ability to separate soluble extracellular DNA pools without cell lysis from intracellular and particle-complexed DNA pools may enable new insights into the preservation of DNA in sediments (Lever *et al.* 2015). Overall we did not observe any differential degradation (i.e. preferential loss of certain taxa) for eukaryotic groups in Nylandssjön sediments. This might result from rather quick sedimentation processes in an environment undisturbed by sediment resuspension and transportation. Moreover, the lake bottom of Nylandssjön is known to be anoxic, a condition known to be favor DNA preservation (Coolen & Gibson 2009).

For the varves older than 15 years (dated before 1991, ~ 7 cm depth in sediment), the structure of microbial eukaryotes communities was found to be stable and very similar between the two cores. This suggests that, when possible, avoiding the uppermost sediment layers should allow to produce more reliable reconstructions of past microbial eukaryotic communities inferred from sedimentary DNA.

### **III.6. Conclusion**

Gälman et al. (2008) revealed the temporal changes in C and N contents of the Nylandssjön sediment from 1950 to 2006 after 5 years of burying while here the DNA signal appeared to be quite stable for all sediment strata older than 15 years corresponding in the lake at 7 cm of depth in the sediment. The modifications detected for the uppermost sediment layers are explained by few OTUs for which variations of relative abundances are observed but without significant impact on the total richness of the whole microbial eukaryote assemblage or on the way the different deposits (dated strata) are discriminated in a given core. Clearly, biogeochemical sedimentary profiles (bacterial and archaeal activity, redox gradients with sediment depth) would allow to consolidate the results obtained here. Moreover, because those results were obtained from one shallow and cold lake, such results cannot be directly extrapolated to very different systems (deep warm lakes, shallow lake with oxic bottom...) in which environmental conditions prevailing for DNA preservation during early diagenesis could largely differ.

Overall, the present study suggest that precautions should be taken more particularly for the Fungi group that could vary in terms of richness and relative abundance from one core to another.. Clearly, all paleogenetics studies using lacustrine sediments would benefit of systematic study of DNA preservation during burying into sediments; this research will contribute to a better understanding of the fossil record in lacustrine systems.

### III.7. Supporting information

Figure III.S1. Quantities of DNA extracted for each sample (in  $\mu\text{g.g}^{-1}$  of wet sediment).

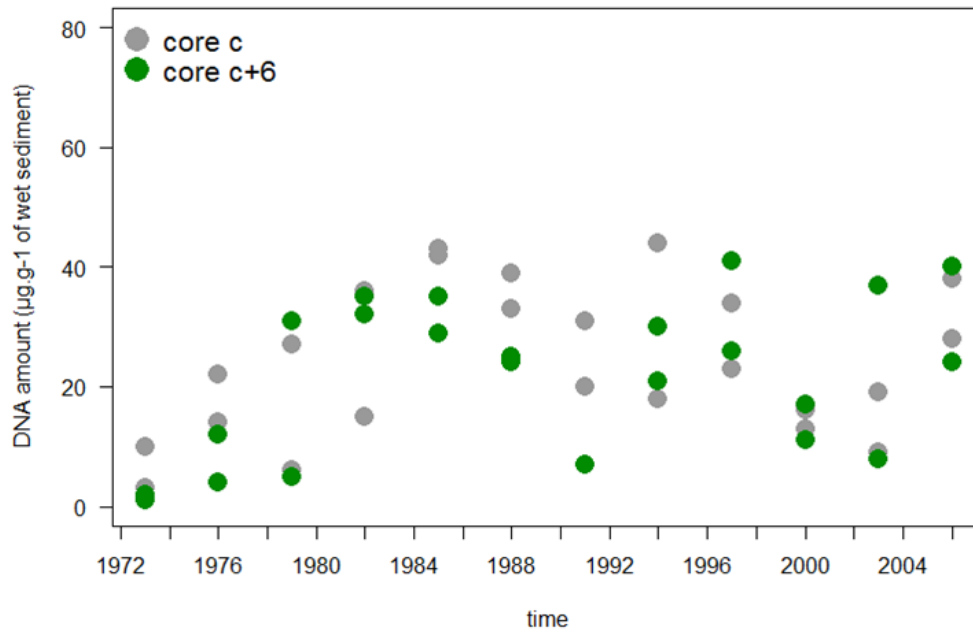
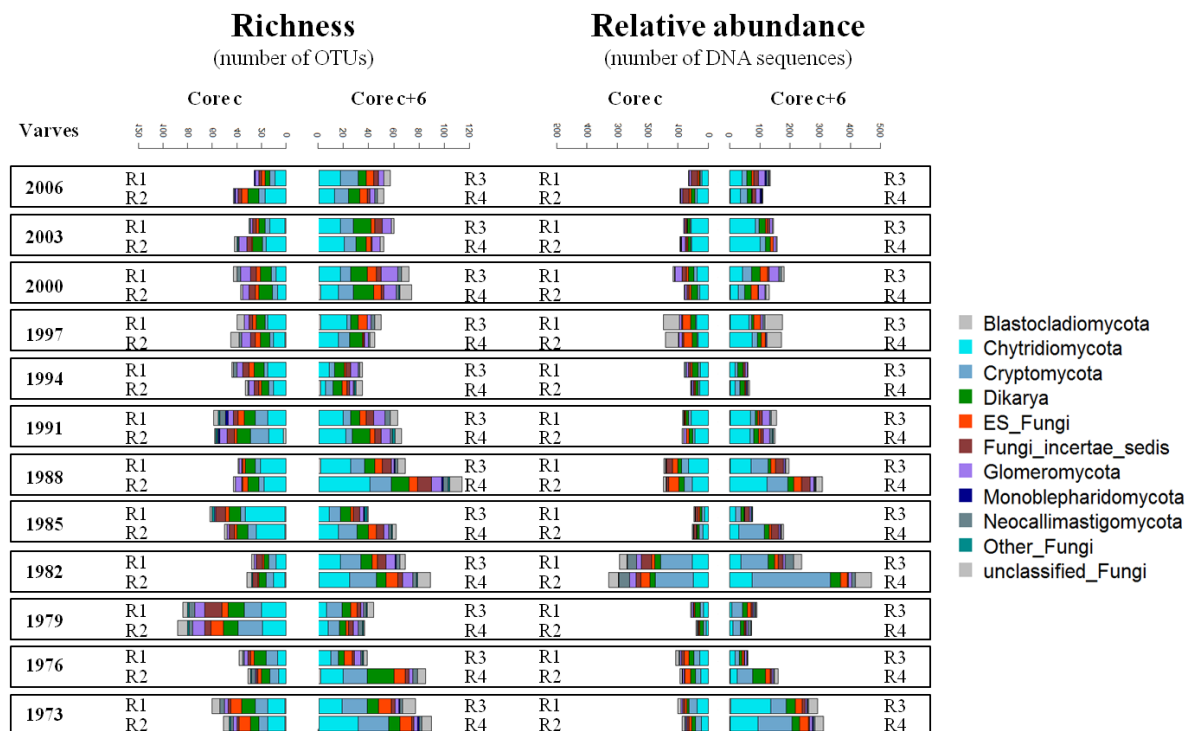


Figure III.S2. Distribution of the richness and relative abundance of fungal groups in the 48 molecular inventories.





Microbial eukaryotes groups	1982 varve								1979 varve								1976 varve								1973 varve											
	c				c+6				c				c+6				c				c+6				c				c+6							
	N725A		N725B		N1331A		N1331B		N728A		N728B		N1334A		N1334B		N731A		N731B		N1337A		N1337B		N734A		N734B		N1340A		N1340B					
	#reads	#OTUs	#reads	#OTUs	#reads	#OTUs	#reads	#OTUs	#reads	#OTUs	#reads	#OTUs	#reads	#OTUs	#reads	#OTUs	#reads	#OTUs	#reads	#OTUs	#reads	#OTUs	#reads	#OTUs	#reads	#OTUs	#reads	#OTUs	#reads	#OTUs						
Alveolata	Apicomplexa	12	3	21	4	13	4	22	3	11	4	7	4	20	3	20	3	2	1	6	3	6	2	21	4	3	2	1	1	4	3	7	3			
	Ciliophora	256	33	245	29	345	35	281	28	59	18	87	17	219	28	238	29	242	36	227	37	108	29	66	25	815	60	776	52	370	43	238	37			
	Dinophyceae	60	29	63	33	18	17	23	14	22	12	12	9	15	9	15	11	28	17	20	16	5	5	18	17	14	11	26	13	19	13	19	12			
	Perkinsea	17	8	2	2	4	4	15	6	3	3	1	1	12	8	11	7	13	8	11	8	5	5	15	7	17	12	14	5	21	9	12	7			
	Voromonas	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	2	1	5	1	0	0	0	0	0	0	1	1	1		
unclassified Alveolata	242	15	234	18	66	11	183	12	71	10	60	9	136	12	45	9	190	15	121	14	73	8	104	10	203	19	180	16	250	16	163	16				
Amoebozoa	Tubulinea	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
Cryptophyta	Cryptomonadales	0	0	1	1	0	0	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
	Cryptophyta_3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
	Cryptophyta_4	1	1	2	1	3	3	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	Pyrenomonadales	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	unclassified Cryptophyta	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Haptophyceae	Coccolithales	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	Isochrysidales	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	Pavlovales	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	Phaeocystales	1	1	0	0	1	1	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	Pyrmniales	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
unclassified Haptophyceae	0	0	2	1	2	1	3	2	2	1	0	0	3	2	0	0	0	0	0	0	1	1	4	2	0	0	0	0	1	1	0	0	0	0		
Opisthokonta	Choanoflagellida	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
	Fungi	293	84	329	88	239	69	469	89	58	38	40	31	91	44	73	37	107	60	94	51	61	39	161	85	100	57	86	52	291	77	310	90			
unclassified Opisthokonta	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
Rhizaria	Cercozoa	44	19	34	17	32	15	50	21	15	9	7	5	8	6	5	13	9	8	7	11	9	20	15	12	6	11	6	36	18	30	14	14			
	unclassified Rhizaria	14	9	12	5	8	6	19	10	10	4	6	4	8	7	3	3	6	5	8	4	23	8	9	4	2	2	3	3	16	6	10	7			
Stramenopiles	Bacillariophyta	89	17	110	15	62	13	60	12	69	13	67	13	15	7	18	3	143	34	118	17	53	15	69	18	97	18	92	18	67	10	99	15			
	Bicosoecia	2	2	3	0	0	0	4	4	0	0	1	1	1	1	1	1	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	Bolidophyceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	Chrysochyceae	107	49	124	45	180	57	120	44	255	69	306	72	348	77	373	85	160	66	164	62	106	37	123	49	89	42	93	43	113	46	124	48			
	Dityochophyceae	1	1	0	0	0	0	1	1	0	0	1	1	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	Eustigmatophyceae	4	2	3	2	3	3	1	1	11	3	9	2	9	2	14	3	5	2	7	4	1	1	1	1	1	1	1	1	1	1	2	2	4	3	
	Labyrinthulomycetes	141	23	143	22	205	23	109	12	460	33	455	30	745	40	875	37	166	12	152	18	92	11	100	14	95	14	79	16	68	9	50	8			
	MAST	3	2	2	1	0	0	0	0	2	2	0	0	0	0	0	0	0	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	Oikomonadaceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	Olmycetes	0	0	0	0	5	3	8	4	1	1	2	2	3	3	0	0	6	2	4	3	8	5	6	2	2	2	2	2	2	15	3	9	3		
	Pinguicophyceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	PX_clade	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	Raphidophyceae	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Synurophyceae	26	13	50	16	43	13	33	12	51	14	105	11	119	13	197	17	58	18	48	14	26	9	49	15	32	9	21	10	15	6	30	14			
unclassified Stramenopile	773	53	1073	60	442	56	994	29	3072	86	3483	61	2873	61	2777	79	314	113	3403	113	3544	88	3441	101	3294	121	3653	120	3761	51	3923	111	111			
Viridiplantae	Chlorophyta	3371	328	3010	288	2796	232	3085	179	1267	174	825	126	854	159	756	149	1223	175	1092	180	1561	145	1273	177	709	131	457	101	444	77	459	91			
	Streptophyta	2	2	2	2	1	2	1	0	0	0	0	1	1	1	1	3	3	2	2	1	1	3	3	1	1	0	0	0	0	0	0	0	0		
	unclassified Viridiplantae	38	9	30	8	28	12	9	5	16	3	20	3	17	7	27	2	12	6	12	5	9	7	3	3	8	6	0	0	1	1	2	2			

Table III.S2. Results obtained from SIMPER analysis comparing the molecular inventories of the two cores considering the 6 most recent strata. The table presents the percentage of contribution of the 39 OTUs to the dissimilarity between the two cores; only OTUs that contributed to more than 0.2 % of the total dissimilarity are presented. The mean number of DNA sequences obtained for each of these OTUs in C and C+6 cores are presented as well as their taxonomic affiliation.

Taxon	Contrib. %	Mean #reads in core c	Mean #reads in core c+6	Microbial eukaryote group	Taxonomic resolution
OTU189	16.99	1220	1670	Chlorophyta	Chlorophyta
OTU9301	10.66	470	657	unclassified_Alveolata	unclassified_Alveolata
OTU14361	6.64	381	14	Ciliophora	<i>Condylostoma</i> sp.
OTU14437	6.54	518	200	unclassified_Stramenopiles	unclassified_Stramenopiles
OTU1790	3.58	112	242	unclassified_Stramenopiles	unclassified_Stramenopiles
OTU8942	2.86	243	158	Chlorophyta	<i>Chlamydomonas</i> sp.
OTU12178	2.10	145	145	Chlorophyta	<i>Choricystis</i> sp.
OTU17253	2.07	150	234	Chlorophyta	<i>Scherffelia</i> sp.
OTU2575	1.40	54	74	Chlorophyta	<i>Desmodesmus</i> sp.
OTU14887	1.39	131	76	Chlorophyta	Chlamydomonadales
OTU11734	1.18	25	71	Chlorophyta	Chlorophyta
OTU7045	1.11	3	62	Synurophyceae	<i>Mallomonas</i> sp.
OTU9294	0.99	47	27	unclassified_Alveolata	unclassified_Alveolata
OTU17044	0.89	45	75	Chlorophyta	<i>Pseudoscourfieldia</i> sp.
OTU18164	0.77	42	59	unclassified_Alveolata	unclassified_Alveolata
OTU11165	0.74	35	52	Dinophyceae	Dinophyceae
OTU2190	0.59	58	33	unclassified_Alveolata	unclassified_Alveolata
OTU9717	0.54	42	23	Ciliophora	<i>Arcuospathidium</i> sp.
OTU3322	0.50	37	39	Chlorophyta	Chlorophyta
OTU13955	0.49	51	37	Chlorophyta	Chlamydomonadales
OTU11926	0.48	23	23	Dinophyceae	Dinophyceae
OTU10554	0.43	6	25	Synurophyceae	<i>Mallomonas</i> sp.
OTU9995	0.43	40	31	Chlorophyta	Chlorophyta
OTU8250	0.42	21	9	Chlorophyta	Trebouxiophyceae
OTU15442	0.41	17	11	Ciliophora	<i>Diophrys</i> sp.
OTU4893	0.40	22	7	Pavlovales	<i>Rebecca</i> sp.
OTU18543	0.37	15	25	Chrysophyceae	<i>Cyclonexis</i> sp.
OTU17048	0.36	23	9	unclassified_Stramenopiles	Pirsonia
OTU968	0.35	28	12	Bacillariophyta	Fragilariaceae
OTU2507	0.34	18	11	unclassified_Haptophyceae	unclassified_Haptophyceae
OTU13894	0.34	15	7	unclassified_Rhizaria	unclassified_Rhizaria
OTU16432	0.30	22	11	Labyrinthulida	Labyrinthulida
OTU18636	0.30	5	16	Chlorophyta	Chlorophyta
OTU15595	0.27	0	15	Synurophyceae	<i>Mallomonas</i> sp.
OTU9402	0.24	17	11	Fungi	Chytridiaceae
OTU16895	0.23	8	16	Chlorophyta	<i>Ulothrix</i> sp.
OTU17132	0.23	9	7	Fungi	Fungi
OTU14512	0.21	7	11	Chlorophyta	Scenedesmales
OTU12256	0.21	14	4	Labyrinthulida	Labyrinthulida



# **IV. Impact des forçages environnementaux au cours des 2200 dernières années sur les communautés microbiennes eucaryotes des lacs Bourget (France) et Igaliku (Groenland)**

## **Long-term dynamics in microbial eukaryotes communities: a paleolimnological view based on sedimentary DNA**

Accepté dans Molecular Ecology

Eric Capo<sup>1</sup>, Didier Debroas<sup>2,3</sup>, Fabien Arnaud<sup>4</sup>, Typhaine Guillemot<sup>5</sup>, Vincent Bichet<sup>5</sup>, Laurent Millet<sup>5</sup>, Emilie Gauthier<sup>5</sup>, Charly Massa<sup>5</sup>, Anne-Lise Develle<sup>4</sup>, Cécile Pignol<sup>4</sup>, Franck Lejzerowicz<sup>6</sup>, Isabelle Domaizon<sup>1\*</sup>

<sup>1</sup>CARTELE, INRA, Université de Savoie Mont Blanc, 74200, Thonon-les-bains, France

<sup>2</sup>Clermont Université, Université Blaise Pascal, CNRS, UMR 6023 LMGE, 63171 Aubière, France

<sup>3</sup>CNRS, UMR 6023, LMGE, Campus Universitaire des Cézeaux 63171 Aubière, France

<sup>4</sup>CNRS, UMR 5204 EDYTEM, Université Savoie Mont Blanc, Le Bourget du Lac Cedex, France

<sup>5</sup>Laboratoire Chrono-Environnement, UMR 6249 CNRS, Université de Bourgogne-Franche-Comté, Besançon,

<sup>6</sup>Department of Genetics and Evolution, University of Geneva, Geneva, Switzerland

### **Contribution des auteurs :**

- Conception de l'étude: ID, EC, DD, VB, LM
- Acquisition des données: EC, TG, ALD, CP, FA, VB, LM, CM, ID
- Traitements bioinformatiques: DD, EC
- Analyses et interprétations des données: EC, ID, DD, FA, TG, VB, LM, ALD, CP, FL,
- Ecriture du papier: EC, ID, DD, FA, TG, VB, LM, EG, ALD, CP, FL

### **Résumé**

La dynamique temporelle des communautés de micro-eucaryotes lacustres a été étudiée au cours des 2200 dernières années en appliquant le séquençage massif (MiSeq) à l'ADN sédimentaire de carottes sédimentaires provenant de deux lacs (lac du Bourget, lac Igaliku). Une large diversité de taxons a été révélée au sein des communautés des deux lacs avec plus de 16 000 OTUs distribuées dans 50 groupes de micro-eucaryotes. De nombreux groupes jusque là rarement étudiés dans les archives sédimentaires lacustres notamment les



Chlorophytes, les Dinophycées, et les Ciliés semblent apparaître comme des biomarqueurs intéressants des changements d'état trophique. Les données obtenues suggèrent que l'eutrophisation du lac du Bourget dans les années 1940s est le forçage le plus fort pour la communauté de micro-eucaryote et ses effets ont surpassé ceux liés aux fluctuations climatiques. Dans le cas du lac Igaliku, les changements dans la structure de la communauté du lac Igaliku (moins exposé aux forçages anthropiques locaux) ont été détectés de manière concomitante avec les fluctuations climatiques majeures de ces 2 derniers millénaires (réchauffement médiéval, petit âge glaciaire, réchauffement récent). Cette analyse met par ailleurs en évidence le rôle joué par la biosphère rare dans les réarrangements temporels de la communauté micro-eucaryote. Ces données ont mis en évidence le potentiel de l'ADN sédimentaire pour décrire la diversité des micro-eucaryotes lacustres sur plusieurs millénaires.

## IV.1. Abstract

Assessing the extent to which changes in lacustrine biodiversity are affected by anthropogenic or climatic forces requires extensive paleolimnological data. We used high-throughput sequencing to generate time-series data encompassing over 2200 years of microbial eukaryotes (protists and Fungi) diversity changes from the sedimentary DNA record of 2 lakes (Lake Bourget in French Alps and Lake Igaliku in Greenland). From 176 samples, we sequenced a large diversity of microbial eukaryotes, with a total 16 386 Operational Taxonomic Units (OTUs) distributed within 50 phylogenetic groups. Thus, microbial groups, such as Chlorophyta, Dinophyceae, Haptophyceae and Ciliophora that were not previously considered in lacustrine sediment record analyses appeared to be potential biological markers of trophic status changes. Our data suggest that shifts in relative abundance of extant species, including shifts between rare and abundant taxa, drive ecosystem responses to local and global environmental changes. Community structure shift events were concomitant with major climate variations (more particularly in Lake Igaliku). However, this study shows that the impacts of climatic fluctuations may be overpassed by the high-magnitude eutrophication impacts, as observed in the eutrophicated Lake Bourget. Overall, our data show that DNA preserved in sediment constitutes a precious archive of information on past biodiversity changes.

Keywords: protists, lake, climate change, eutrophication, eDNA.

## IV.2. Introduction

Microbial eukaryotes (i.e. protists and Fungi) are key components of aquatic systems, as they largely contribute to biogeochemical processes and ecosystems functioning (e.g. Falkowski *et al.* 2008). They represent a diverse and abundant fraction of planktonic assemblages in freshwater and marine environments where they play various ecological roles (e.g., primary producers, phagotrophs, saprotrophs, and parasites) (Lefranc *et al.* 2005, Oikonomou *et al.* 2015, Massana *et al.* 2015, de Vargas *et al.* 2015). Hence, their responses to environmental fluctuations may directly affect aquatic ecosystems functioning (Caron & Hutchins 2013). However, there are knowledge gaps about how microbial eukaryotes communities change over time and at which time scales, and how environmental conditions shape these communities. Recent research has endeavored to understand the relative impacts of warming and eutrophication on phytoplankton composition and productivity, and has shown the

existence of interactions between these two forcing factors (e.g. Gallina *et al.* 2013, Rigosi *et al.* 2014). Up to now, such long-term environmental monitoring studies rarely investigated the whole diversity of microbial eukaryotes. As introduced by Caron & Hutchins (2013), the effects of climate change on these microbial eukaryotes communities remains poorly known, which is most probably due to the lack of long-term time-series that are a prerequisite for understanding naturally- and anthropogenically- induced modifications in biological communities. The paleolimnological approach is ideal to study the long-term dynamics of lacustrine populations, and allows studying the effects of disturbances due to climate change and eutrophication (e.g. Smol 1992, Battarbee *et al.* 2001). However, the proxies traditionally studied, mainly based on the morphological remains of diatoms, zooplankton, chironomids, reflect only a small part of the total planktonic diversity.

The recent emergence of molecular analyses based on the sequencing of environmental DNA has opened up new areas of inquiry, extended to the spatial and temporal survey of the whole spectrum of the diversity of planktonic eukaryotes (de Vargas *et al.* 2015, Massana *et al.* 2015, Debroas *et al.* 2015). Based on this approach, sedimentary DNA (Sed-DNA) can be used to reach a more complete reconstruction of planktonic biodiversity, including taxa with no visible diagnostic features. Up to now, paleolimnological studies involving the analysis of preserved DNA focused on specific microbial groups, such as cyanobacteria (Domaizon *et al.* 2013), diatoms (Stoof-Leichsenring *et al.* 2012, 2015) or haptophytes (Coolen *et al.* 2004, Randlett *et al.* 2014). Recently it was demonstrated that lacustrine Sed-DNA constitutes an appropriate material for paleo-reconstruction of the huge diversity of planktonic microbial eukaryotes since DNA derives from all taxa and reflects mostly past communities (Capo *et al.* 2015).

The potential limits of paleo-reconstruction based on sedimentary DNA are mainly linked to the fate of DNA in the environment, with potential risks of biotic degradation, abiotic decay leading to DNA fragmentation, and potential differential preservation according to environmental conditions and taxa (Dell'Anno & Corinaldesi 2004, Corinaldesi *et al.* 2005, 2008, Boere *et al.* 2011, Willerslev & Cooper 2005, Torti *et al.* 2015). However, aquatic sediments have proven to be rich in DNA with preservation being favored by cold and anoxic conditions (Boere *et al.* 2011a, Thomsen & Willerslev 2015), and the analysis of sedimentary DNA to reconstruct past aquatic biological assemblages has already been applied successfully (Domaizon *et al.* 2013, Lejzerowicz *et al.* 2013, Hou *et al.* 2014, Stoof-Leichsenring *et al.* 2015). One study recently investigated the whole diversity of aquatic eukaryotes to document

their response in terms of community composition to freshwater/marine transitions over the last 10 000 years of Black Sea history (Coolen *et al.* 2013). Clearly, fundamental questions in microbial ecology would benefit from the integration of DNA-based paleo-reconstructions. More specifically, targeting eukaryotic DNA would allow addressing the challenging question of the effect of environmental changes on the microbial eukaryotes assemblages. This approach would also allow exploring the role of the rare microbial biosphere in biological responses to environmental disturbances, revealing rare taxa that contribute to deterministic community dynamics (Shade & Gilbert 2015, Logares *et al.* 2015). This information would provide original contribution to the active debates on the role of these low abundance taxa, which are considered either to perform specific ecological functions within the ecosystem or to contribute to the functional redundancy and act as a reservoir of diversity to buffer the impacts of environmental stress (e.g. Lynch & Neufeld 2015, Coveley *et al.* 2015).

In the present study, we investigate the temporal dynamics of microbial eukaryotes communities of two lakes based on the high-throughput sequencing (HTS) of Sed-DNA. These two lakes present contrasted settings; one located in continental Europe (Lake Bourget) is submitted to intensive land-use and catchment area urbanization (Giguet-Covex *et al.* 2010, Millet *et al.* 2010), whereas the other one (Lake Igaliku) is located in Greenland and only faced episodic human forcing (local agriculture activities; Guillemot *et al.* 2015). The considered temporal framework allows to cover the main climatic periods of the last 2 millennia: Medieval Warm Period (MWP), Little Ice Age (LIA) and Current Warm Period (CWP) (McKay and Kaufman, 2014). We hypothesize that these climatic fluctuations induced significant modifications within microbial eukaryotes assemblages in both lakes, whereas marked eutrophication periods led to drastic rearrangements of community over the last century, more particularly in Lake Bourget. Our main objectives are (i) to identify the rearrangements that occurred through time within the communities and describe the temporal changes in terms of community diversity, composition and structure, (ii) to ascribe to these rearrangements the relative contribution of photosynthetic and heterotrophic taxa, as well as of the rare biosphere, and (iii) to document the diversity changes in microbial eukaryotes communities with respect to two majors forces of environmental fluctuations enacting either locally (i.e. anthropogenic pressure leading to eutrophication) or globally (i.e. climate forcing reflected by temperature fluctuations).

## IV.3. Materials and methods

### IV.3.1. Study sites

Lake Bourget (45°44'N 5°51'E 231 m a.s.l., maximum depth: 147 m) is a deep monomictic lake located in European Alps and currently classified as oligo-mesotrophic (Jacquet *et al.* 2014). A marked eutrophication, described from both water column monitoring (long-term observatory on alpine lakes [http://www6.inra.fr/soere-ola\\_eng](http://www6.inra.fr/soere-ola_eng)) and paleolimnological studies (Giguet-Covex *et al.* 2010, Alric *et al.* 2013, Jenny *et al.* 2013, Berthon *et al.* 2014, Perga *et al.* 2015) was reported for this lake. The early eutrophication was observed from the mid 1940s followed by a marked eutrophication in the 1960s-1980s with the highest phosphorus concentrations ( $> 100 \mu\text{g.L}^{-1}$ ). This lake has undergone a restoration leading to a progressive re-oligotrophication from the 2010s (Jacquet *et al.* 2014).

Lake Igaliku (61°00'N 45°26'W 15 m a.s.l., maximum depth: 26 m) is a small and dimictic lake located in southern Greenland and currently oligotrophic. This lake has been preserved from the marked eutrophication. In Greenland, only two periods of agropastoral activities were inventoried (i) the Norse settlement period between AD 986 and the middle of the 15<sup>th</sup> century (Arneborg *et al.* 2011, Massa *et al.* 2012, Guillemot *et al.* 2015) (ii) the recent occupation, since the 1920s, marked by moderate eutrophication around the 1980s (Austrheim *et al.* 2008, Perren *et al.* 2012, Millet *et al.* 2014).

These two lakes underwent changes in climatic conditions over the last two millennia. A temporal framework divided into four climatic periods is used here, in accordance with common practice in paleoclimatic studies (Sicre *et al.* 2008, Büntgen *et al.* 2011, McKay & Kaufman 2014): (i) the pre-medieval period ( $< 700$  AD), the Medieval Warming Period (MWP; ca. AD 700- ca. AD 1300), the Little Ice Age (LIA; ca. AD 1300- ca. AD 1900) and the Current Warming Period (CWP; during the last century).

### IV.3.2. Field sampling & laboratory sub-sampling

The sediment cores were sampled in 2013 in the deepest area of the lakes Bourget (core LDB13-I; IGSN:IEFRA00C0, [www.geosamples.org](http://www.geosamples.org)) and Igaliku (core Iga-2013) using UWITEC piston and gravity corers (90 mm in diameter), respectively. Each core was halved longitudinally: one half was kept for dating and sedimentological parameters and the second half was dedicated to biological analyses. Forty-five and 43 sediment slices of 0.5 to 1 cm in thickness were sub-sampled from Lake Bourget and from Lake Igaliku sediment cores,

respectively. A specific effort of sampling was realized for the two warmer periods (medieval and recent periods) in both lakes. For each sediment layer, only the center of each slice was kept for DNA analysis to avoid any extraneous or cross-contamination due to contact with the plastic liner at the periphery. Then, the samples were immediately placed at - 20 °C and kept frozen until further molecular analyses (within the following month). All laboratory procedures were carried out in a dedicated laboratory, following strict precautions to ensure the authenticity of the results; in addition, blanks were introduced during the sub-sampling step (opened tube containing ultrapure water) in order to detect any aerial contamination.

### **IV.3.3. Dating of sediment cores**

For the sediment core from Lake Bourget, chemical composition was measured by XRF core scanning, following a sampling step of 5 mm both in studied and reference cores (Figure **IV.S1**, Supporting information). Sediment core from Lake Bourget is a composite sequence (i.e. made of 3 sections; LDB13-I-01A, LDB13-I-02A, LDB13-I-02B). It was dated by chemical wiggle matching by comparison with a set of 2 previously well-dated cores (Arnaud *et al.* 2012). It is based upon the hypothesis that chemical changes in sediment composition present wiggles that are synchronous in the whole deep lake basin. An age-depth model developed on 2 cores sampled in 2001 (LDB01-01, IGSN: EDYLDB001; LDB01-02, IGSN: EDYLDB002) and 2004 (LDB04-I, IGSN: IEFRA003S) by Arnaud *et al.* (2012) was used as a reference. The age-depth model was based on 7 clearly-identified chronological tie points and using (i) varve counting from 2013 back to 1930 (ii)  $^{210}\text{Pb}$  measurements back to 1900 AD and (iii)  $^{14}\text{C}$  ages back to 250 BC. This age-depth model was applied to LDB13-I by correlating the ratio K/Ti (measured by XRF core scanning) with LDB04 (Table **IV.S1** & Figure **IV.S1**, Supporting information). The age-uncertainty is less than 3 years and 5 years for varve counting and  $^{210}\text{Pb}$  dating respectively (Giguet-Covex *et al.* 2010, Jenny *et al.* 2014), and about +/- 100 years for the  $^{14}\text{C}$ -based age model (Arnaud *et al.* 2012; Figure **IV.S1**).

Sediment core from Lake Igaliku is a unique sequence (Iga-2013). An age-depth model was developed on a core retrieved in 2007 (Iga-2007) by Massa *et al.* (2012). It was based on 16 accelerator mass spectrometry (AMS) radiocarbon dates and on  $^{210}\text{Pb}$  measurements coupled with  $^{137}\text{Cs}$  to precisely date the last century. According to Massa *et al.* (2012), the Iga-2007  $^{210}\text{Pb}$  model has a mean error between  $\pm 25$  to 1.5 years during the 20th century whereas the Iga-2007  $^{14}\text{C}$  revealed a mean age error of  $\pm 25$  years during the Middle-Age. Iga-2007 and Iga-2013 were correlated thanks to magnetic susceptibility data, measured with a Geotek

Multi Sensor Core Logger (Chrono-Environnement laboratory, University of Bourgogne/Franche-Comté; Table **IV.S1**). Around 50 tie-points were identified between the two cores in order to perform stratigraphic alignment. For reasons of clarity, only ten are shown in Figure **IV.S1**.

#### **IV.3.4. Molecular analysis**

For each sediment layer, two DNA extractions were performed on 0.5 g of wet sediment by using the NucleoSpin® Soil kit according to the manufacturer instructions (Macherey-Nagel, Düren, Germany). The concentration of bulk DNA ranged from 5 to 125 ng.µL<sup>-1</sup> (mean 24 ng.µL<sup>-1</sup>). The V7 region of 18S rRNA gene was PCR amplified on a 260 bp long fragment, from ca. 25 ng of environmental DNA extract for each sample and using the general eukaryotic primers 960F (5'-GGCTTAATTTGACTCAACRCG-3') (Gast *et al.* 2004) and NSR1438 (5'-GGGCATCACAGACCTGTTAT-3') (Van de Peer *et al.* 2000). The combinations of tagged primers were designed to balance primers usage frequencies, and therefore to reduce the frequency of mistagging events that result in cross-contaminations (Esling *et al.* 2015). We multiplexed 176 samples using 176 out of the 400 possible combinations with 20 forward and 20 reverse tagged primers. Each PCR was performed in duplicate in a total volume of 25 µl containing 3 µl of 10X NH<sub>4</sub> reaction buffer, 1.2 µl of 50 mM MgCl<sub>2</sub>, 0.25 µl of BioTaq (Bioline), 0.24 µl of 10 mM dNTP, 0.36 µl of 0.5 µg/µl BSA and 1 µl of each primer. The amplification conditions consisted of an initial denaturation at 94 °C for 10 min followed by 35 cycles of 1 min at 94 °C, 1 min at 55 °C and 1 min 30 s at 72 °C. The amplicons were then subjected to a final 10-min extension at 72 °C. The two PCR products obtained for a given DNA extract were pooled and purified using Illustra™ GFX™ PCR DNA and Gel Band Purification Kit (GE Healthcare Life Sciences, Velizy Villacoublay, France) and quantified using Quant-it PicoGreen kit (Invitrogen, Carlsbad, CA, USA) on a Fluoroskan Ascent™ FL (ThermoLabsystems, Beverly, MA, USA). Finally, the purified amplicons corresponding to each sample were pooled at equimolar concentrations and sent to Fasteris SA (Geneva, Switzerland) for library preparation and paired-end (2 x 250 bp) sequencing on a MiSeq Illumina instrument (San Diego, CA, USA).

Specific precautions and choices for paleogenetics work:

The isolation of DNA from sediment was carried out in a specific DNA laboratory, physically separated from the molecular biology laboratory and employing rigorous controls against contamination. This included (i) separate stations for sampling, DNA extractions and PCR, (ii) the use of sterile disposable labware, (iii) the introduction of blanks (negative controls) for extraction and amplification steps, (iv) duplicates for extractions and for sequencing of each sediment sample.

The choice of the barcode region was based on the following verifications: (i) the region amplified was chosen after a thorough study of the coverage by the “universal” probes found in publications. This allowed verifying that V4 and V7 regions are good candidates in terms of coverage while V1 and V9 regions give the lower results. (ii) the length of the barcode was considered: from our samples, we verified the ability to amplify the V4 and V7 regions for all DNA extracts. For the longer fragment (V4 ; ~350 bp) we detected a loss of amplification in the sediment cores for sediment older than 900 yrs most probably due to DNA fragmentation; consequently we selected the V7 region (around 250 bp) that showed positive amplifications for all our samples and allowed to obtain an adequate quality of taxonomic identification with the bioinformatics tools used. (iii) the simulation performed by Taïb *et al.* (2013) showed that there were no statistical significant differences for the taxonomic identification of amplicons (200 bp) by a similarity method (i.e. BLAST) between the V4 and V7 regions. Based on these three verifications, the probes 960F and NSR1438 appeared to be good candidates in terms of coverage, length, and taxonomic assignment.

#### **IV.3.5. Bioinformatics processing**

The paired-end reads were merged together using UPARSE tools (option `-fastq_mergepairs` with a minimal overlap equal to 150 and no mismatch allowed) (Edgar 2013) allowing to obtain 24 088 960 raw DNA sequences. These sequences were then submitted to three cleaning procedures: (i) no undefined bases (Ns), (ii) a minimum sequence length of 200 bp, and (iii) no sequencing error in the forward and reverse primers. Putative chimeras were detected by UCHIME (Edgar *et al.* 2011). The combining of these cleaning procedures with a sequencing depth (2 x 250 bp) covering all the amplicon length (~ 220 bp without the primers) has allowed to sequence twice each base and to minimize drastically the sequencing errors. After this cleaning step and demultiplexing process, the remaining sequences were clustered at a 95 % similarity threshold with UPARSE 7.0 (Edgar 2013) to obtain the seed OTUs (option `-cluster_fast`). The delineating of the similarity threshold was chosen following



the recommendations from Mangot *et al.* (2013) who estimated the errors during PCR and sequencing by introducing an internal 18S standard (a unique alien DNA sequence) and demonstrated that 95 % was a relevant cut-off threshold for OTUs delineation allowing to limit the artificial inflation of OTUs number. This threshold was chosen after sequencing of a mock community in the same experimental conditions (primers, sequencing platform and bioinformatics pipelines), we obtained 16 OTUs from the 16 species in the mock community when applying the 95% threshold (data deposited in the Dryad Digital Repository).

The taxonomic affiliation was performed by BLAST against the SSURef SILVA database (Pruesse *et al.* 2007) after application of the following selection criteria: length > 1200 bp, quality score > 75 % and a pintail value > 50; SILVA database was enriched by lacustrine DNA sequences originating from various studies (Lefranc *et al.* 2005, Lepère *et al.* 2006, 2008, Mangot *et al.* 2013). The taxonomy of an OTU corresponded to the best hit given by the similarity search. If an OTU was associated with several best hits (hits having the same identity), then the taxonomy was the common taxonomy of these hits named also the lowest common ancestor. In addition, a specific taxonomy based on previous publications (Lefranc *et al.* 2005, Lepère *et al.* 2006, 2008, Mangot *et al.* 2013) allows to identify typical freshwater clades (e.g. Cryptophyta\_1, 2, Perkinsea\_1). OTUs affiliated to metazoans and embryophytes were removed from these analyzes in order to consider only microbial eukaryotes (protist and Fungi at the basis of trophic networks). After these different steps (cleaning, clustering and taxonomic affiliation), 13 044 506 quality-filter DNA sequences were clustered in 80 081 OTUs classified as microbial eukaryotes. For each sample, the OTUs which were detected only in one of the two duplicates were removed from the analysis; leading to 18 583 remaining OTUS. This additional cleaning step allows to minimize the presence of falsely rare OTUs due to PCR bias as recently pointed out by Lange *et al.* (2015) and Egge *et al.* (2015). For each remaining OTU, the numbers of DNA sequences obtained for each duplicate were then summed. For all samples, the alpha diversity was estimated using Hill's numbers (i.e. estimation of the effective number of species; Hill 1973, Jost 2007, Chao *et al.* 2014). Different Hill numbers  $qD$  are defined by their "order"  $q$  (Hill 1973), the parameter  $q$  controls the sensitivity of the measure to species relative abundance.  $0D$  equals species richness ( $q=0$ ),  $1D$  is the exponential form of Shannon entropy ( $q=1$ ) and  $2D$  is a transformation of Gini–Simpson measures ( $q=2$ ) (Gotelli & Chao 2013). These metrics were computed using the functions from the R package iNEXT (Chao *et al.* 2014, Hsieh *et al.* 2016). In order to perform comparisons between samples, the number of DNA sequences was standardized at 35

973 for all samples resulting to a total number of clean DNA sequences of 3 165 624 (13.1 % and 24.2 % of the raw and quality-filtered sequencing dataset, respectively). From these data, the final dataset (i.e. rarefied OTU table with the 16 386 OTUs expressed in terms of relative number of DNA sequences for each sediment layer) was used for all the further analyses on beta diversity. The data (demultiplexed sequences) are deposited in the Dryad Digital Repository (<http://datadryad.org/>).

#### **IV.3.6. Data treatment and statistical analyses**

Rarity: OTUs accounting for less than 0.1 % of the total DNA sequences were considered as rare OTUs. For each lake time series, 2 categories of rarity were considered, namely for OTUs that are (i) permanently rare (ii) oscillating (i.e. exhibit large temporal oscillations between abundant and rare). The existence, number and identity of taxa that exhibited temporal shifts from rare to abundant states (and inversely from abundant to rare) were recorded.

Ecological inference: We performed a basic functional annotation in order to discriminate OTUs corresponding to non-pigmented and pigmented taxa. Ecological inferences were based on the literature available about the presence/absence of chloroplasts in microbial eukaryotes cell. Some microbial groups include only phototrophs (mainly represented by Bacillariophyta, Chlorophyta, Haptophyta) and their OTUs were logically assigned to pigmented taxa. OTUs distributed within groups of free-living heterotrophs (mainly represented by phagotrophic Ciliophora, Choanoflagellida, Cercozoa, Amoebozoa, osmotrophic Fungi) and parasites (Apicomplexa, Perkinsea, Oomycetes) were reliably assigned to non-pigmented groups. When OTUs belonging to Chrysophyceae, Dinophyceae and Cryptophyta, which gather both pigmented and non-pigmented taxa, were not precisely identified (i.e. at least at the genus level) they were then assigned to “Unknown” ecological role.

Change points in diversity (Hill’s numbers): Bayesian change points (bcp) analysis was used to reveal timing changes in the diversity metrics ( $0D$  = OTU richness,  $1D$  = transformed Shannon and  $2D$  = transformed Gini-Simpson indices). The bcp analysis provides estimates of posterior probabilities for the presence of a change point between consecutive dates and of

posterior mean values of each index. The R package “bcp” (Erdman & Emerson 2007) was used for this analysis.

Community structure modifications: A chronological clustering was used to detect the timing and hierarchy of marked breakpoints in the community structure. For this purpose, a Multivariate Regression Tree (MRT) analysis was performed considering the overall OTUs table or using separately the non-pigmented and the pigmented OTUs tables. This time-constrained analysis was performed according to the instructions of Borcard *et al.* (2011), the size of each tree was selected using the graphs of the relative error (Figure **IV.S2**). The R package “mvpart” (De’ath 2002) was used for this analysis. In addition, non-metric multidimensional scaling (NMDS) analysis was used to summarize the main dimensions of modifications observed in the community structure. For each lake, a Bray-Curtis dissimilarity matrix and a Jaccard dissimilarity matrix were then built from the OTUs table and the NMDS graphic visualizations were then plotted using the R package “vegan” (Oksanen *et al.* 2015). To identify the OTUs that contributed to the dissimilarity over time, SIMPER analyzes were conducted using PAST v.2.11 (Hammer *et al.* 2001) considering groups of samples significantly discriminated by ANOSIM analyzes (p-values < 0.0001).

Table IV.1. The table presents (i) for each lake, the numbers of DNA sequences, OTUs and oscillating OTUs within each taxonomic group (ii) the numbers of common OTUs in the two lakes and core OTUs found in all samples.

Taxonomic groups	Lake Bourget			Lake Igaliku			Common OTUs	Core OTUs	
	#DNA sequences	#OTUs	#Oscillating OTUs	#DNA sequences	#OTUs	#Oscillating OTUs			
Alveolata	Apicomplexa	15464	144	4	7582	110	1	77	
	Ciliophora	241068	1185	27	194702	1157	32	723	9
	Colpodellidae	25	3		1	1		1	
	Dinophyceae	89488	894	13	42236	674	6	473	3
	Perkinsea	26424	146	2	32500	130	3	83	4
	Voromonas	159	13		5	3		2	
	unclassified Alveolata	312973	964	26	64600	750	13	571	15
Amoebozoa	Discosea	80	7		88	4		1	
	Tubulinea	58	13		1532	29	1	11	
Cryptophyta	Cryptomonadales	91	5		187	4		3	
	Cryptophyta 3	14	2		3	1		1	
	Cryptophyta 4	346	12		21	5		4	
	Pyrenomonadales	240	6		14	4		3	
	unclassified Cryptophyta	153	9		36	6		4	
Euglenozoa	Kinetoplastida	109	11		8307	40	5	10	
Haptophyta	Coccolithales				1	1			
	Isochrysidales				22	3			
	Phaeocystales	18619	26	2	716	6		5	
	Prymnesiales	302	8		131	7		5	
	Zygodisciales	1	1		4	2			
	unclassified Haptophyceae	74601	43	2	15329	37	4	22	2
Heterolobosea	Schizopyrenida	1849	82		48606	328	10	74	
Nucleariidae	Nucleariidae	206	3		3	1		1	
Opisthokonta	Choanoflagellida	211	9		24	8		4	
	Fungi	403422	3611	63	156225	2570	45	1644	10
	Opisthokonta incertae sedis	29	2		1	1			
Rhizaria	Cercozoa	117174	1156	18	416460	1589	29	738	9
	unclassified Rhizaria	33137	644	4	228359	1128	9	519	5
Rhodophyta	Bangiophyceae	418	11		2878	12	1	3	
	Florideophyceae	30	2		7	1		1	
Stramenopiles	Bacillariophyta	14697	199	9	9575	197	6	105	
	Bicosoecida	109460	262	4	81186	369	9	179	6
	Bolidophyceae	22	3		2	2		2	
	Chrysophyceae	11088	212	5	41370	290	9	138	2
	Dictyochophyceae	280	7		471	10	1	2	
	Eustigmatophyceae	4417	49	3	882	33		28	
	Labyrinthulida	333	19		2890	29	3	16	
	MAST	1150	34	1	653	23	1	13	
	Oikomonadaceae	29	2		18	3		2	
	Oomycetes	14137	158	2	2694	111		69	
	Phaeothamniophyceae	1	1		2	1		1	
	Pinguiphyceae	24	2		2	1		1	
	PX clade	649	20		367	21		11	
	Raphidophyceae	324	6		30	3		3	
	Synurophyceae	2458	60	1	15765	100	4	44	1
	unclassified Stramenopiles	24444	325	6	84358	477	13	243	4
Viridiplantae	Chlorophyta	94975	940	22	83064	929	29	509	3
	Streptophyta	2158	127		2004	117		59	
	unclassified Viridiplantae	1448	21		926	16	1	9	
Total	1618785	11459	214	1546839	11344	235	6417	73	

## **IV.4. Results**

### **IV.4.1. OTU richness and taxonomic composition**

A total of 16 386 OTUs was assigned to 50 phylogenetic groups. The number of OTUs varies from 1172 to 2011 OTUs in Lake Bourget samples and from 1208 to 2643 OTUs in Lake Igaliku. The total number of OTUs is quite similar in the two lakes (11 459 and 11 344 for Lake Bourget and Lake Igaliku, respectively), which have in common 40 % of the OTUs (Table **IV.S2**). The OTU richness distribution within each phylogenetic group is very similar for the two lakes (Table **IV.1**, Figure **IV.S3**). The non-pigmented groups (mainly Fungi, Ciliophora and Cercozoa) dominate the OTU richness in both lakes, with 60 % and 58 % of the OTUs found in lakes Bourget and Igaliku, respectively. Conversely, pigmented taxa account for only 18 % and 17 % of the OTUs found in lakes Bourget and Igaliku, respectively. Among pigmented taxa, Chlorophyta, Dinophyceae and Bacillariophyta (i.e. diatoms) exhibit the highest OTUs richness (Table **IV.1**). Less than 1 % of the OTUs are persistent throughout the entire time-series dataset (94 and 107 OTUs in lakes Bourget and Igaliku, respectively). Interestingly, 73 of them occur not only in both lakes but also in every sequenced sample. Although these OTUs gather 51 % of the DNA sequences and hence include the most abundant OTUs, their relative abundance profiles differ markedly between the two lakes. In Lake Bourget, the 10 most abundant OTUs representing around 30 % of the total number of DNA sequences are affiliated to Ciliophora, Fungi, Haptophyceae and Bicosoecida, while in Lake Igaliku they are affiliated to Cercozoa, Ciliophora, unclassified Rhizaria and unclassified Stramenopiles.

### **IV.4.2. Temporal tipping points detected for the structure of eukaryotes community**

The major change point in the composition of microbial eukaryotes community of Lake Bourget is placed during the 1940s by the MRT analysis (Figure **IV.1**), which corresponds to when eutrophication started (Figure **IV.S4**). This change is concomitant with a significant increase of diversity metrics ( $0D$ ,  $1D$  and  $2D$ ), as reported by the Bayesian change point (bcp) analysis (Figure **IV.2**). A second change point characterizes another shift in the community structure in the 1960s, accompanied by a decrease of the above indices (Figures **IV.1** and **IV.2**).

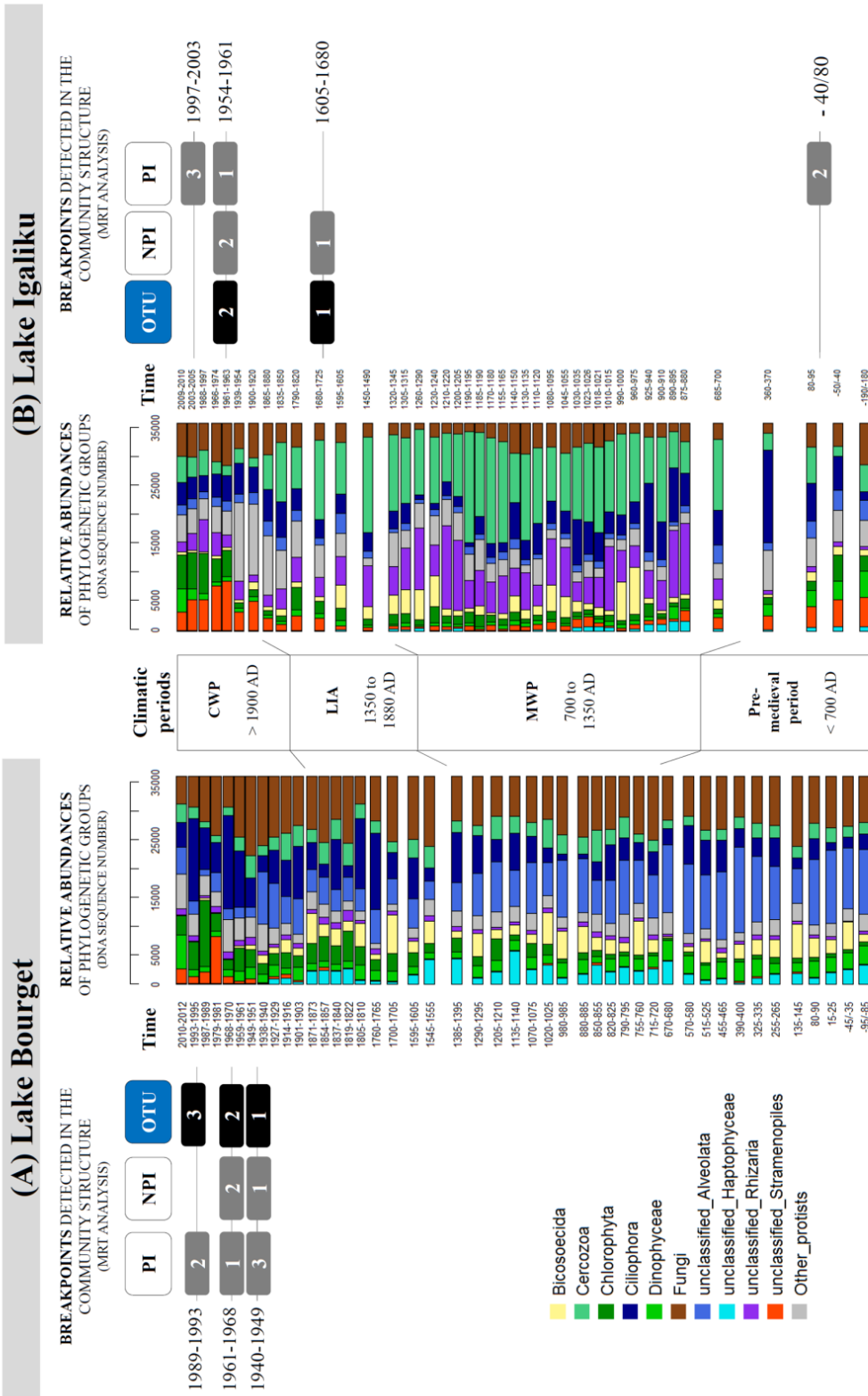


Figure IV.1. For Lake Bourget (A) and Lake Igaliku (B) : Temporal dynamics of the relative abundance of each phylogenetic group of microbial eukaryotes and breakpoints detected by MRT analysis considering the whole list of OTUs and separately non-pigmented OTUs (NPI) and pigmented OTUs (PI). The breakpoints of first, second and third importance are respectively represented by the numbers 1, 2 and 3.

For Lake Igaliku, the main change point in the whole community structure occurred between 1605 and 1680 AD, i.e. during the coolest period of the LIA (Figure **IV.1**). The second most important change point occurred in the 1960s. The OTUs richness tends to increase progressively from AD 700, while the diversity estimates derived from Shannon and Simpson indices show an increase only after LIA cooling (Figure **IV.2**).

While non pigmented OTUs exhibit synchronous shifts with the whole community, the temporal rearrangements of the pigmented community follow a different timing and remarkably their major change point occurred in the 1960s for the 2 lakes (Figure **IV.1**).

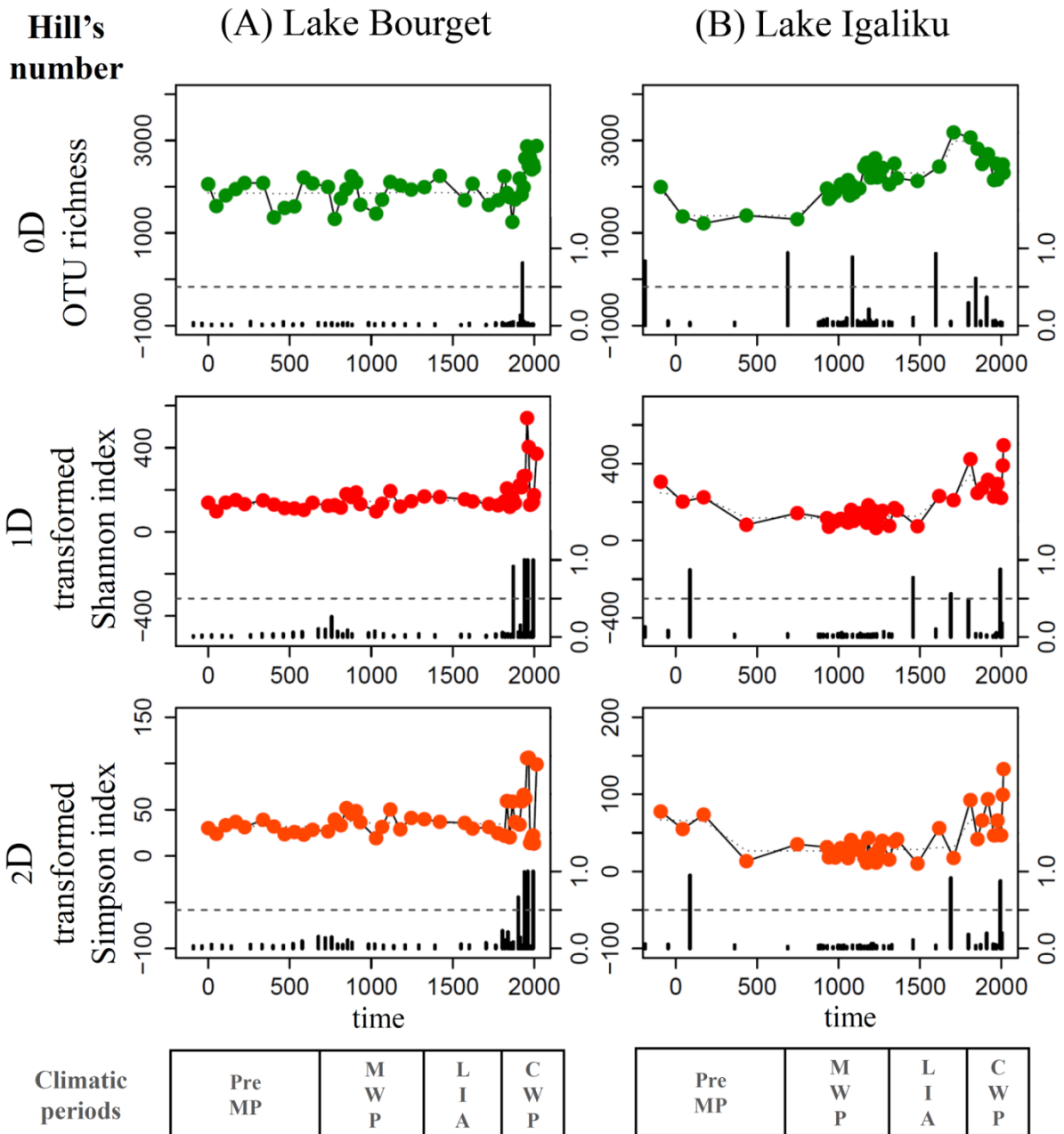


Figure IV.2. Time series of Hill's numbers (OTU richness, transformed Shannon and Simpson indices) for Lake Bourget (A) and Lake Igaliku (B). In the upper part of the graph, the colored line shows the metrics values and the dashed line shows the sample mean for the time series. The results of bcp analysis are plotted on the lower part of each graph showing the posterior probability of a change point at a given time. Timing changes are considered as relevant if the posterior probability values are equal or superior to 0.5. Four climatic periods are mentioned: Pre-medieval Period (preMP), Medieval Warming Period (MWP), Little Ice Age (LIA) and Current Warming Period (CWP).



#### IV.4.3. Community rearrangements over time

For the analysis of beta diversity, both Jaccard (presence/absence of OTUs) and Bray-Curtis (relative abundances of OTUs) indices were considered. Using non-metric multidimensional scaling (nMDS), similar patterns of samples dissimilarity were visualized for Jaccard and Bray-Curtis indices (Figure IV.S5). Further exploration of dissimilarity data was therefore continued by considering only relative abundances (Bray Curtis indices).

Three groups of samples are discriminated for the Lake Bourget time series (Figure IV.3): before the 1940s, from the 1940s to the 1960s and after 1960s. In Lake Bourget, the non-pigmented OTUs are the major contributors to the rearrangements over time (Figure IV.4). Altogether, they explain 63 % of the total dissimilarity between these 3 sample groups. The main non-pigmented contributors to the 1940s rearrangements are the Ciliophora OTU3409 and 2 Fungi OTUs assigned to Cryptomycota (OTU39169) and Chytridiomycetes (OTU8852). Two of these OTUs are also determinants to the 1960s rearrangement, but then exhibit opposite dynamics (increase *versus* decrease in terms of relative DNA sequences abundance; Figure IV.4). The pigmented taxa entail 2 major contributors to the community rearrangements, including one Haptophyceae (OTU18217), which DNA sequences abundance drops since the 1940s, and one Chlorophyceae reliably assigned to *Chlamydomonas* sp. (OTU42983) characterized by a marked increase since the 1960s (Figure IV.4).

In Lake Igaliku time series, 4 groups of samples are defined: before AD 700, from AD 700 to 1350, from AD 1450 to 1880 and after 1880 AD (Figure IV.3) consistently with the climatic periods previously described (MWP, LIA and CWP). The main contributors to the rearrangements are non-pigmented taxa with mainly 5 OTUs belonging to Cercozoa and to unclassified Rhizaria; these 5 OTUS are particularly abundant during the period AD 700-1300 (Figure IV.4). The Table IV.S3 presents the list of the 97 and 118 OTUs identified as the main contributors to dissimilarity in Lake Bourget and Lake Igaliku, respectively.

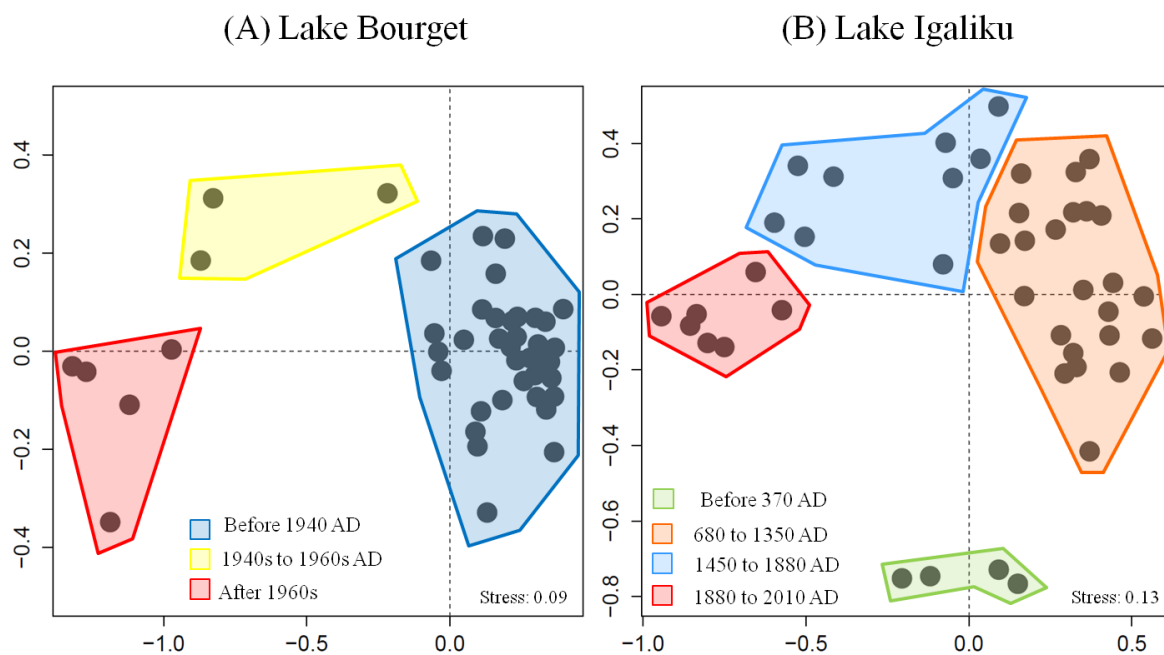
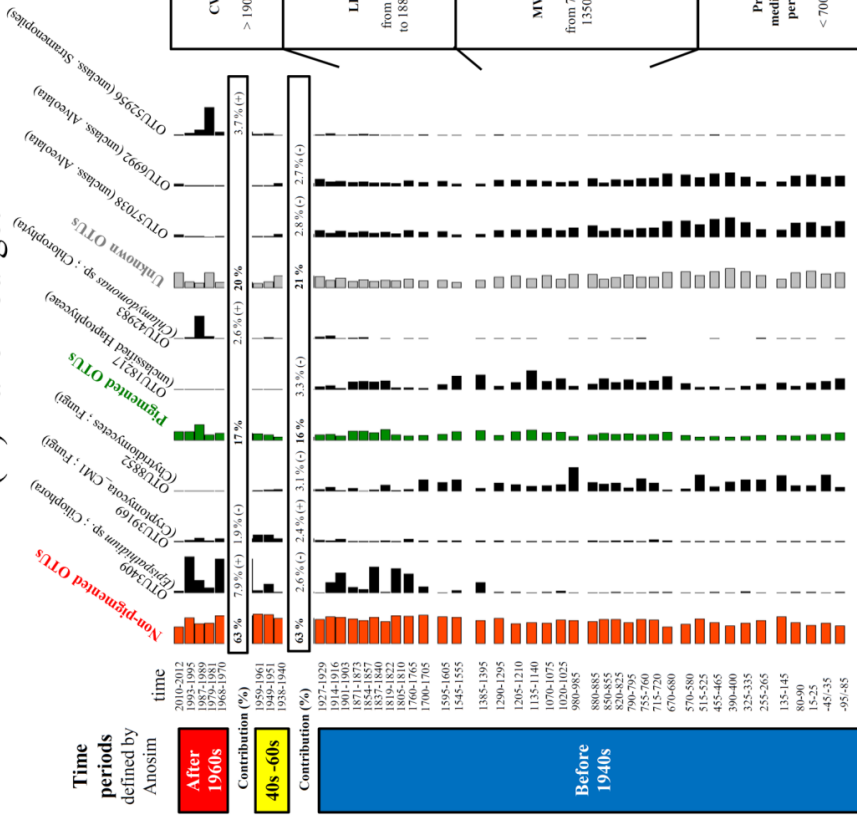


Figure IV.3. Non-metric multidimensional scaling analysis of microbial eukaryotes communities based on Bray-Curtis dissimilarity are presented for Lake Bourget (A) and Lake Igaliku (B) time series. The colors of each sample correspond to the groups of samples identified by ANOSIM analysis ( $P < 0.001$ ).

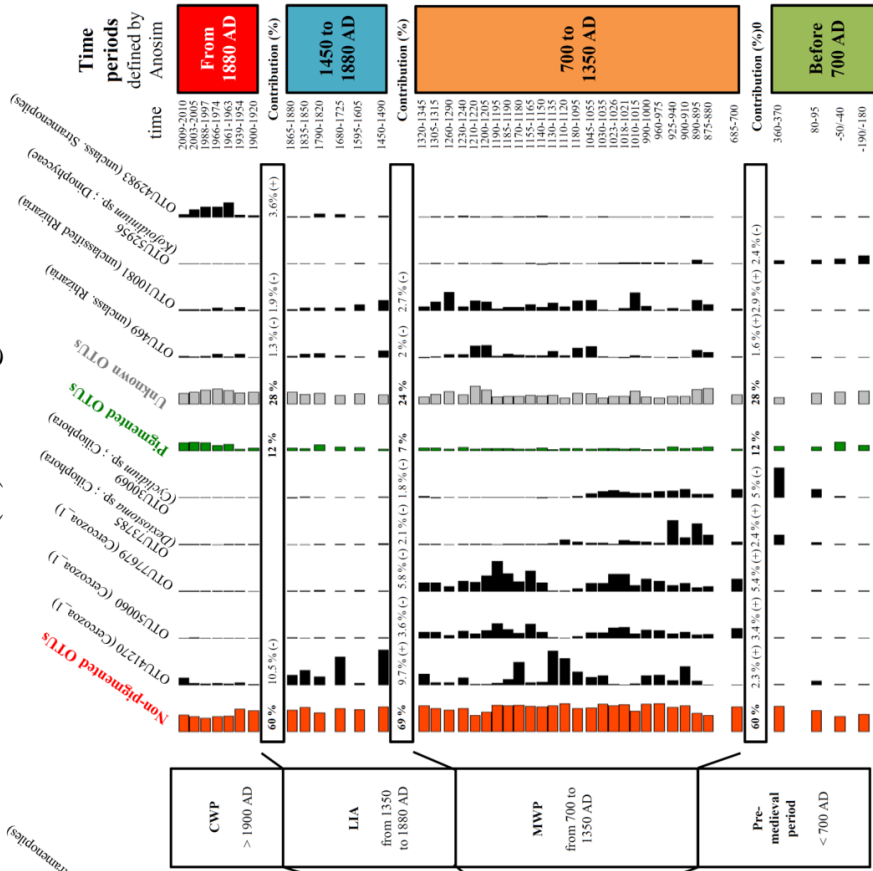
Figure IV.4. Temporal relative abundances of the non-pigmented, pigmented and unknown OTUs, as well as the OTUs that contributed to more than 2 % of the total dissimilarity over time, are presented for Lake Bourget (A) and Lake Igaliku (B). The considered time periods are those defined by ANOSIM analysis. The contribution of OTUs (expressed as %) to the community dissimilarity between consecutive time periods are obtained from SIMPER analysis. The sign "+" means an increase of the relative abundance of the OTU between the 2 compared periods, while the sign "-" means a decrease of this relative abundance.

### (A) Lake Bourget



Relative abundances (% of DNA sequences)

### (B) Lake Igalku



Relative abundances (% of DNA sequences)

The permanently rare OTUs, representing 10 545 and 10 442 OTUs in Lake Bourget and Lake Igaliku respectively, contribute to less than 20 % of the temporal rearrangements over time. The oscillating OTUs (214 and 235 OTUs in Lake Bourget and Lake Igaliku, respectively ; Table IV.1 & Figure IV.5), being abundant or rare for long time periods, predominantly contribute to the dissimilarity in both lakes time series. 72 % and 81 % of the major contributors, in Lake Bourget and Lake Igaliku respectively, are indeed oscillating taxa (Table IV.S3). Among these oscillating taxa, which became particularly abundant since the 1960s, we noticed 5 Chlamydomonadaceae OTUs that exhibit similar temporal dynamics in both lakes. Chlorophyta globally increase (in terms of relative DNA sequences abundance) since the 1960s in both lakes, and a shift of dominance from Ulvophyceae (Lake Bourget) or Trebouxiophyceae (Lake Igaliku) to Chlorophyceae OTUs is observed (Figure IV.6, Table IV.S3).

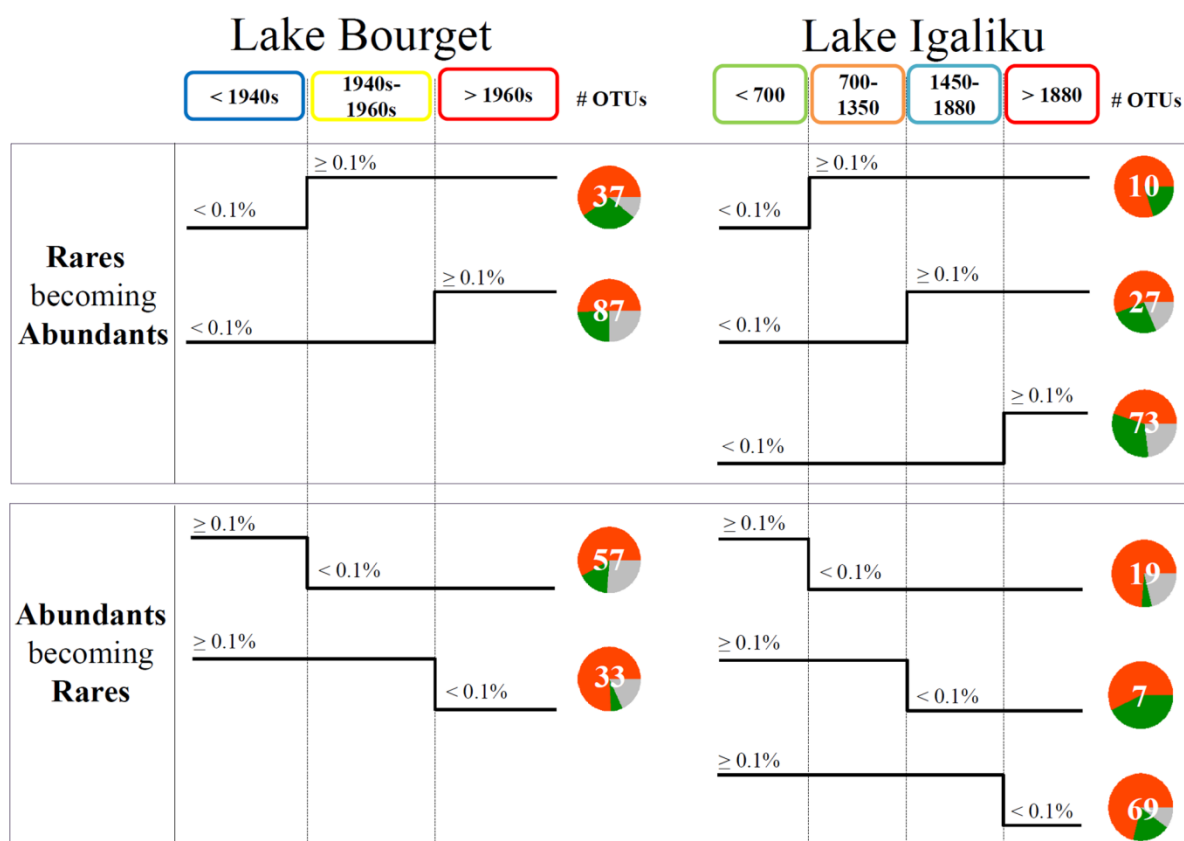


Figure IV.5. Temporal dynamics of oscillating OTUs in Lake Bourget and Lake Igaliku. Oscillating OTUs shifted between abundant (> 0.1 %) and rare states (< 0.1 % of the number of DNA sequences, on average for each period).

The selected time periods are those defined by the ANOSIM analysis. The number of OTUs is presented within pie-charts showing the proportion of non-pigmented (orange), pigmented (green) and unknown (gray) OTUs.

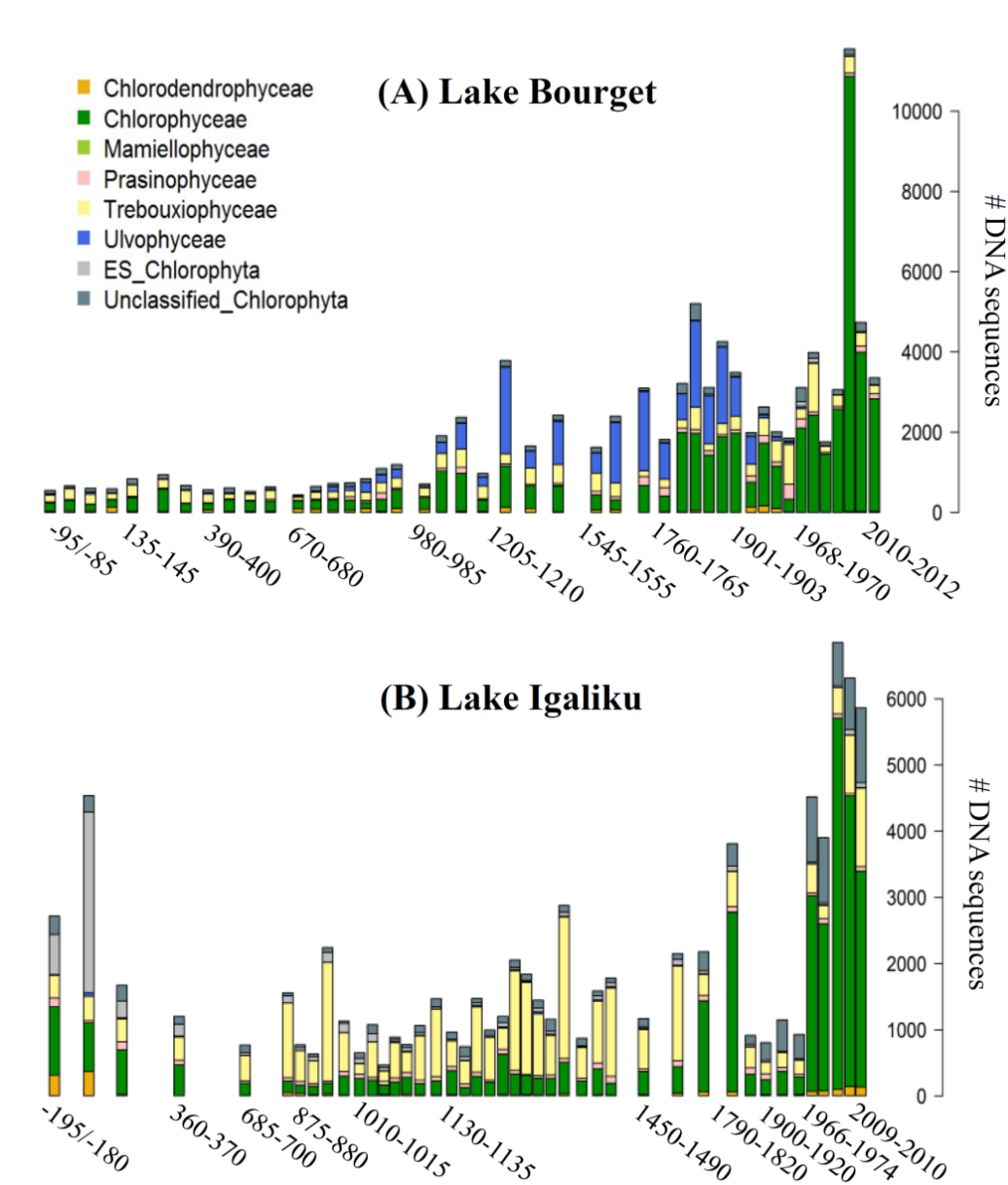


Figure IV.6. Composition and temporal dynamics of the chlorophytan groups in Lake Bourget (A) and in Lake Igaliku (B) based on their relative abundance in terms of number of DNA sequences.

## IV.5. Discussion

### IV.5.1. Methodological aspects

Various issues should be addressed to achieve robust findings using paleo-environmental DNA, some of which can be integrated rather easily in paleo-reconstruction studies (like controls of non-contamination, optimization of protocols for fragmented DNA) while other issues are complex to clarify as for instance the problem of potential differential preservation that is influenced by both physical, chemical and biological parameters (e.g. Boere *et al.* 2011a, Jørgensen *et al.* 2012, Lejzerowicz *et al.* 2013, Torti *et al.* 2015).

In this study, in addition to the strict clean-lab procedures applied, the blanks used at various analytical steps allowed to verify the absence of contamination by external DNA. Only one of the control of aerial contamination was positive for PCR amplification. The sequencing results obtained for this positive blank showed that the OTUs found to be contaminants were mainly represented by Fungi and Alveolata, which account for 95 % of the DNA sequences (25% of Saccharomycetales, 21 % of Agaricomycetes, 10% of unclassified Ciliophora and 39 % of unclassified Alveolata). None of these dominant OTUs were retrieved in the DNA inventory obtained for our environmental samples.

We also considered the potential limitations due to the level of DNA preservation and to the targeted barcodes' length. The preservation of DNA is indeed a key issue in paleogenetics (e.g. Dell'Anno *et al.* 2002, Corinaldesi *et al.* 2008, Boere *et al.* 2011a). The diagenetic processes can degrade sedimentary DNA into shorts fragments thus preventing or limiting the efficiency of PCR amplification especially when the targeted barcodes are long (i.e. > 300 bp). In this study we tested the DNA amplification using 2 different lengths of barcodes (~ 400 bp and ~260 bp). For the longer fragment, we detected a loss signal in the sediment cores from ~ AD 900, showing a potential limit in the amplification of this barcode (data not shown). The 260 bp-long DNA fragment was successfully amplified over the entire sediment cores (up to ~ 100 BCE) and was thus used for reconstructing the past eukaryotic diversity. We revealed the presence in the sedimentary archives of all the taxonomic groups that are usually observed in planktonic molecular surveys (e.g. Debroas *et al.* 2015). Additionally, we did not observe any modification in richness that could be visibly due to diagenetic processes (loss of OTUs when sediment ages, for instance).

Another potential bias in the reconstruction of past biological diversity based on sedimentary DNA is the potential differential preservation of DNA linked to physical and chemical

conditions in the lake as well as to the source organisms, i.e. taxa with resistant cells as resting propagules allowing physical protection for the DNA *versus* taxa that leave their DNA with a limited protection (simple cell membrane easily lysed) or release DNA in the environment (extracellular DNA adsorbed on mineral and organic particles) (Corinaldesi *et al.* 2005, 2008, Pedersen *et al.* 2015). We carefully considered the changes in the relative decrease or disappearance of certain groups in our temporal dataset in order to verify if the decrease of some taxa in the oldest sediment layers could be due to a low preservation capacity and inversely for the taxa producing resistant cells. The taxa known to arbor rigid materials (Bacillariophyta, Dinophyceae) did not show a particularly high abundance in the oldest sediments. We paid special attention to Chlorophyceae since several OTUs within this group (especially *Chlamydomonas* OTUS) were largely represented in upper layers and had a decreasing relative abundance from the 1960s. Chlorophyceae are known to exhibit resistant rigid cell wall and zygospore (Imam *et al.* 1985, VanWinkle-Swift & Rickoll 1997), consequently this group does not have a particularly high sensitivity to degradation during burial in sediment in comparison to other unicellular eukaryotes. The only group for which we suspect a very low level of preservation is the group of Cryptophyta, which is known to be abundant in Lake Bourget plankton over the last 20 yrs but is however poorly retrieved in the recent sediments, as already reported by Capo *et al.* (2015), and is not retrieved in ancient sediments either (this study).

Although careful calibration studies are still to be conducted for DNA proxy as with any paleolimnological proxy, based on methodological precautions applied here and the stringency of our data treatment (analytic and bioinformatics procedures allowing to minimize the inflation of false rare OTUs) we are confident on the validity of these molecular data to describe microbial eukaryotes communities including the rare biosphere.

#### **IV.5.2. Sedimentary DNA reveals past microbial eukaryotes community**

The high richness sequenced from both lakes entails OTUs belonging to all the phylogenetic groups of microbial eukaryotes encountered in lacustrine systems. High latitude ecosystems generally exhibit a reduced biodiversity (aquatic plants and animals) compared to lower latitude systems, which is mainly due to environmental constraints and remoteness (Vincent *et al.* 2013). In our data we could not detect any significant difference in terms of OTU richness between the arctic Lake Igaliku and the temperate Lake Bourget, whatever the considered

time period. The magnitude of the OTU richness characterizing our time-series datasets as well as the OTU distribution within taxonomic groups were within the expectations of previous HTS inventories performed on lake water columns, with Fungi, Ciliophora and Chlorophyta being the most diverse groups (Charvet *et al.* 2014, Oikonomou *et al.* 2015, Kammerlander *et al.* 2015, Debroas *et al.* 2015). Only the study reported by Coolen *et al.* (2013) about the reconstruction of temporal changes within aquatic eukaryotes communities over the Holocene in the Black Sea, is rather comparable to this study. A thorough comparison between these 2 studies would however be tricky due to the numerous methodological differences (sequencing technology, targeted barcode regions ...). From the dominant groups of microbial eukaryotes retrieved in the two studied lakes, we identified a pool of core/persistent taxa, which included OTUs affiliated to Cercozoa, Fungi, Ciliophora, Haptophyceae and Chlorophyta. There is little evidence for the existence of persistent, core microbiomes based on time-series environmental diversity data, especially for microbial eukaryotes (Shade & Gilbert 2015). Our results support recent observations showing the existence of persistent bacterial taxa being mostly represented by abundant OTUs (Gibbons *et al.* 2013). More particularly we show that changes in the dominant OTUs (from one lake to another or through time) primarily reflected shifts in the relative abundance of the persistent OTUs.

#### **IV.5.3. A temporal view on dominance and rarity**

We found a core of 73 OTUs that is not only ubiquitous (detected in all samples) but that also comprises the most sequenced OTUs (i.e. most abundant). This is consistent with the general pattern observed for HTS microbial inventories showing the dominance (in terms of number of DNA reads) of few OTUs, i. e. hyperdominance concept as reported de Vargas *et al.* (2015) in oceans. However, we revealed that the relative abundance of more than half of these persistent OTUs displayed oscillating patterns and were even rare during long time periods. This implies that a formal hyperdominance concept should incorporate knowledge on the temporal dynamics of microbial eukaryotes communities. Shade & Gilbert (2015) recently emphasized the importance of temporal investigations into microbial dynamics in order to determine the role of the rare taxa in ecosystem resilience and stability. The highly diversified rare biosphere is considered either as a pool of active taxa living at low abundance, and/or a reservoir of dormant taxa that provide ecosystems with a biological buffering capacity to withstand environmental changes, and hence is generally associated to the idea of functional



redundancy (Lynch & Neufeld 2015, Logares *et al.* 2014). The rare biosphere represents a dynamic but overlooked fraction of lacustrine diversity. It has been demonstrated, from 18S rRNA, that dominance patterns among microbial eukaryotes could change at short or intermediate time scales ranging from day to seasons or years (Nolte *et al.* 2010, Jones *et al.* 2013, Mangot *et al.* 2013, Egge *et al.* 2015). This suggests that rare taxa take active part into the rearrangements of community, especially when these are subject to disturbance (Sjöstedt *et al.* 2012, Shade *et al.* 2014, Baltar *et al.* 2015, Shade & Gilbert 2015). Our analyses suggest that the eukaryotic rare biosphere participate in the ecosystemic response to disturbances linked to both anthropogenic and climatic changes, illustrating the potential of the paleolimnological approach to detect which rare taxa exhibit a deterministic behavior. Further paleogenetics studies should be dedicated to the study of the rare biosphere dynamics inferred from sedimentary DNA in order to reinforce our observations and test the generalization of such patterns.

#### **IV.5.4. First clues of synchrony between MECs shifts and climatic fluctuations**

Recent paleoecological studies investigated the effects of climate (i.e. mainly temperature) on biological taxa but generally focused on specific photosynthetic groups such as Cyanobacteria, Haptophyta or Chlorophyta (Domaizon *et al.* 2013, Hou *et al.* 2014, Savichtcheva *et al.* 2015). The whole community of microbial eukaryotes (including heterotrophs) play key roles in food webs, they are seldom accounted for in studies dealing with climate impacts (Caron & Hutchins 2013, Coolen *et al.* 2013). In the case of Lake Bourget, we detect no marked effect of climatic fluctuations on community dynamics. However, it is noteworthy that, by excluding the data obtained for the last century, we detect changes concomitant with climatic modifications in this lake. Indeed, excluding the recent eutrophication period and considering data for the period -100/1900 AD, we verify the existence of a shift in the eukaryotic community structure at the end of the medieval warming (around 1350 AD, when temperatures started to decrease; data not shown). The recent environmental forcings (mainly eutrophication and recent warming effects) seem to erase the more subtle effects of ancient climatic fluctuations in this lake.

In contrast, we detect major changes in Lake Igaliku community structure that could be related to main climatic periods. We identified a major change point during the 17<sup>th</sup> century, which corresponds to the coolest period of the last 2200 years (between AD 1640 and 1780; Millet *et al.* 2014). Moreover, the microbial eukaryotes community appeared to be structured

according to the main climatic periods (MWP, LIA, CWP; McKay & Kaufman, 2014). Such modifications in microbial assemblages are not necessarily related to the direct effects of temperature but could also result from complex indirect effects of climate change (modifications in the hydrological conditions in the watershed and lake's tributaries). The previous studies based on the morphological remains (Perren *et al.* 2012, Millet *et al.* 2014) reported subtle changes over the last 1450 yrs and until 1980 AD. Our results are rather different since both marked modifications in the richness and structure of microbial eukaryotic assemblages are detected concomitantly with known climatic fluctuations. Although our results do not allow us ranking the importance of climate and anthropogenic drivers, we hypothesize that the changes in microbial eukaryotic structure that occurred during the MWP are primarily linked to climatic modifications. The shift is indeed concomitant with the beginning of the warming (700 AD) while the human occupation started in the late 900s. In addition there is evidence that Norse agriculture during this period had almost no impact on the trophic state of the Lake Igaliku, except subtle changes of its flora and fauna (Perren *et al.* 2012, Millet *et al.* 2014). Here, we highlight profound changes in the dominant taxa between the 3 consecutive climatic periods, this paleolimnological view provides new material for generating hypotheses regarding the impact of climate on eukaryotic taxa. More particularly, our data suggest that the relative abundance of Cercozoa OTUs could be linked to the temperature fluctuations with the highest record of this group during the MWP. The transition between the MWP and the LIA was indeed marked by a strong decrease of Cercozoa OTUs relative abundance and of other closely related OTUs. Cercozoans predate bacteria and micro-algae in planktonic and benthic ecosystems, but may also parasite large phytoplanktonic taxa (Tilman *et al.* 1999). The positive response of Cercozoa to the increase of temperature was also recently reported in arctic peat (Tveit *et al.* 2015) but extrapolating to various lake systems requires further data and research.

#### **IV.5.5. The impacts of marked eutrophication seems to overpass that of other forcing factors**

The eutrophication of the water column constitutes a strong disturbance for fish, zooplankton and phytoplankton assemblages (e.g. Jeppesen *et al.* 2000, Paerl 2006, Schindler 2006, Ibelings *et al.* 2007). Here, from our data we can reasonably assume that the impact of eutrophication is detectable on the whole diversity of microbial eukaryotes community. In Lake Bourget, its effect seems to override that of climatic fluctuations as measured by the

response of community OTU richness and structure. Indeed, the main changes were detected concomitantly with key events characterizing the process of eutrophication, namely when it started (in the 1940s; **Figure IV.S4**) and when the phosphorus concentration peaked (in the 1960s; **Figure IV.S4**), while no significant change was detected before. Other studies based on various biological proxies highlighted the effect of eutrophication in Lake Bourget since the 1940s (Millet *et al.* 2010, Berthon *et al.* 2013, Alric *et al.* 2013, Perga *et al.* 2015). This eutrophication indeed was reported as the dominant driver of the relative abundance cyanobacterial species (Domaizon *et al.* 2013, Savichtcheva *et al.* 2015) but also of the composition and size structure of *Daphnia* sp. population (Alric *et al.* 2013). Here, our data tend to confirm the eutrophication threshold effect that causes a disruption of the biological compartments affecting a large diversity of microorganisms, and most probably the whole food web structure. While the first shift observed in the 1940s can be attributed to the eutrophication (based on the congruence of several proxies and studies), for the second change point occurring in the 1960s, the impact of climate cannot yet be excluded (increase of air temperature in the 1960s and from the 1990s), and the respective contribution of local anthropogenic pressure and global climatic conditions is still uncertain.

#### **IV.5.6. Toward ‘new’ biological markers from sedimentary DNA**

Sed-DNA provides a huge amount of information about past and diverse communities, including potential biological markers that have never been explored before. For instance we reveal that the genus *Chlamydomonas* was constantly present throughout the last 2200 years in both lakes but only dominated Chlorophyta since the 1960s. This shift within the structure of the Chlorophyta community could be a potential indicator of recent increase in trophic status and temperature conditions. Indeed, *Chlamydomonas* is ubiquitous and abundant in numerous lakes, but seems particularly successful in eutrophic conditions (Naselli-Flores & Barone 2000, Lepère *et al.* 2008, Chen *et al.* 2008). Algal remains preserved in subsurface lake sediments have been previously analyzed to reconstruct Chlorophyta abundance profiles, showing their increase during the eutrophication period (60s-70s) in Lake Biwa (Tsugeki *et al.* 2009). However, the taxonomic resolution of such analyses was limited to the few taxa for which morphological remains could be identified.

Haptophyceae is another pigmented group that responded markedly to the modification of trophic state. Haptophyta taxa disappeared from the whole assemblage at the beginning of the eutrophication process of Lake Bourget in the 1940s. The sensitivity of this group to

eutrophication is consistent with previous knowledge about their preference for oligotrophic conditions (Hou *et al.* 2014, Unrein *et al.* 2014).

Additionally, based on the assignation of OTUs to basic trophic modes (i.e. pigmented and non-pigmented), we show that the changes observed since the 1960s were detected not only for Chlorophyta and Haptophyta, but more globally in the structure of the whole pigmented community, and in both lakes. This observation is consistent with growing body of evidence that local human activities impose a widespread force shaping the structure of phytoplankton communities. This interpretation is in line with that of Taranu *et al.* (2015), based on pigment analyses of 108 sedimentary time series from north temperate-subarctic lakes.

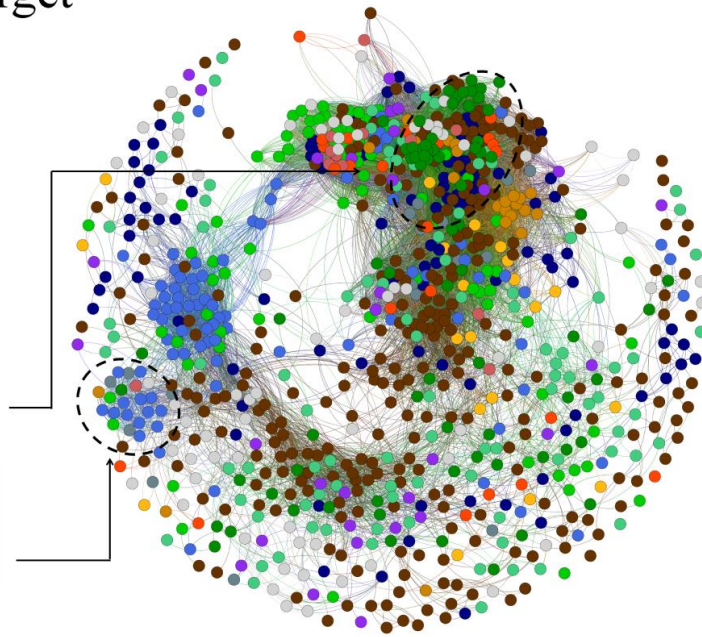
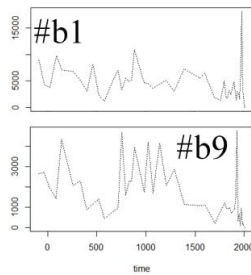
Regarding non pigmented taxa, Ciliophora appeared as an interesting group to focus on. This group is known to be a good indicator of trophic status (Beaver & Crisman 1989, Kisand & Zingel 2000, Pawlowski *et al.* 2016) and has already been used for saprobic index (Foissner 1988, Madoni 2005). Here we detected more particularly the response of *Epispathidium* sp., but to our knowledge, this genus was not particularly described as an indicator of trophic modification. The application of ciliates metabarcoding for bioindication is still to be developed (Pawlowski *et al.* 2016), most probably the optimization of such HTS biomonitoring would represent an interesting potential in paleolimnology.

For some other non-pigmented taxa responding to environmental changes over the past 2 millennia, the OTUS could not be precisely assigned to referenced taxa and consequently their ecological roles are unknown hampering any hypothesis to be drawn regarding their link to climatic or trophic pressures. This is the case for instance for several OTUs within Cercozoa found dominant during the warmer medieval period in Lake Igaliku.

The on-going characterization of undescribed microbial eukaryotes should provide a more complete view of the ecological roles of these taxa and thus improve the accuracy and broaden the scope of microbial DNA inventories (Zinger *et al.* 2012, del Campo *et al.* 2014, de Vargas *et al.* 2015), with application both in limnology and paleolimnology.

## (A) Lake Bourget

Metrics	Values
Number of nodes	913
Number of edges	25575
Average Degree	28,012
Diameter	10
Density	0,061
Weighted Modularity	0,493
Number of modules	28
Clustering Coefficient	0,633
Average path length	3,624



## (B) Lake Igaliku

Metrics	Values
Number of nodes	902
Number of edges	29621
Average Degree	32,839
Diameter	8
Density	0,073
Weighted Modularity	0,655
Number of modules	37
Clustering Coefficient	0,705
Average path length	3,321

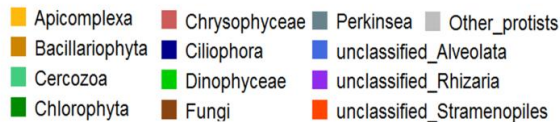
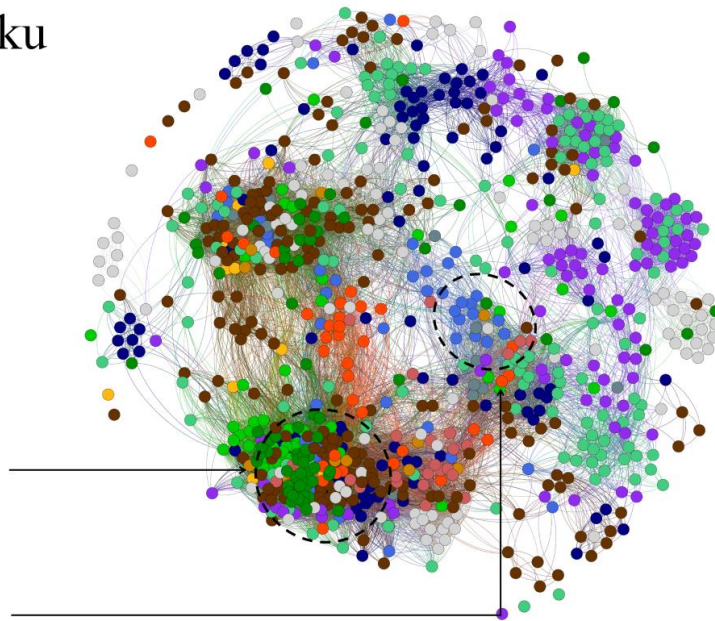
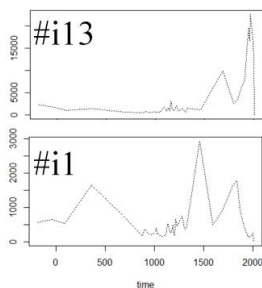


Figure IV.7. Co-occurrence networks for Lake Bourget (A) and Lake Igaliku (B). The nodes represented each one OTU. The colors of nodes represented their taxonomic affiliation. The edges correspond to the Pearson correlation values calculated between each pair of OTUs. Only the edges with values superior to 0.5 for the visualization and the calculations of topological indices (see Table). In each network, the modules found to include significant common clusters between the two networks were surrounded by dotted lines and graph presenting their relative abundance over time were plotted.

## IV.6. Additional analysis

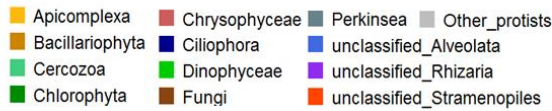
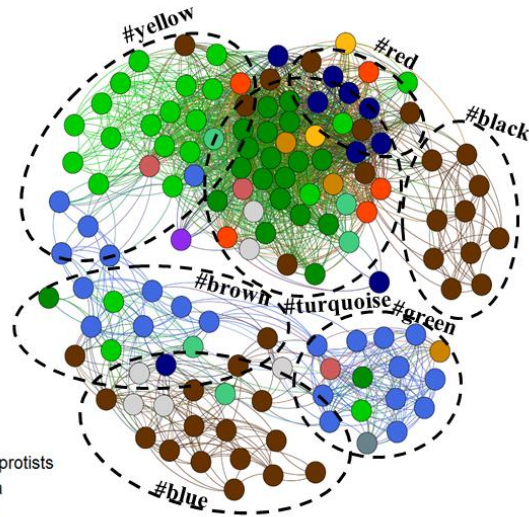
### IV.6.1. Application of network analysis to the view of the temporal dynamics of microbial eukaryotes communities via the network analysis time series

For each lake independently, co-occurrence networks were built from the relative abundance datasets of OTUs. The construction of weighted correlation networks was performed following Langfelder & Horvath (2008) using the R package WGCNA. In order to reduce the complexity of the 2 datasets, all the permanently rare OTUs (OTUs that never reached more than 36 DNA sequences i.e., 0.1 % of the number of DNA sequences per sample) were removed. The relative abundance datasets were standardized (using Hellinger transformation calculated by *decostand* function from the vegan R package; Oksanen *et al.* 2015) for further analyzes. The thresholding power values of 6 were chosen for both networks. The adjacency matrices, obtained by the calculation of Pearson correlation values between each pair of OTUs, were transformed into topological overlap matrix to minimize the possible errors. Subsequently the OTUs were clustered in modules, with a minimum number of 8 OTUs per module. In the two individual networks, only the OTUs (i.e., the nodes) linked by correlations (i.e., the edges) superior to 0.5 were kept for the visualisation and the calculation of network topological indices (average degree, diameter, density, weighted modularity, clustering coefficient and average path length) realized using the Gephi software (Bastian *et al.* 2009). In order to detect clusters of highly-connected OTUs present in both networks, we used a procedure from the WGCNA package based on the construction of a consensus network from the adjacency matrices of the two individual networks,

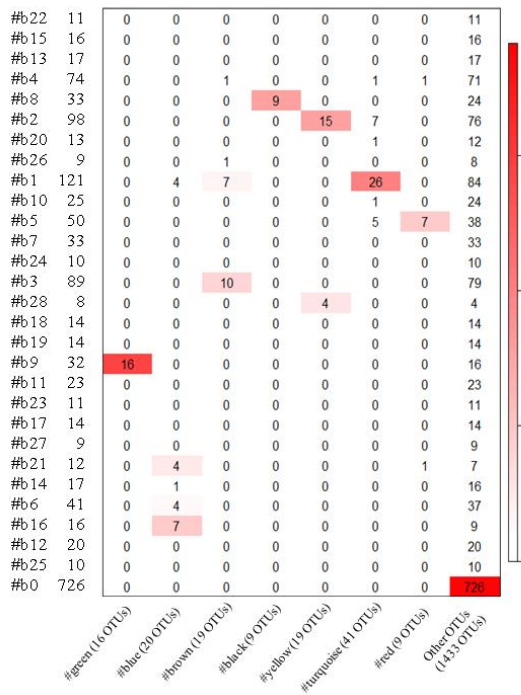
The resulting network from Lake Bourget was constituted by 913 nodes linked by 25 575 edges while in the Lake Igaliku network, 902 OTUs and 29 261 edges were found (**Figure IV.7**). The diameter, showing the distance between the two most distant nodes, was similar for the 2 networks. The two networks were constituted by highly connected OTUs (mean number of edges per node superior to 28) structured among densely connected groups of nodes (modules) and forming clustered topology as expected for environmental networks. Only groups of nodes with more than 8 OTUs were arbitrarily considered as modules; 20 modules were retrieved in both network. In the consensus network, 133 OTUs (out of the total 1566 OTUs) were retrieved within 7 clusters of highly-connected OTUs (called consensus modules; **Figure IV.8**). The consensus modules #turquoise and the module #green exhibit the highest number of OTUs.



(A) Consensus network



(B) Correspondence with Annecy network



(C) Correspondence with Bourget network

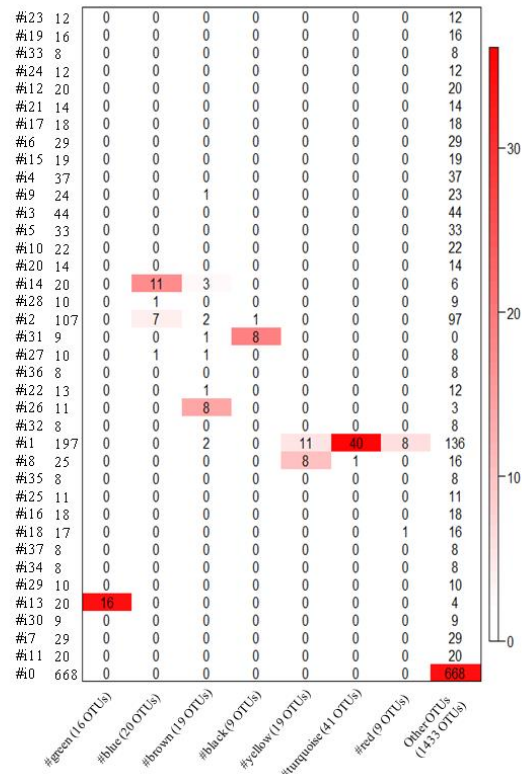


Figure IV.8. Consensus network. (A) Only the OTUs retrieved within the 7 consensus modules were represented. Correspondence of the (B) Annecy set-specific and (C) Bourget set-specific with the consensus modules. Each row of the table corresponds to one module from the individual networks, and each column corresponds to one consensus module. Numbers in the table indicate the number of OTUs in the intersection of the corresponding modules. Coloring of the table encodes  $-\log(p)$ , with  $p$  being the Fisher's exact test  $p$ -value for the overlap of the two modules. The stronger the red color, the more significant the overlap is.

The module #turquoise was mainly composed by Chlorophyta OTUs found to be highly abundant in the recent period from both lake time series. We hypothesize that the increase in the recent period of Chlorophyta could originate from the eutrophication in Lake Bourget from the 1940s to the 1970s (Giguët-Covex *et al.* 2010, Alric *et al.* 2013) and the moderate eutrophication in Lake Igaliku around the 1980s (Austrheim *et al.* 2008, Perren *et al.* 2012, Millet *et al.* 2014). The module #blue was mainly composed by unclassified Alveolata. Those OTUs peaked at different time in both time series. No clear relationship with known environmental changes can be proposed due to the lack of data about environmental parameters for the whole time period.

To conclude, network analyzes allow to reveal common responses of microbial eukaryotes communities from both lakes that could be related to forcing factors changes but further analyzes should be considered to better understand the links between environmental data changes and molecular inventories rearrangements. For instance, additional molecular analyzes based on the PCR amplifications (qPCR or high-throughput sequencing) of specific microbial groups (such as ciliates) and a better characterization of environmental conditions changes over the considered time periods in those lakes would be interesting to confirm the patterns observed in the present network analyzes.



## IV.7. Supporting information

Figure IV.S1. Dating of sediment cores.

Lake Bourget (A): Age depth model based on chemical matching (gray crosses) and varve counting (red points) proposed for LDB13-I (a). The pink area represents the part of the sequence which suffers of chronological uncertainties and has to be considered with caution. Chemical matching using downcore variations of K/Ti ratio in LDB04-I (Arnaud et al., 2012) and LDB13-I (b). The seven tie points identified are noted A to G. Lake Igaliku (B): Age depth model for Iga-2007 and correlation between Iga-2007 and Iga-2011 cores based on  $^{14}\text{C}$ ,  $^{210}\text{Pb}$  and  $^{137}\text{CS}$  data (dots) according to Massa et al.,(2012) (a). The correlation is based on magnetic susceptibility data (b).

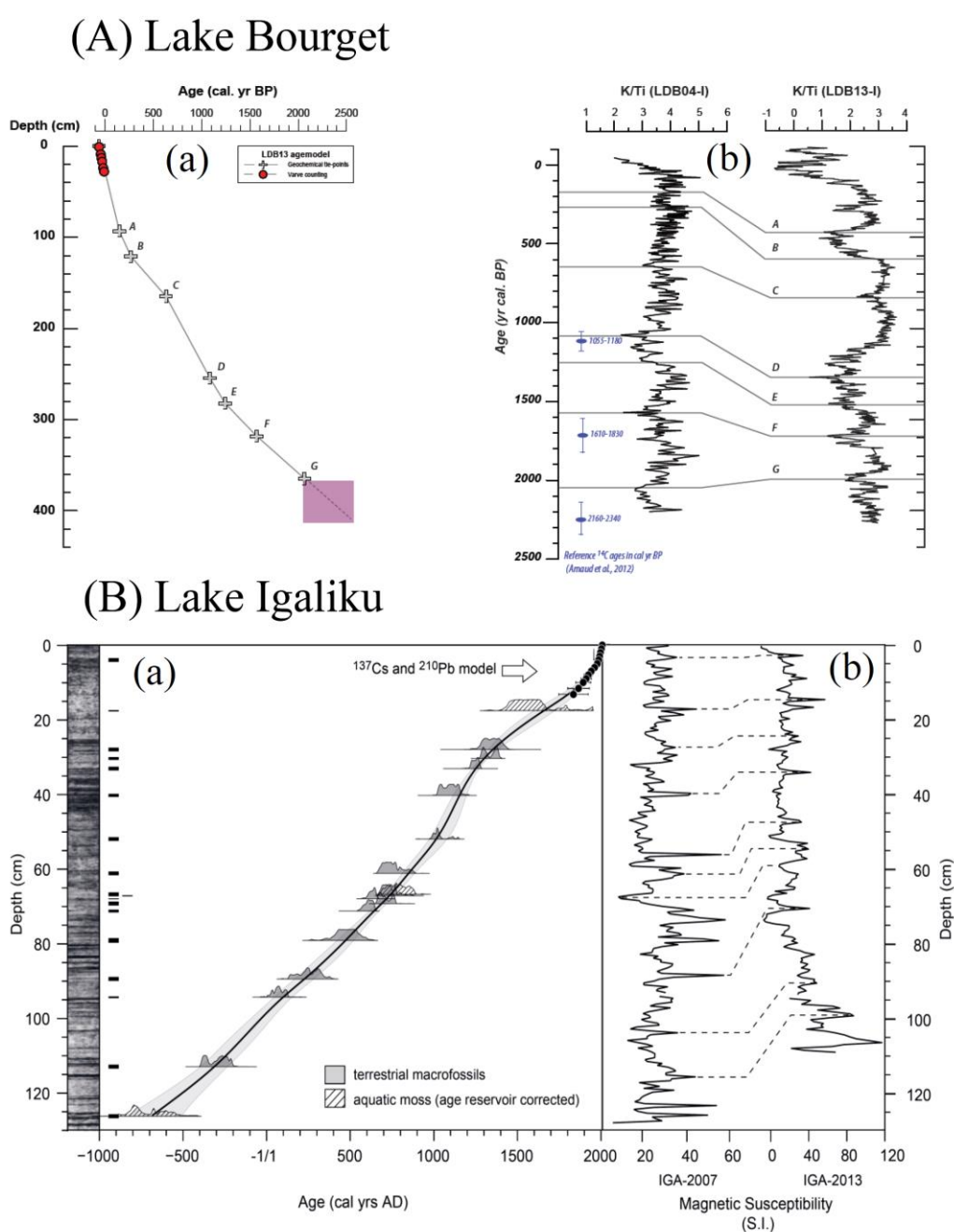


Figure IV.S2. Figure showing from the MRT analysis, the relative error RE and the cross-validated relative error CVRE for the 3 studied datasets (all OTUs, non-pigmented OTUs, pigmented OTUs) in Lake Bourget (A) and Lake Igaliku (B). The solution with the smallest CVRE is indicated (red point), as well as CVRE error bars. The green bars indicate the number of times that the solution was selected as the best during the cross-validation iterations. Black circles show the size of tree selected in the present study.

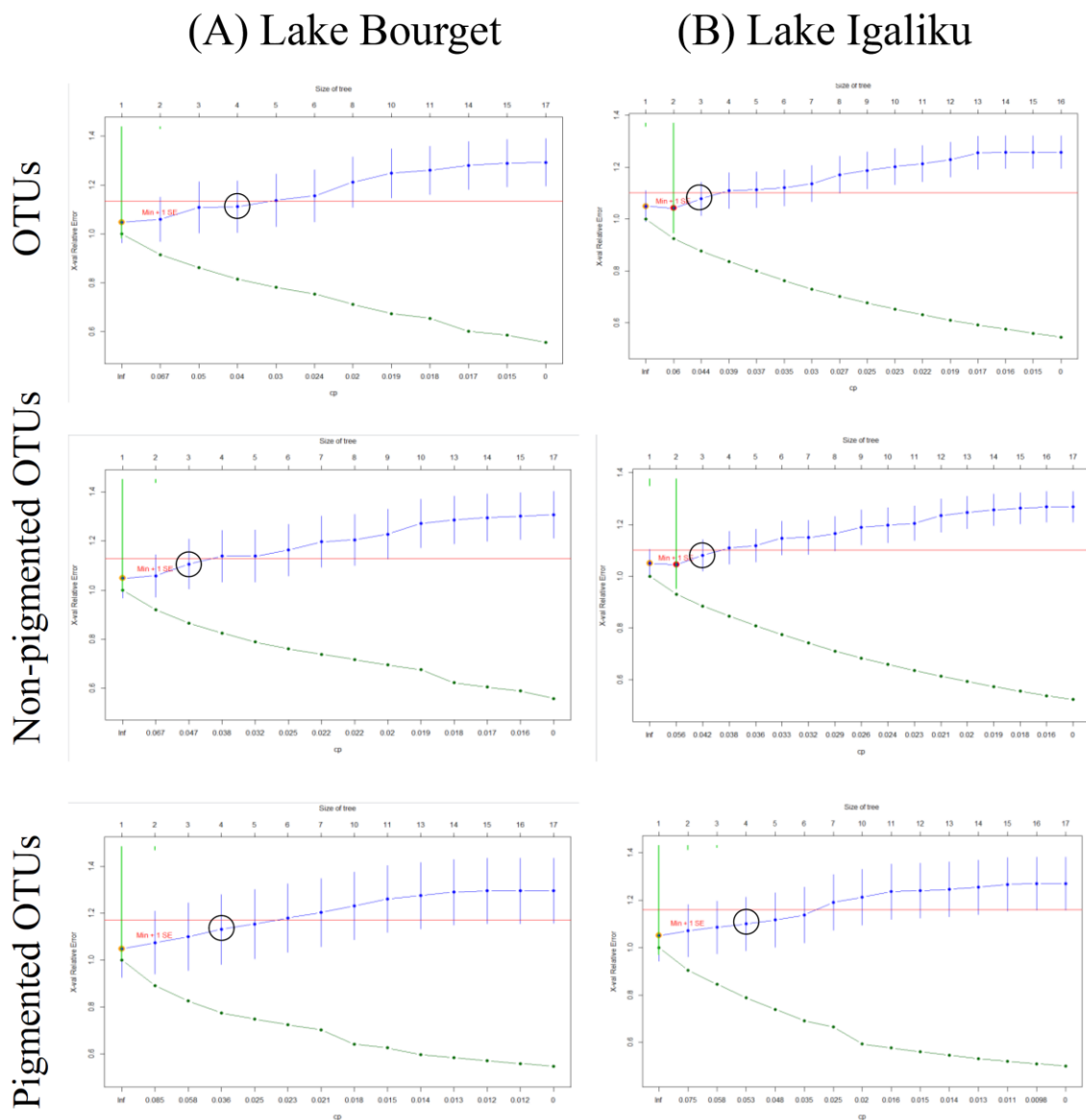


Figure IV.S3. Distribution of richness (in terms of OTUs number) within each phylogenetic group in Lake Bourget (A) and Lake Igaliku (B).

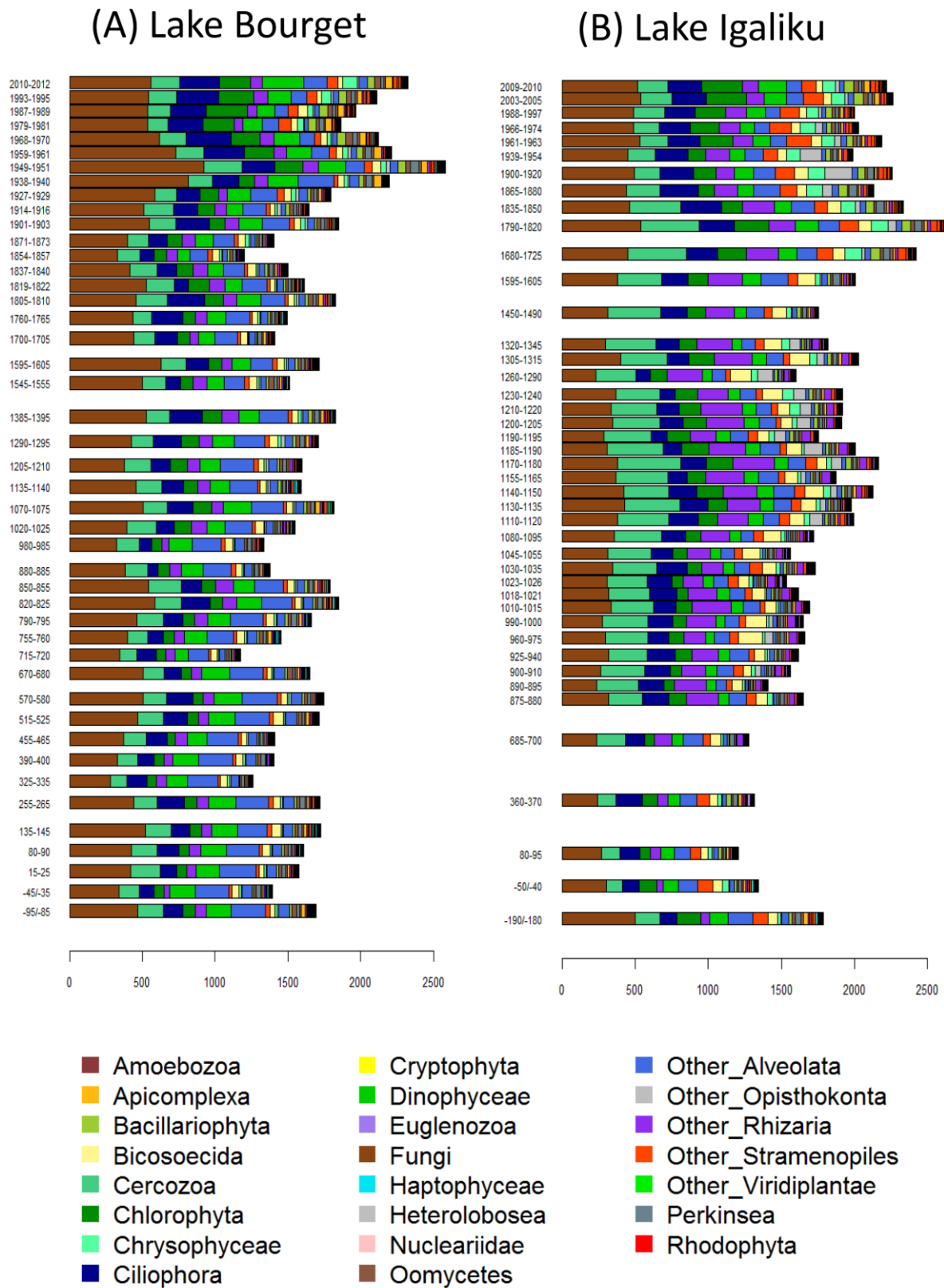


Figure IV.S4. Temporal changes over the last century for (A) the anomalies of air temperature (annual average in  $10^{-1} \text{ }^{\circ}\text{C}$ ) ; Data from Auer et al. 2007 (B) Daphnia-inferred total phosphorus concentrations (annual mean concentrations over the 0-20 m water depths  $\pm$  95 confidence intervals) in  $\mu\text{g P.L}^{-1}$  ; Data from Alric *et al.* 2013; Berthon *et al.* 2014.

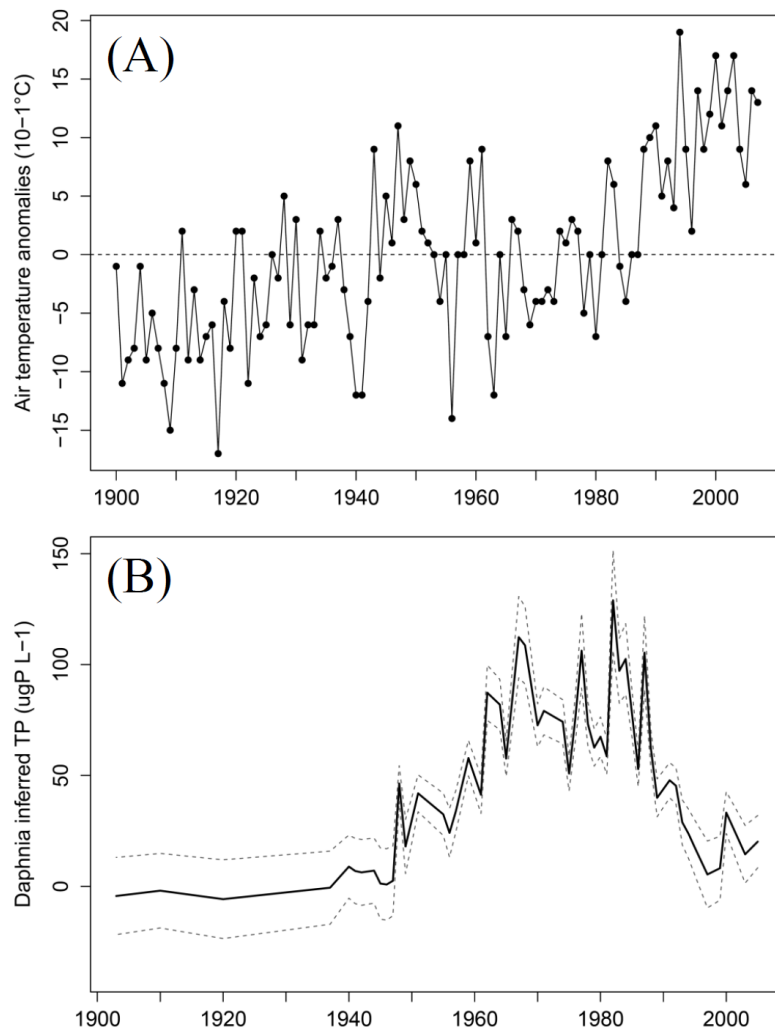


Figure IV.S5. Non-metric multidimensional scaling analysis based on Bray-Curtis (relative abundance of OTUs) and Jaccard (presence/absence of OTUs) indices. Plots are presented for Lake Bourget and Lake Igaliku time series. The colors correspond to the groups of samples identified by ANOSIM analysis ( $p < 0.05$ ).

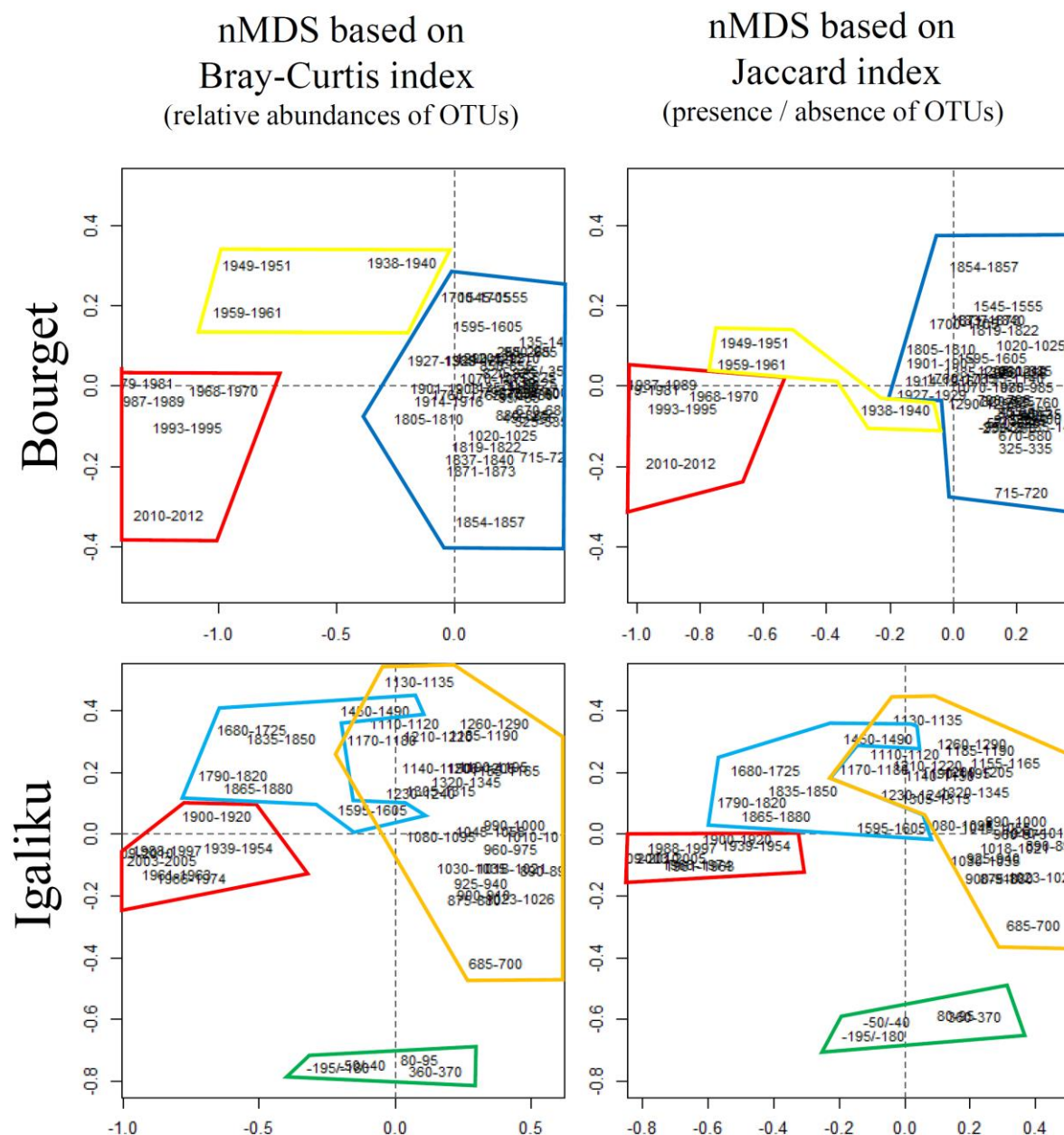


Table IV.S1. (A) Chronology and XRF results from LDB13-I core (B) Chronology and magnetic susceptibility from Iga-2013 core.

See online supporting information

Table IV.S2. Richness and diversity indices obtained for each sample.

Lake	Sample	Composite_depth (cm)	Dating / Time (AD)	# Quality-filter DNA sequences	# Clean DNA sequences	# all OTUs	# Pigmented-OTUs	# Non-pigmented-OTUs	estimated OTU richness	Transformed Shannon diversity	Transformed Simpson diversity
Bourget	BourgetAC0135	1-2	2010-2012	67422 86429	35973	2327	474	1224	2871	373.92	99.10
	BourgetAC0139	9-10	1993-1995	65278 66824	35973	2110	473	1193	2393	174.61	13.45
	BourgetAC0141	13-14	1987-1989	93600 87332	35973	1967	475	1140	2444	140.14	22.19
	BourgetAC0143	17-18	1979-1981	105600 93144	35973	1865	390	1080	2393	132.84	18.94
	BourgetAC0146	24-25	1968-1970	75410 81283	35973	2120	411	1310	2615	129.99	13.98
	BourgetAC0148	28-29	1959-1961	71932 56817	35973	2213	369	1413	2392	394.93	103.21
	BourgetAC0151	33-34	1949-1951	74332 56725	35973	2583	415	1662	2861	540.48	104.57
	BourgetAC011	39-40	1938-1940	78846 97686	35973	2196	289	1344	2610	265.15	61.96
	BourgetAC013	45-46	1927-1929	52083 50840	35973	1796	312	1044	1897	265.45	65.94
	BourgetAC015	51-52	1914-1916	49599 58571	35973	1643	248	1011	1774	211.92	58.41
	BourgetAC017	57-58	1901-1903	52580 79086	35973	1847	284	1102	2100	225.09	34.39
	BourgetAC021	69-70	1871-1873	71354 80966	35973	1405	230	784	1673	136.23	36.14
	BourgetAC023	75-76	1854-1857	23975 74427	35973	1201	197	684	1219	163.49	58.79
	BourgetAC025	81-82	1837-1840	52360 96285	35973	1503	253	876	1739	119.27	20.37
	BourgetAC027	87-88	1819-1822	61264 75026	35973	1613	290	918	1813	206.98	59.22
	BourgetAC029	92-93	1805-1810	90034 74289	35973	1829	321	1090	2175	145.32	21.97
	BourgetAC033	104-105	1760-1765	65030 74086	35973	1494	231	887	1691	124.72	24.21
	BourgetAC037	117-118	1700-1705	73964 74414	35973	1408	217	858	1585	133.01	31.35
	BourgetAC040	132-133	1595-1605	83010 72875	35973	1716	253	1085	2014	145.03	29.63
	BourgetAC042	138-139	1545-1555	34502 86948	35973	1511	223	909	1638	155.54	35.95
	BourgetAC048	156-157	1385-1395	103497 75353	35973	1827	308	1056	2277	166.00	36.90
	BourgetAC052	168-169	1290-1295	81173 74142	35973	1709	295	947	2038	167.78	39.70
	BourgetAC056	180-181	1205-1210	72316 76505	35973	1599	285	823	1901	145.59	41.44
	BourgetAC060	192-193	1135-1140	98907 83517	35973	1590	271	892	1992	122.53	29.46
	BourgetAC064	205-206	1070-1075	70369 65940	35973	1817	334	997	2149	194.90	50.33
	BourgetAC068	217-218	1020-1025	30095 68295	35973	1550	267	846	1722	134.39	31.36
	BourgetAC072	226-227	980-985	53683 37907	35973	1334	228	687	1388	97.79	19.40
	BourgetAC080	250-251	880-885	71817 58721	35973	1378	231	718	1576	131.63	36.48

	BourgetAC0 82	256-257	850-855	68781 54219	35973	1792	271	1041	2048	187.73	48.22
	BourgetAC0 84	262-263	820-825	67604 75581	35973	1848	285	1087	2169	167.68	46.20
	BourgetAC0 86	268-269	790-795	70824 57945	35973	1663	257	926	1913	182.55	51.99
	BourgetAC0 88	274-275	755-760	73337 68370	35973	1452	243	793	1710	117.21	33.49
	BourgetAC0 90	280-281	715-720	103656 45908	35973	1172	164	697	1268	126.59	39.31
	BourgetAC0 92	286-287	670-680	78124 88932	35973	1648	259	888	1973	124.70	26.69
	BourgetAC0 96	298-299	570-580	71304 77757	35973	1745	267	968	2090	138.37	28.36
	BourgetAC0 98	304-305	515-525	85389 83423	35973	1717	258	958	2109	103.01	22.92
	BourgetAC0 100	310-311	455-465	51344 53778	35973	1412	202	765	1545	111.70	25.88
	BourgetAC0 102	316-317	390-400	58575 50686	35973	1403	216	708	1555	112.47	23.33
	BourgetAC0 104	322-323	325-335	68072 20820	35973	1262	188	651	1321	129.27	31.74
	BourgetAC0 106	328-329	255-265	54690 78841	35973	1718	269	938	1989	152.33	40.25
	BourgetAC0 109	338-339	135-145	71295 74394	35973	1723	274	980	2027	130.95	30.86
	BourgetAC0 111	343-344	80-90	66374 74899	35973	1608	260	874	1878	152.33	36.96
	BourgetAC0 113	349-350	15-25	57599 71752	35973	1573	257	839	1789	139.56	33.19
	BourgetAC0 115	355-356	-45/-35	83385 57689	35973	1394	245	715	1612	96.75	23.88
	BourgetAC0 117	361-362	-95/-85	100680 67388	35973	1691	271	921	2006	141.59	31.08
	IgalikuAC0I 2	0.5-1	2009- 2010	51935 64872	35973	2219	510	1171	2300	494.95	132.74
	IgalikuAC0I 4	1.5-2	2003- 2005	72110 76788	35973	2266	486	1189	2482	390.05	99.46
	IgalikuAC0I 6	2.5-3	1988- 1997	69699 65154	35973	2003	372	1074	2325	221.95	46.92
	IgalikuAC0I 8	3.5-4	1966- 1974	67563 54544	35973	2027	365	1087	2156	293.87	66.00
	IgalikuAC0I 10	4.5-5	1961- 1963	72377 72894	35973	2186	417	1158	2499	289.03	47.05
	IgalikuAC0I 12	5.5-6	1939- 1954	54399 66223	35973	1988	275	1158	2144	229.17	46.76
	IgalikuAC0I 14	6.5-7	1900- 1920	83385 80309	35973	2258	347	1279	2705	315.55	93.74
	IgalikuAC0I 16	7.5-8	1865- 1880	61191 77180	35973	2133	292	1172	2491	271.17	66.07
	IgalikuAC0I 18	8.5-9	1835- 1850	67694 65261	35973	2334	312	1351	2823	246.57	41.74
	IgalikuAC0I 20	9.5-10.5	1790- 1820	76119 58735	35973	2643	480	1465	3064	422.69	92.52
	IgalikuAC0I 24	13.5-14.5	1680- 1725	97760 72618	35973	2425	440	1300	3174	209.54	17.42
	IgalikuAC0I 28	17.5-18.5	1595- 1605	86505 72809	35973	2009	264	1104	2436	231.56	56.16
	IgalikuAC0I 32	21.5-22.5	1450- 1490	68297 81142	35973	1755	240	1035	2125	74.72	10.12
	IgalikuAC0I 36	25.5-26.5	1320- 1345	64168 66478	35973	1821	221	1061	2184	157.31	41.70

Igaliku  
u

IgalikuACOI 38	27-27.5	1305- 1315	68018 75059	35973	2029	309	1132	2501	168.83	38.73
IgalikuACOI 40	28-28.5	1260- 1290	79413 76546	35973	1600	210	899	2049	76.49	15.31
IgalikuACOI 42	29-29.5	1230- 1240	59583 111070	35973	1920	306	1075	2406	154.65	39.63
IgalikuACOI 44	30-30.5	1210- 1220	68345 55300	35973	1917	284	1025	2270	133.00	30.22
IgalikuACOI 46	31-31.5	1200- 1205	56010 50282	35973	1912	295	1067	2204	117.28	24.00
IgalikuACOI 48	32-32.5	1190- 1195	67578 67090	35973	1754	286	937	2204	65.44	11.40
IgalikuACOI 50	33-33.5	1185- 1190	78612 86017	35973	2004	320	1121	2615	106.13	23.21
IgalikuACOI 52	34-34.5	1170- 1180	58366 83506	35973	2163	316	1201	2552	156.89	20.59
IgalikuACOI 54	35-35.5	1155- 1165	64819 49786	35973	1875	227	1058	2195	98.50	18.63
IgalikuACOI 56	36-36.5	1140- 1150	56322 51378	35973	2125	320	1203	2375	183.80	43.35
IgalikuACOI 58	37-37.5	1130- 1135	83734 59954	35973	1977	254	1231	2517	92.75	11.03
IgalikuACOI 60	38-38.5	1110- 1120	75710 53776	35973	1995	272	1213	2424	143.95	17.98
IgalikuACOI 62	39-39.5	1080- 1095	48225 56848	35973	1721	204	1070	1964	142.50	32.27
IgalikuACOI 64	40-40.5	1045- 1055	63083 61289	35973	1561	197	955	1866	102.63	27.92
IgalikuACOI 66	41-41.5	1030- 1035	55355 71825	35973	1730	226	1042	2046	159.79	40.72
IgalikuACOI 68	42-42.5	1023- 1026	39037 78357	35973	1534	187	912	1812	120.94	26.40
IgalikuACOI 70	43-43.5	1018- 1021	53625 75381	35973	1618	195	945	1905	110.58	25.08
IgalikuACOI 72	44-44.5	1010- 1015	66907 96865	35973	1693	219	946	2144	93.69	17.19
IgalikuACOI 74	45-45.5	990- 1000	94816 52746	35973	1649	180	997	2057	102.46	28.42
IgalikuACOI 76	46-46.6	960-975	76025 74946	35973	1663	202	1037	2050	113.29	30.18
IgalikuACOI 78	47-47.5	925-940	68792 54198	35973	1619	222	940	1879	98.59	18.00
IgalikuACOI 80	48-48.5	900-910	65799 47185	35973	1565	193	906	1813	104.27	22.64
IgalikuACOI 82	49-49.5	890-895	51825 77678	35973	1407	158	828	1736	71.85	18.32
IgalikuACOI 84	50-50.5	875-880	60876 69023	35973	1652	240	881	1962	117.13	31.46
IgalikuACOI 92	57.5-58.5	685-700	59119 51157	35973	1277	158	707	1297	142.85	35.21
IgalikuACOI 100	65.5-66.5	360-370	73358 76605	35973	1317	247	706	1378	82.02	13.37
IgalikuACOI 108	73.5-74.5	80-95	48005 111961	35973	1208	198	654	1208	224.96	73.68
IgalikuACOI 116	81.5-82.5	-50/-40	58389 51221	35973	1346	267	685	1357	202.28	54.98
IgalikuACOI 124	89.5-90.5	-195/- 180	85378 75823	35973	1789	342	963	1996	305.45	77.71



Table IV.S3. For Lake Bourget (A) and Lake Igaliku (B): results obtained from SIMPER analysis considering the periods defined by ANOSIM analysis. The table presents the percentage of contribution to temporal rearrangements for the OTUs that contributed to more than 0.2 % of the total dissimilarity, the ecological inference of these OTUs, the detection of their potential oscillating pattern, and their taxonomic resolution of each OTU with the percentage of identity to the best hit. The sign (+) represents an increase and the sign (-) represents a decrease of the relative abundance of the OTU between the 2 compared periods.

(A) Lake Bourget

OTUs	% of contribution to the rearrangements		Oscillating pattern	Ecological inference	Taxonomic group	Best Taxonomic resolution	% identity
	before 1940s vs	1940-60s vs after					
OTU3409	2.63 (-)	7.93 (+)		Npi	Ciliophora	<i>Epispathidium</i> sp.	89.9
OTU52956	0.33 (+)	3.67 (+)		Unk	unclassified_Stramenopiles	unclassified_Stramenopiles	91.3
OTU42983	0.24 (+)	2.58 (+)		Pi	Chlorophyta	<i>Chlamydomonas</i> sp.	98.5
OTU39169	2.43 (+)	1.87 (-)		Npi	Fungi	Cryptomycota_CM1	92.7
OTU5230	1.38 (+)	1.56 (-)		Unk	unclassified_Alveolata	unclassified_Rhizaria	79.7
OTU58202	1.51 (-)	1.30 (-)		Unk	unclassified_Alveolata	unclassified_Alveolata	93.4
OTU74525	1.19 (+)	1.18 (-)		Npi	Fungi	Cryptomycota	92.4
OTU65060	< 0.20	1.13 (+)	yes	Npi	Fungi	unclassified_Chytridiomycetes	97.0
OTU35618	1.16 (+)	1.08 (+)	yes	Npi	Apicomplexa	Apicomplexa	97.1
OTU38055	1.26 (-)	1.03 (+)		Npi	Ciliophora	Haptorida	91.9
OTU15701	1.73 (-)	0.93 (+)	yes	Npi	Bicosoecida	Bicosoecida_clade	79.6
OTU23762	< 0.20	0.90 (+)	yes	Pi	Chrysophyceae	Chrysosaccus sp.	90.3
OTU4897	0.85 (+)	0.81 (-)	yes	Npi	Fungi	unclassified_Cryptomycota	94.3
OTU12810	0.79 (-)	0.80 (-)	yes	Npi	Ciliophora	Alveolate_1	94.3
OTU9807	< 0.20	0.77 (+)	yes	Npi	Fungi	unclassified_Fungi	95.5
OTU72919	1.21 (-)	0.68 (+)		Npi	Ciliophora	Arcuospathidium sp.	97.0
OTU37932	1.11 (-)	0.66 (-)		Npi	Ciliophora	Spathidium sp.	95.1
OTU12403	0.56 (+)	0.60 (-)	yes	Npi	Ciliophora	Alveolate_1	96.6
OTU8809	0.64 (+)	0.60 (-)		Npi	Perkinsea	Perkinsea_1	91.7
OTU58675	< 0.20	0.60 (+)	yes	Npi	Fungi	unclassified_Fungi	92.9
OTU14646	0.61 (+)	0.60 (-)		Pi	Chlorophyta	Trebouxiophyceae	98.5
OTU71469	0.35 (+)	0.59 (-)	yes	Npi	Fungi	Penicillium sp.	94.0
OTU62642	< 0.20	0.58 (+)	yes	Unk	unclassified_Stramenopiles	unclassified_Stramenopiles	94.0
OTU62261	0.53 (+)	0.58 (-)	yes	Npi	Ciliophora	Alveolate_1	95.4
OTU45249	0.52 (+)	0.55 (-)	yes	Npi	Ciliophora	Alveolate_1	93.9
OTU59827	0.51 (+)	0.51 (-)	yes	Pi	Chlorophyta	Lagerheimia sp.	98.5
OTU76908	0.57 (+)	0.51 (-)		Npi	Apicomplexa	Apicomplexa	97.4
OTU19500	0.46 (+)	0.51 (-)	yes	Npi	Ciliophora	Alveolate_1	93.5
OTU13101	< 0.20	0.50 (+)	yes	Npi	Cercozoa	Cercozoa	87.5
OTU57038	3.78 (-)	0.49 (-)		Unk	ES Alveolata	ES Alveolata	96.3
OTU6992	2.66 (-)	0.49 (-)		Unk	unclassified_Alveolata	unclassified_Alveolata	95.5
OTU3855	0.47 (+)	0.48 (-)	yes	Unk	unclassified_Alveolata	unclassified_Alveolata	75.2
OTU64308	< 0.20	0.48 (+)	yes	Pi	Bacillariophyta	unclassified_Thalassiosiraceae	97.4
OTU579	0.26 (+)	0.46 (+)	yes	Unk	Rhizaria	unclassified_Rhizaria	90.8
OTU39097	0.30 (-)	0.43 (+)	yes	Npi	Fungi	unclassified_Pleosporales	97.3
OTU9270	0.23 (+)	0.40 (+)	yes	Pi	Chlorophyta	unclassified_Chlamydomonadaceae	95.1
OTU62873	1.96 (-)	0.39 (-)	yes	Npi	Bicosoecida	Bicosoecida_clade	79.9
OTU8852	3.10 (-)	0.39 (-)	yes	Npi	Fungi	Chytridiomycetes	76.6
OTU57850	0.36 (+)	0.39 (-)	yes	Npi	Ciliophora	Alveolate_1	94.6
OTU39050	< 0.20	0.38 (+)	yes	Npi	Fungi	unclassified_Fungi	88.8
OTU32080	0.08	0.38 (+)	yes	Pi	Bacillariophyta	unclassified_Fragilariaceae	97.8
OTU49494	< 0.20	0.35 (+)	yes	Unk	unclassified_Alveolata	unclassified_Alveolata	97.0
OTU43283	0.68 (-)	0.35 (-)		Npi	Ciliophora	unclassified_Polycostidae	95.4
OTU76634	0.29 (+)	0.35 (-)	yes	Npi	Fungi	unclassified_Blastocladales	94.0
OTU72274	0.21 (+)	0.34 (+)	yes	Pi	Chlorophyta	unclassified_Chlamydomonadaceae	95.5
OTU920	1.03 (-)	0.33 (-)	yes	Npi	Cercozoa	Plasmiodiophora sp.	80.7
OTU17747	< 0.20	0.33 (+)	yes	Npi	Fungi	unclassified_Chytridiomycota	83.5
OTU16279	< 0.20	0.32 (+)	yes	Pi	Chlorophyta	Mychonastes sp.	97.0
OTU721	0.32 (+)	0.31 (-)		Npi	Fungi	Cryptomycota_CM1	96.6
OTU44626	0.30 (+)	0.30 (-)		Npi	Fungi	unclassified_Fungi	87.2
OTU14098	0.29 (+)	0.29 (-)	yes	Npi	Fungi	Cryptomycota_CM1	91.6
OTU34798	< 0.20	0.29 (+)	yes	Pi	Bacillariophyta	unclassified_Thalassiosiraceae	95.1
OTU29821	0.26 (+)	0.28 (-)	yes	Npi	Apicomplexa	Apicomplexa	94.9
OTU27689	< 0.20	0.26 (+)	yes	Pi	Chlorophyta	unclassified_Chlamydomonadaceae	95.8
OTU74712	0.28 (+)	0.26 (-)		Npi	Fungi	unclassified_Cryptomycota	96.2
OTU59774	< 0.20	0.26 (+)	yes	Pi	Dinophyceae	unclassified_Dinophyceae	94.1
OTU53281	< 0.20	0.26 (+)	yes	Npi	Ciliophora	unclassified_Ciliophora	96.6
OTU52551	< 0.20	0.25 (+)	yes	Npi	Fungi	unclassified_Chytridiomycota	90.1
OTU18150	0.27 (+)	0.25 (-)	yes	Unk	unclassified_Alveolata	unclassified_Alveolata	95.8
OTU78496	< 0.20	0.25 (+)	yes	Pi	Chlorophyta	Wislouchiella sp.	98.1
OTU40951	< 0.20	0.24 (+)	yes	Unk	unclassified_Alveolata	unclassified_Alveolata	98.5

OTU69537	0.32 (+)	0.24 (-)		Pi	Chlorophyta	Desmodesmus sp.	93.0
OTU74400	< 0.20	0.24 (+)	yes	Npi	Ciliophora	<i>Trachelophyllum</i> sp.	95.9
OTU65130	0.23 (+)	0.23 (-)		Pi	Chlorophyta	unclassified_Nephroselmidophyceae	91.8
OTU28684	0.21 (-)	0.23 (-)	yes	Npi	Fungi	unclassified_Cordycipitaceae	97.7
OTU44653	< 0.20	0.23 (+)	yes	Pi	Chlorophyta	unclassified_Scenedesmaeae	97.4
OTU10926	< 0.20	0.22 (+)		Npi	Oomycetes	unclassified_Oomycetes	90.2
OTU9027	< 0.20	0.22 (-)	yes	Npi	Cercozoa	Plasmodiophora sp.	82.0
OTU40933	0.25 (-)	0.22 (-)	yes	Npi	Fungi	unclassified_Dothideomycetes	90.2
OTU45224	0.29 (+)	0.22 (-)		Pi	Chlorophyta	unclassified_Scenedesmaeae	97.0
OTU65365	0.21 (+)	0.21 (-)		Npi	Fungi	Cryptomycota	92.5
OTU11888	< 0.20	0.21 (+)	yes	Unk	unclassified_Alveolata	unclassified_Alveolata	97.0
OTU50995	< 0.20	0.21 (+)	yes	Npi	Cercozoa	Cercozoa_1	98.1
OTU3092	0.23 (+)	0.20 (-)		Pi	Dinophyceae	<i>Symbiodinium</i> sp.	94.7
OTU62983	< 0.20	0.20 (+)	yes	Npi	Fungi	unclassified_Fungi	90.2
OTU58568	0.25 (+)	0.20 (-)	yes	Npi	Cercozoa	Gymnophrys sp.	80.4
OTU41369	< 0.20	0.20 (+)	yes	Pi	Eustigmatophyceae	Nannochloropsis sp.	97.7
OTU79385	< 0.20	0.20 (+)	yes	Npi	Fungi	unclassified_Fungi	89.2
OTU16608	< 0.20	0.20 (+)	yes	Npi	Fungi	unclassified_Chytridiomycota	93.2
OTU69102	< 0.20	0.20 (+)	yes	Npi	Ciliophora	Semispithidium sp.	93.3
OTU47888	0.90 (-)	< 0.20		Unk	Dinophyceae	unclassified_Dinophyceae	94.4
OTU56519	0.27 (-)	< 0.20	yes	Npi	Cercozoa	Cercozoa_2	93.2
OTU11783	0.38 (-)	< 0.20	yes	Npi	Perkinsea	Perkinsea_2	84.8
OTU16518	0.55 (-)	< 0.20		Pi	Dinophyceae	Symbiodinium sp.	92.3
OTU21266	0.21 (-)	< 0.20	yes	Npi	Fungi	unclassified_Pleosporales	97.3
OTU51341	0.87 (-)	< 0.20	yes	Pi	Chlorophyta	Pseudendoclonium sp.	62.5
OTU21720	0.51 (-)	< 0.20	yes	Npi	Fungi	Metarhizium sp.	97.3
OTU68720	0.30 (-)	< 0.20	yes	Unk	unclassified_Alveolata	unclassified_Alveolata	90.7
OTU38660	0.20 (-)	< 0.20	yes	Npi	Bicosoecida	Bicosoecida_clade	78.8
OTU51404	0.30 (-)	< 0.20	yes	Npi	Cercozoa	Plasmodiophora	79.4
OTU35209	1.02 (-)	< 0.20	yes	Npi	Fungi	Polyporales	98.5
OTU18217	3.31 (-)	< 0.20	yes	Pi	Haptophyceae	unclassified_Haptophyceae	59.3
OTU9200	0.26 (-)	< 0.20	yes	Npi	Fungi	Metarhizium sp.	95.8
OTU3689	0.79 (-)	< 0.20	yes	Npi	Ciliophora	unclassified_Colpodiidae	63.4
OTU21649	0.72 (-)	< 0.20	yes	Pi	Phaeocystales	Phaeocystis sp.	63.0
OTU77985	0.23 (-)	< 0.20	yes	Npi	Fungi	unclassified_Chytridiomycota	79.9
OTU22472	0.40 (-)	< 0.20	yes	Pi	Haptophyceae	unclassified_Haptophyceae	61.9

## (B) Lake Igaliku

OTUs	ù of contribution to the rearrangements			Oscillating OTUs	Ecological inference	Taxonomic group	Best Taxonomic resolution	% identity
	before 700 AD	700-1350 AD vs	1450-1880					
OTU52956	< 0.20	0.50 (+)	3.64 (+)	yes	Unk	unclassified_Stramenopiles	unclassified_Stramenopiles	91.3
OTU41720	2.31 (+)	9.73 (-)	10.53 (-)	yes	Npi	Cercozoa	Cercozoa_1	91.3
OTU10081	2.88 (+)	2.68 (-)	1.89 (-)	yes	Unk	unclassified_Rhizaria	unclassified_Rhizaria	95.2
OTU8790	< 0.20	0.20 (+)	1.48 (+)	yes	Npi	Schizopyrenida	<i>Naegleria</i> sp.	89.2
OTU469	1.62 (+)	2.01 (-)	1.28 (-)	yes	Pi	unclassified_Rhizaria	unclassified_Rhizaria	81.8
OTU69865	< 0.20	1.30 (+)	1.23 (-)	yes	Npi	Cercozoa	Cercozoa_1	85.1
OTU5758	< 0.20	< 0.20	1.21 (+)	yes	Npi	Schizopyrenida	<i>Naegleria</i> sp.	91.1
OTU42178	< 0.20	1.4 (+)	1.19 (-)	yes	Npi	Ciliophora	<i>Dexiostoma</i> sp.	98.1
OTU44926	< 0.20	0.20 (+)	1.13 (+)	yes	Npi	Cercozoa	Cercozoa_1	84.1
OTU57824	1.35 (-)	0.22 (-)	1.12 (+)	yes	Unk	unclassified_Stramenopiles	unclassified_Stramenopiles	98.1
OTU25969	1.62 (-)	0.29 (-)	1.11 (+)	yes	Unk	unclassified_Stramenopiles	unclassified_Stramenopiles	95.9
OTU28768	< 0.20	0.36 (+)	0.980 (+)	yes	Npi	Fungi	unclassified_Fungi	94.7
OTU31016	0.22 (+)	0.75 (+)	0.85 (+)	yes	Unk	Chrysophyceae	unclassified_Chrysophyceae	90.0
OTU67130	< 0.20	< 0.20	0.83 (+)	yes	Unk	unclassified_Rhizaria	unclassified_Rhizaria	93.3
OTU15373	< 0.20	0.83 (+)	0.79 (-)	yes	Npi	Perkinsea	Perkinsea_1	95.9
OTU62744	1.82 (+)	1.94 (-)	0.77 (-)	yes	Npi	Cercozoa	Cercozoa_1	91.1
OTU9270	< 0.20	< 0.20	0.75 (+)	yes	Pi	Chlorophyta	unclassified_Chlamydomonadaceae	95.1
OTU74258	0.48 (+)	0.75 (+)	0.73 (-)	yes	Npi	Ciliophora	<i>Paradiaphrys</i> sp.	67.0
OTU77679	5.43 (+)	5.80 (-)	0.71 (-)	yes	Npi	Cercozoa	Cercozoa_1	93.3
OTU40537	0.76 (+)	1.02 (-)	0.71 (-)	yes	Npi	Schizopyrenida	<i>Naegleria</i> sp.	90.1
OTU69058	< 0.20	< 0.20	0.68 (+)	yes	Unk	unclassified_Rhizaria	unclassified_Rhizaria	93.3
OTU44063	0.73 (-)	< 0.20	0.67 (+)	yes	Unk	unclassified_Stramenopiles	unclassified_Stramenopiles	96.7
OTU20865	< 0.20	0.56 (+)	0.66 (+)	yes	Unk	Chrysophyceae	Chrysophyceae	96.3
OTU72274	< 0.20	< 0.20	0.65 (+)	yes	Pi	Chlorophyta	unclassified_Chlamydomonadaceae	95.5
OTU14600	1.68 (-)	1.09 (-)	0.62 (-)	yes	Npi	Ciliophora	<i>Cyclidium</i> sp.	70.7
OTU38055	0.23 (-)	< 0.20	0.60 (+)	yes	Npi	Ciliophora	unclassified_Haptorida	91.9
OTU49044	< 0.20	< 0.20	0.59 (+)	yes	Unk	unclassified_Rhizaria	unclassified_Rhizaria	97.8
OTU58202	0.24 (-)	0.54 (+)	0.57 (-)	yes	Unk	unclassified_Alveolata	unclassified_Alveolata	93.4
OTU76121	< 0.20	0.53 (+)	0.55 (-)		Npi	Cercozoa	Cercozoa_1	95.1
OTU8239	< 0.20	0.53 (+)	0.54 (-)		Npi	Ciliophora	Spathidium sp.	95.1
OTU72919	< 0.20	0.39 (+)	0.53 (+)	yes	Npi	Ciliophora	Arcuospathidium sp.	97.0
OTU53307	1.06 (+)	1.07 (-)	0.53 (-)	yes	Npi	Bicosoecida	Bicosoecida_clade	93.9
OTU57300	< 0.20	0.44 (+)	0.53 (-)	yes	Npi	Fungi	unclassified_Chytridiomycota	97.4
OTU37932	< 0.20	< 0.20	0.52 (+)	yes	Npi	Ciliophora	Spathidium sp.	95.1
OTU65258	0.24 (-)	< 0.20	0.50 (+)	yes	Npi	Fungi	unclassified_Fungi	95.9
OTU14646	0.92 (+)	0.90 (-)	0.46 (-)	yes	Pi	Chlorophyta	unclassified_Trebouxiophyceae	98.5
OTU44090	< 0.20	0.34 (+)	0.45 (+)	yes	Npi	Fungi	unclassified_Chytridiomycota	96.2
OTU27689	< 0.20	< 0.20	0.45 (+)	yes	Pi	Chlorophyta	unclassified_Chlamydomonadaceae	95.8
OTU78496	< 0.20	< 0.20	0.45 (+)	yes	Pi	Chlorophyta	Wislouchiella sp.	98.1

OTU30205	< 0.20	0.47 (+)	0.45 (-)		Npi	Perkinsea	Perkinsea_1	97.4
OTU47951	0.94 (+)	0.93 (-)	0.43 (-)	yes	Npi	Bicosoecida	<i>Bicosoecida_clade</i>	93.5
OTU73785	2.45 (+)	2.14 (-)	0.41 (-)	yes	Npi	Ciliophora	<i>Dexiostoma sp.</i>	89.3
OTU50060	3.40 (+)	3.59 (-)	0.39 (-)	yes	Npi	Cercozoa	Cercozoa_1	93.7
OTU4556	< 0.20	0.21 (+)	0.38 (+)	yes	Npi	Schizopyrenida	<i>Naegleria sp.</i>	83.6
OTU23374	< 0.20	< 0.20	0.37 (+)	yes	Pi	Chlorophyta	unclassified_Chlorophyta	92.9
OTU52816	< 0.20	< 0.20	0.37 (+)	yes	Npi	Ciliophora	Alveolate_1	91.3
OTU50100	< 0.20	< 0.20	0.36 (+)	yes	Npi	Ciliophora	unclassified_Stichotrichia	97.8
OTU21170	< 0.20	< 0.20	0.35 (+)	yes	Pi	Chlorophyta	unclassified_Chlorophyta	95.9
OTU49516	< 0.20	< 0.20	0.35 (+)	yes	Npi	Fungi	<i>Pilaira sp.</i>	69.2
OTU28155	< 0.20	0.34 (+)	0.34 (-)		Npi	Cercozoa	Cercozoa_1	96.2
OTU78367	0.24 (+)	0.35 (+)	0.32 (-)		Npi	Fungi	<i>Paraglomus sp.</i>	68.1
OTU70723	< 0.20	< 0.20	0.31 (+)	yes	Unk	unclassified_Stramenopiles	unclassified_Stramenopiles	95.9
OTU57856	0.70 (+)	0.70 (-)	0.31 (-)	yes	Npi	Bicosoecida	<i>Bicosoecida_clade</i>	90.7
OTU6992	0.52 (-)	< 0.20	0.30 (+)	yes	Unk	unclassified_Alveolata	unclassified_Alveolata	95.5
OTU37640	< 0.20	< 0.20	0.30 (+)	yes	Npi	Fungi	unclassified_Fungi	96.2
OTU51285	< 0.20	0.29 (+)	0.29 (-)		Unk	unclassified_Rhizaria	unclassified_Rhizaria	92.2
OTU57038	0.76 (-)	< 0.20	0.27 (+)	yes	Unk	unclassified_Alveolata	unclassified_Alveolata	96.3
OTU1825	< 0.20	0.28 (+)	0.27 (-)		Pi	Synurophyceae	<i>Mallomonas sp.</i>	64.4
OTU4212	< 0.20	< 0.20	0.26 (+)	yes	Unk	unclassified_Rhizaria	unclassified_Rhizaria	91.8
OTU51304	0.27 (+)	0.36 (-)	0.26 (-)	yes	Npi	Schizopyrenida	unclassified_Naegleria	93.6
OTU18335	0.58 (-)	0.21 (+)	0.25 (+)	yes	Unk	Chrysophyceae	unclassified_Chromulinaceae	98.5
OTU16011	< 0.20	0.26 (+)	0.25 (-)		Npi	Perkinsea	Perkinsea_1	97.7
OTU59774	< 0.20	< 0.20	0.24 (+)		Unk	Dinophyceae	unclassified_Dinophyceae	94.1
OTU29548	< 0.20	0.24 (+)	0.24 (-)		Npi	Fungi	<i>Sphenospora sp.</i>	64.8
OTU76103	< 0.20	0.24 (+)	0.24 (-)		Npi	Cercozoa	Cercozoa_1	93.5
OTU14786	< 0.20	< 0.20	0.23 (+)	yes	Pi	Synurophyceae	<i>Chrysophaerella sp.</i>	95.1
OTU501	< 0.20	0.23 (+)	0.23 (-)	yes	Npi	Cercozoa	Cercozoa_1	93.1
OTU44342	< 0.20	< 0.20	0.22 (+)	yes	Npi	Ciliophora	Alveolate_1	93.9
OTU48827	< 0.20	< 0.20	0.22 (+)	yes	Pi	Chlorophyta	Chlamydomonadaceae	98.9
OTU7733	< 0.20	0.23 (+)	0.22 (-)		Npi	Fungi	<i>Pluteus sp.</i>	67.6
OTU48783	< 0.20	< 0.20	0.22 (-)	yes	Pi	Bacillariophyta	<i>Aulacoseira sp.</i>	97.0
OTU3409	0.45 (-)	< 0.20	0.21 (+)	yes	Npi	Fungi	<i>Candida sp.</i>	90.2
OTU12810	< 0.20	< 0.20	0.20 (+)	yes	Npi	Ciliophora	Alveolate_1	94.3
OTU65886	< 0.20	< 0.20	0.20 (+)		Npi	Apicomplexa	Sarcocystidae	90.9
OTU10338	1.87 (+)	1.98 (-)	< 0.20	yes	Npi	Cercozoa	Cercozoa_1	93.3
OTU30069	4.97 (-)	1.78 (-)	< 0.20	yes	Npi	Ciliophora	<i>Cyclidium sp.</i>	91.5
OTU33699	1.18 (+)	1.35 (-)	< 0.20	yes	Unk	unclassified_Rhizaria	unclassified_Rhizaria	94.9
OTU49566	1.20 (+)	1.27 (-)	< 0.20		Unk	unclassified_Rhizaria	unclassified_Rhizaria	92.9
OTU8603	0.87 (+)	0.93 (-)	< 0.20		Npi	Perkinsea	Perkinsea_1	89.3
OTU51842	0.78 (+)	0.87 (-)	< 0.20	yes	Unk	unclassified_Rhizaria	unclassified_Rhizaria	89.2
OTU42111	0.69 (+)	0.80 (-)	< 0.20	yes	Pi	unclassified_Haptophyceae	unclassified_Haptophyceae	91.6
OTU37586	0.44 (+)	0.56 (-)	< 0.20	yes	Unk	unclassified_Rhizaria	unclassified_Rhizaria	91.2
OTU20991	0.46 (+)	0.53 (-)	< 0.20	yes	Npi	Fungi	<i>Hygrocybe sp.</i>	63.1
OTU16885	0.40 (+)	0.46 (-)	< 0.20		Unk	unclassified_Rhizaria	unclassified_Rhizaria	90.2
OTU74554	0.97 (-)	0.36 (-)	< 0.20	yes	Npi	Ciliophora	<i>Cyclidium sp.</i>	93.9
OTU14892	0.28 (+)	0.32 (-)	< 0.20		Npi	Fungi	<i>Polychytrium sp.</i>	88.6
OTU50456	2.38 (-)	0.29 (-)	< 0.20	yes	Unk	Dinophyceae	<i>Kofoidinium sp.</i>	89.9
OTU35799	0.24 (+)	0.27 (-)	< 0.20	yes	Npi	Fungi	<i>Hygrocybe sp.</i>	63.8
OTU66436	0.20 (+)	0.26 (-)	< 0.20		Npi	Ciliophora	unclassified_Stichotrichia	97.4
OTU28653	< 0.20	0.22 (-)	< 0.20		Unk	unclassified_Rhizaria	unclassified_Rhizaria	88.7
OTU56375	0.64 (-)	0.21 (+)	< 0.20	yes	Npi	Cercozoa	Cercozoa_1	93.3
OTU70059	< 0.20	0.2 (-)	< 0.20		Npi	Cercozoa	Cercozoa_1	93.7
OTU77864	1.20 (-)	< 0.20	< 0.20	yes	Unk	Dinophyceae	<i>Kofoidinium sp.</i>	88.3
OTU13159	0.95 (-)	< 0.20	< 0.20		Pi	Chrysophyceae	<i>Dinobryon sp.</i>	97.4
OTU1716	0.88 (-)	< 0.20	< 0.20	yes	Npi	Kinetoplastida	<i>Trypanosoma sp.</i>	80.3
OTU19424	0.83 (-)	< 0.20	< 0.20	yes	Pi	Chlorophyta	unclassified_Chlorophyta	95.1
OTU49699	0.78 (-)	< 0.20	< 0.20		Pi	Synurophyceae	<i>Mallomonas sp.</i>	97.7
OTU59454	0.68 (-)	< 0.20	< 0.20	yes	Pi	Bangiophyceae	<i>Porphyra sp.</i>	51.6
OTU20579	0.57 (-)	< 0.20	< 0.20	yes	Pi	Chlorophyta	unclassified_Chlorophyta	95.6
OTU8852	0.46 (-)	< 0.20	< 0.20	yes	Npi	Fungi	unclassified_Chytridiomycota	76.6
OTU26778	0.42 (-)	< 0.20	< 0.20	yes	Npi	Cercozoa	<i>Massisteria sp.</i>	85.4
OTU4623	0.40 (-)	< 0.20	< 0.20	yes	Npi	Kinetoplastida	<i>Bodo sp.</i>	91.3
OTU15701	0.36 (-)	< 0.20	< 0.20	yes	Npi	Bicosoecida	<i>Bicosoecida_clade</i>	79.6
OTU18217	0.36 (-)	< 0.20	< 0.20	yes	Pi	unclassified_Haptophyceae	unclassified_Haptophyceae	59.3
OTU9425	0.36 (-)	< 0.20	< 0.20		Npi	Apicomplexa	unclassified_Apicomplexa	98.9
OTU47050	0.33 (-)	< 0.20	< 0.20	yes	Npi	Cercozoa	<i>Massisteria sp.</i>	87.3
OTU33951	0.32 (-)	< 0.20	< 0.20	yes	Npi	Cercozoa	unclassified_Cercozoa	86.2
OTU62873	0.31 (-)	< 0.20	< 0.20	yes	Npi	Bicosoecida	<i>Bicosoecida_clade</i>	79.9
OTU31695	0.27 (-)	< 0.20	< 0.20	yes	Npi	Cercozoa	Cercozoa_1	61.9
OTU49198	0.26 (-)	< 0.20	< 0.20	yes	Unk	unclassified_Stramenopiles	unclassified_Stramenopiles	90.3
OTU55316	0.26 (-)	< 0.20	< 0.20	yes	Npi	Kinetoplastida	<i>Bodo sp.</i>	93.9
OTU53146	0.25 (-)	< 0.20	< 0.20	yes	Pi	Chlorophyta	unclassified_Chlamydomonadales	95.5
OTU6418	0.25 (-)	< 0.20	< 0.20	yes	Npi	Fungi	<i>Dekkera sp.</i>	91.0
OTU21309	0.24 (-)	< 0.20	< 0.20	yes	Npi	Cercozoa	unclassified_Cercozoa	88.1
OTU26093	0.24 (-)	< 0.20	< 0.20	yes	Npi	Fungi	<i>Dissophora sp.</i>	98.9
OTU4505	0.23 (-)	< 0.20	< 0.20	yes	Pi	Chlorophyta	<i>Scherffelia sp.</i>	95.8
OTU47814	0.22 (-)	< 0.20	< 0.20	yes	Npi	Ciliophora	<i>Condylostoma sp.</i>	61.3
OTU16518	0.21 (-)	< 0.20	< 0.20		Pi	Dinophyceae	<i>Symbiodinium sp.</i>	92.3

# **V. Impact des forçages environnementaux au cours des 100 dernières années sur les communautés microbiennes eucaryotes des lacs Annecy et Bourget (France)**

## **Tracking a century of changes in microbial eukaryotic diversity in lakes driven by nutrient enrichment and climate warming**

Papier en préparation pour soumission dans *Environmental Microbiology*

Eric Capo<sup>1</sup>, Didier Debroas<sup>2,3</sup>, Fabien Arnaud<sup>4</sup>, Marie-Elodie Perga<sup>1</sup>, Cécile Chardon<sup>1</sup>, Isabelle Domaizon<sup>1\*</sup>

<sup>1</sup>CARTELE, INRA, Université de Savoie Mont Blanc, 74200, Thonon-les-bains, France

<sup>2</sup>Université Clermont Auvergne, Université Blaise Pascal, Laboratoire « Microorganismes : Génome et Environnement », BP 10448, F-63000 Clermont-Ferrand, France

<sup>3</sup>CNRS, UMR 6023, LMGE, Campus Universitaire des Cézeaux 63171 Aubière, France

<sup>4</sup>CNRS, UMR 5204 EDYTEM, Université Savoie Mont Blanc, Le Bourget du Lac Cedex, France

### **Contribution des auteurs :**

- Conception de l'étude: ID, EC, DD
- Acquisition des données: EC, CC
- Traitements bioinformatiques: DD
- Analyses et interprétations des données: EC, ID, DD, FA, MEP, CC
- Ecriture du papier: EC, ID, DD, FA, MEP, CC

### **Resumé**

Afin d'évaluer les effets des forçages anthropiques locaux (ici modification des teneurs en phosphore) et climatiques (changements de température) sur les communautés de micro-eucaryotes lacustres au cours du siècle dernier, nous avons appliqué le séquençage massif à l'ADN sédimentaire de deux lacs (lac d'Annecy et lac du Bourget). Ces deux lacs ont été exposés au cours du siècle dernier au même forçage climatique mais à des niveaux d'eutrophisation différents (eutrophisation marquée dans le lac du Bourget et faible dans le lac d'Annecy). Les données acquises montrent que dans les deux lacs, plusieurs groupes

eucaryotes (Apicomplexés, Cercozoaires, Chrysophycées, Ciliés, Chlorophytes and Dinoflagellés) ont été impactés par l'enrichissement en phosphore, avec des effets positifs ou négatifs sur leur richesse et abondance relative (proportion dans les inventaires ADN). La modélisation statistique appliquée aux inventaires moléculaires montre que, avant les années 1980s, le phosphore est un facteur plus structurant que les fluctuations de température. Au-delà des années 80s, l'augmentation des températures est identifiée comme un prédicteur de la dynamique de plusieurs groupes eucaryotes. L'augmentation des températures concomitante avec la re-oligotrophisation des deux lacs est notamment marquée par l'émergence d'OTUs de Dinophycées (en termes de richesse et d'abondance relative). L'application des analyses de réseaux de co-occurrence à ces données temporelles a permis de montrer que la réponse des assemblages eucaryotes est également détectable au travers de la dynamique des modules (OTUs fortement connectées) constituant les réseaux écologiques.

## **V.1. Abstract**

High-throughput sequencing of sedimentary DNA (sed-DNA) was utilized to reconstruct the temporal dynamics of microbial eukaryote communities (MECs) at a centennial scale in two re-oligotrophicated lakes that were exposed to different levels of phosphorus enrichment. The temporal changes within the MECs were expressed in terms of richness, composition and community structure to investigate their relationships with two key forcing factors (i.e., nutrient enrichment and climate warming). Various groups, including Apicomplexa, Cercozoa, Chrysophyceae, Ciliophora, Chlorophyceae and Dinophyceae, responded to [P] enrichment levels with either positive or negative impacts on their richness and relative abundance. For both lakes, statistical modelling demonstrated that [P] was a dominant contributor to MECs modifications before the 1980s; after the mid-80s, the contribution of air temperature increased and potentially surpassed the contribution of [P]. Co-occurrence network analysis revealed that some clusters of taxa (i.e., modules) composed mainly of Dinophyceae and unclassified Alveolata were strongly correlated to air temperature in both lakes. Overall, our data showed that sed-DNA constitutes a precious archive of information on past biodiversity changes, allowing the study of the dynamics of numerous eukaryotic groups that were not traditionally considered in paleo-reconstructions.

## V.2. Introduction

Eutrophication and climate warming are both known to strongly alter the functioning of lacustrine systems by changing the diversity, structure and metabolism of biological communities at multiple trophic levels (Williamson *et al.* 2009, Adrian *et al.* 2009, Moss 2012, Kraemer *et al.* 2016). Because these two forcing factors fluctuate at relatively long-term time scales and display either antagonistic or synergetic impacts, their respective effects are still not known for most biological communities, particularly within microbial lineages (Falkowski & Oliver 2007, Caron & Hutchins, 2013). Recent research has confirmed the interactive impacts of warming and eutrophication on phytoplankton composition and productivity (e.g., Gallina *et al.* 2013, Rigosi *et al.* 2014); however, environmental studies have rarely investigated the diversity of microbial eukaryotes (here defined as Protists and Fungi), despite their abundance and key roles in biogeochemical processes and ecosystem functioning (as primary producers, bacterivores, planktivores, saprotrophs, and parasites) (Falkowski *et al.* 2008, Caron *et al.* 2009). The use of high-throughput sequencing (HTS) tools via the analysis of 18S rRNA genes retrieved directly from environmental samples now allows for deep assessments of the composition of the microbial eukaryotic communities (de Vargas *et al.* 2015, Filker *et al.* 2016, Giner *et al.* 2016). Though their responses to environmental changes have potentially large functional implications, significant knowledge gaps about how environmental conditions shape microbial eukaryotic communities (MECs) remain.

Long-term time-series are a prerequisite for understanding natural and human-induced modifications in biological communities. In such an attempt, paleolimnological approaches have already provided valuable insights into the drivers of biological assemblages over long time scales (e.g. Battarbee *et al.* 2001, Smol 2008, Perga *et al.* 2015). Morphological remains preserved in sediments are classically used as biomarkers to track changes of various biological groups (e.g., diatoms, cladocerans, chironomids; Smol *et al.* 2005, Millet *et al.* 2010, Alric *et al.* 2013); however, these classical proxies reflect only a small part of the lacustrine biodiversity and rarely include the entire diversity of MECs because many eukaryotic taxa do not leave diagnostic morphological features in sediments. The recent emergence of DNA tools in paleolimnological studies made it possible to track the temporal dynamics of MECs from archived DNA. Coolen *et al.* (2013) first illustrated how the application of HTS to sedimentary DNA (sed-DNA) revealed the responses of eukaryotic communities from the Black Sea to marked changes in climatic conditions over the last

10,000 years (i.e., the shift from freshwater to saline systems). This type of study that aimed at reconstructing the diversity of large biological assemblages is still rare and requires methodological validation. To our knowledge, only Capo *et al.* (2015) showed that sed-DNA was a good reflection of the microbial eukaryotic diversity found in the water column of the freshwater Lake Bourget. These authors confirmed the potential of Sed-DNA for reconstructing past MECs, including various pigmented and non-pigmented groups, such as Chlorophyta, Dinophyceae, Chrysophyceae, Ciliophora, Apicomplexa, Chytridiomycota, and Cryptomycota. Although not without caveats, especially due to the risk of the degradation of ancient DNA and the potential differential preservation of DNA according to taxa or lake conditions (Anderson-Carpenter *et al.* 2011, Boere *et al.* 2011, Torti *et al.* 2015, Kirkpatrick *et al.* 2016), the application of HTS to investigate the past diversity of MECs opens up many new areas of inquiry due to the unique temporal perspectives of investigations for fundamental and applied questions in microbial ecology. When considering the entire microbial eukaryotic community, the co-occurrence networks between interacting microorganisms (pigmented taxa, phagotrophs, parasites and saprotrophs) and the topological features of these networks can be explored. Such networks have already been built to unveil microbial interactions in a wide range of ecosystems, such as soils (Lentendu *et al.* 2014), human microbiota (Faust *et al.* 2012) and, more recently, oceans (Lima-Mendez *et al.* 2015). However, this has rarely been the case for freshwater microbial eukaryotes (Vick-Majors *et al.* 2014, Kammerlander *et al.* 2015) and has never been the case for paleolimnological approaches.

In this study, we applied metabarcoding to sed-DNA to reveal the composition of MECs over the last century. We compared the temporal dynamics of MECs from two deep monomictic lakes subjected to similar climatic fluctuations (recent warming impact; Perga *et al.* 2015) but contrasting eutrophication levels. Lake Annecy has been exposed to a weak increase in phosphorus concentrations over the last century (maximum concentrations of  $\sim 15 \mu\text{gP}\cdot\text{L}^{-1}$  in the 1960s), while Lake Bourget has experienced eutrophication from the 1940s to the 1990s (phosphorus concentration reaching up to  $120 \mu\text{gP}\cdot\text{L}^{-1}$ ). By applying HTS (MiSeq sequencing, Illumina technique) to sed-DNA, we obtained molecular inventories which allowed us to address 4 main questions: (i) What are the changes within MECs in terms of richness, composition and community structure? (ii) How well do phosphorus concentration and air temperature explain the temporal changes detected for these metrics? (iii) Which specific phylogenetic groups or OTUs contributed to the temporal rearrangements in MECs related to



environmental forcing? (iv) From the co-occurrence networks built from these long-term time series, can we detect responses to environmental forcings at the network level?

## V.3. Materials and Methods

### V.3.1. Study sites and forcing factors

Lakes Bourget and Annecy are two deep lakes with surface areas of 42 and 27 km<sup>2</sup> and maximum depths of 145 and 69 m, respectively. These lakes are warm monomictic lakes located at the northwest edge of the French Alps. These two lakes are currently subjected to monitoring surveys (long-term observatory of alpine lakes, INRA Thonon; <http://www6.inra.fr/soere-ola>) that were started in 2004 for Lake Bourget and in 1996 for Lake Annecy. Air temperature (airT; annual mean air temperature anomalies) was considered the main form of climate change in the areas of both lakes because no trend was observed in the last century for annual mean precipitation in the watersheds of both lakes, while a global increase in air temperature was observed over the last 20 years (Alric *et al.* 2013). Lakes Bourget and Annecy were subjected to phosphorus concentration ([P]) increases in the mid-20th century due to industrial effluents and domestic sewage (Giguet-Covex *et al.* 2010, Perga *et al.* 2015) with very different levels. The reconstructed *Daphnia*-inferred TP function indicated that, for Lake Annecy, the [P] values were less than 10 µg.L<sup>-1</sup> until the mid-60s and increased to 10-15 µg.L<sup>-1</sup> until the early 80s. In Lake Bourget, the [P] values were less than 10 µg.L<sup>-1</sup> before the 1940s, and progressively increased until the early 1980s when [P] values were greater than 100 µg.L<sup>-1</sup>. Since the 1980s, a progressive decrease in [P] resulted in [P] values in the late 2000s that were less than 10 and 20 µg.L<sup>-1</sup> for Lake Annecy and Bourget, respectively, and in agreement with the monitored annual water dynamics [TP] (Berthon *et al.* 2014).

The study sites were selected based on their contrasting ecological trajectories and because the environmental conditions of these lakes are well documented (including physical, chemical and biological surveys). Furthermore, both lakes are characterized by conditions that are favourable for DNA preservation in sediment; the deepest zone of each lake at the coring sites is characterized by bottom anoxia, low temperatures and the absence of bioturbation. These parameters have to be considered for paleogenetics applications because they allow for limited DNA degradation in sediment (diagenetic processes occurring during burying in sediment) (Corinaldesi *et al.* 2008, Boere *et al.* 2011, Torti *et al.* 2015). Moreover, the DNA-

based studies applied on sedimentary archives from both lakes highlighted that they were good candidates for molecular paleolimnological studies (based on cyanobacterial and microbial eukaryotic DNA) (Savichtcheva *et al.* 2011, 2015, Domaizon *et al.* 2013, Capo *et al.* 2015).

### **V.3.2. Coring, dating and sub-sampling**

Several sediment cores were taken in 2009 from the deepest areas of Lakes Annecy and Bourget using a UWITEC gravity corer (90 mm in diameter, UWITEC, Mondsee, Austria). Two sediment cores were used for dating, the analysis of major elements and total organic carbon (TOC) and the inference of [P] using a Daphnia transfer function. The cores LDA\_09\_P2\_03 (for Lake Annecy) and LDB\_09\_P2\_02 (for Lake Bourget) were used for molecular analysis. To limit extraneous contamination, only the centre of each core was used for molecular analysis. The sampling interval was chosen to fit a complete number of lamina triplets when laminae were visible. Twenty-four and thirty-five slices were sampled from the Annecy and Bourget sediment cores, respectively. The sediment core dating was performed using one of the reference cores from each lake using radiometric methods ( $^{210}\text{Pb}$ ,  $^{226}\text{Ra}$ ,  $^{137}\text{Cs}$  and  $^{241}\text{Am}$  activities) and varve-counting (after 1930s). The reference cores and the cores dedicated to DNA analyses were correlated using lithological tie points and lamina counting. The detailed protocol and results are described by Jenny *et al.* (2013). The age-uncertainty is less than 3 years and 5 years for varve counting and  $^{210}\text{Pb}$  dating, respectively (Jenny *et al.* 2014), and about +/- 100 years for the  $^{14}\text{C}$ -based age model (Arnaud *et al.* 2012).

### **V.3.3. Molecular analysis**

For each sediment layer, two DNA extractions were performed on 0.5 g of wet sediment using the NucleoSpin® soil kit, according to the manufacturer instructions (Macherey-Nagel, Düren, Germany). The V7 region of the 18S rRNA gene was PCR amplified on a 260-bp-long fragment, from c.a. 25 ng of environmental DNA extracted from each sample and using the general eukaryotic primers 960F (5'-GGCTTAATTTGACTCAACRCG-3') (Gast *et al.* 2004) and NSR1438 (5'-GGGCATCACAGACCTGTTAT-3') (Van de Peer *et al.* 2000). As reported by Capo *et al.* (in press), the choice of the barcode region was based on 1) the representation of the barcode within the reference database (coverage), 2) the length of the amplicon (moderate length required for paleogenetics) and 3) the quality of taxonomic identification with the bioinformatics tools used. The combinations of tagged primers were designed to balance primer usage frequencies and therefore to reduce the frequency of

mistagging events that result in cross-contaminations (Esling *et al.* 2015). Each PCR was performed in duplicate in a total volume of 25 µl containing 3 µl of 10X NH<sub>4</sub> reaction buffer, 1.2 µl of 50 mM MgCl<sub>2</sub>, 0.25 µl of BioTaq (Bioline), 0.24 µl of 10 mM dNTP, 0.36 µl of 0.5 µg/µl BSA and 1 µl of each primer. The amplification conditions included initial denaturation at 94 °C for 10 min followed by 35 cycles of 1 min at 94 °C, 1 min at 55 °C and 1 min 30 s at 72 °C. The amplicons were then subjected to a final 10-min extension at 72 °C. The two PCR products obtained for a given DNA extraction were pooled and purified using Illustra™ GFX™ PCR DNA and Gel Band Purification Kit (GE Healthcare Life Sciences, Velizy Villacoublay, France), and quantified using a Quant-it PicoGreen kit (Invitrogen, Carlsbad, CA, USA) on a Fluoroskan Ascent™ FL (ThermoLabsystems, Beverly, MA, USA). Finally, the purified amplicons for each sample were pooled at equimolar concentrations and sent to FASTERIS SA (Geneva, Switzerland) for library preparation and paired-end (2 x 250 bp) sequencing on a MiSeq Illumina instrument (San Diego, CA, USA).

### **V.3.4. Bioinformatics processing applied for massive sequencing data**

The paired-end reads were merged together using UPARSE tools (option -fastq\_mergepairs with a minimal overlap equal to 150 and no mismatch allowed) (Edgar 2013), allowing us to obtain 17 116 627 raw DNA sequences. These sequences were then subjected to three cleaning procedures: (i) no undefined bases (Ns), (ii) a minimum sequence length of 200 bp, and (iii) no sequencing error in the forward and reverse primers. Putative chimaeras were detected by UCHIME (Edgar *et al.* 2011). The combination of these cleaning procedures with a sequencing depth (2 x 250 bp) that covered the entire amplicon length (~ 220 bp without the primers) allowed us to sequence each base twice to drastically minimize sequencing errors. After this cleaning step and demultiplexing process, the remaining sequences were clustered at a 95 % similarity threshold with UPARSE 7.0 (Edgar 2013) to obtain the seed OTUs (option -cluster\_fast). The delineation of the similarity threshold was conducted following the recommendations from Mangot *et al.* (2013), who estimated errors during PCR and sequencing by introducing an internal 18S standard (a unique alien DNA sequence), and demonstrated that 95 % was a relevant cut-off threshold for OTUs delineation, limiting artificial inflation of the OTU numbers (Mangot *et al.* 2013). The taxonomic affiliation was performed using BLAST against the SSURef SILVA database (Pruesse *et al.* 2007) after applying the following selection criteria: length > 1200 bp, quality score > 75 % and a pintail value > 50. The SILVA database was enriched with lacustrine DNA sequences originating

from various studies (Lefranc *et al.* 2005, Lepère *et al.* 2006, Mangot *et al.* 2013). The taxonomy of an OTU corresponded to the best hit in the similarity search. If an OTU was associated with several best hits (hits having the same identity), then the taxonomy was chosen as the common taxonomy of the hits and also based on the lowest common ancestor. In addition, a specific taxonomy based on previous publications (Lefranc *et al.* 2005, Lepère *et al.* 2006, Mangot *et al.* 2013) allowed us to identify typical freshwater clades (e.g., Cryptophyta\_1, 2, Perkinsea\_1). OTUs affiliated with metazoans and embryophyte taxa were removed from these analyses to consider only unicellular eukaryotes (microbes at the basis of trophic networks). After these different steps (cleaning, clustering and taxonomic affiliation), 5,537,953 quality-filtered DNA sequences were clustered in 91,692 OTUs classified as microbial eukaryotes. For each sample, OTUs detected in only one of the two duplicates were removed from the analysis, leading to 8,545 remaining OTUS. This additional cleaning step allowed us to minimize the presence of falsely rare OTUs due to the PCR bias recently pointed out by Lange *et al.* (2015) and Egge *et al.* (2015). For each remaining OTU, the number of DNA sequences obtained for each duplicate was calculated. To perform comparisons among samples, the number of DNA sequences was standardized at 29,813 for all samples, resulting in a total number of 1,758,967 clean DNA sequences (10.3 % and 31.7 % of the raw and quality-filtered sequencing dataset, respectively). From these data, the final dataset (i.e., rarefied OTU table with the 8,081 OTUs expressed in terms of relative number of DNA sequences for each sediment layer) was used for all further analyses.

*Data Accessibility:* Obtained data were deposited in the Dryad Digital Repository (<http://datadryad.org/>); demultiplexed reads are complemented by a table with the tag-to-samples information.

### **V.3.5. Specific procedures for paleogenetics work**

Strict clean-lab procedures were applied to ensure the authenticity of the results of the molecular inventories from sedimentary archives. For instance, the isolation of DNA from sediment was carried out in a specific DNA laboratory, physically separated from the molecular biology laboratory and employed rigorous controls against contamination. This included separate stations for sampling, DNA extractions and PCR amplifications, the use of sterile disposable materials (labware, gloves...) and the introduction of blanks (negative controls) for extraction and amplification steps, and duplicates for DNA extractions of each sediment sample. All blanks were found negative in the present study.

Stringent bioinformatics treatments were also applied: reads (i.e., DNA sequences) with undefined bases were removed, and only reads with full overlaps between the forward and reverse sequences were kept; and reads with 95 % sequence similarity were clustered into OTUs, and only OTUs retrieved in both replicated molecular inventories were kept for further analyses. These choices were aimed at reducing the potential pitfalls that can lead to an artificial increase in the number of rare OTUs in HTS inventories, and therefore, the reliability of the resulting diversity metrics (Haegeman *et al.* 2013).

### **V.3.6. Data treatments and statistical analyses**

*Alpha-diversity analysis:* Richness (Chao1), diversity (Shannon and Simpson) and evenness (Pielou) indices were separately computed from the OTUs table for each sediment layer using the *diversity* function from the “vegan” package in R (Oksanen *et al.* 2015). For each time series, the permanent rares were defined as the OTUs that never reached more than 30 DNA sequences (0.1 % of the number of DNA sequences per sample).

*Beta-diversity analysis:* Non-metric multidimensional scaling (nMDS) analysis was utilized for each lake. For each lake, a Bray-Curtis dissimilarity matrix was built from the OTUs table and the *metaMDS* and *ordisurf* functions from the “vegan” package in R (Oksanen *et al.* 2015) were used to plot the nMDS results and gradients of phosphorus concentration and air temperature anomalies. The *env.fit* function was applied to the Bray-Curtis dissimilarity matrices (“vegan” package in R) (Oksanen *et al.* 2015) to evaluate the significance of relationships between MECs structure and the two forcing factors ([P] and airT). To reveal temporal patterns in MECs, a pseudo-autocorrelation analysis was used (i.e., correlation of a series with its own lagged values) to quantify the Bray-Curtis dissimilarity between each pair of samples; the time difference (lag) was then plotted against the values of similarity.

*Network analysis:* For each lake MEC, the construction of weighted correlation networks was performed following Langfelder and Horvath (2008) through their WGCNA package in R. To reduce the complexity of the 2 datasets, all permanently rare OTUs (see above for definition) were removed. The relative abundance datasets were standardized (using Hellinger transformation calculated by the *decostand* function from the vegan package in R) for further analyses. The thresholding power values of 6 were chosen for both networks. The adjacency matrices, obtained by the calculation of Pearson correlation values between each pairs of OTUs, were transformed into a topological overlap matrix to minimize possible errors. Subsequently, the OTUs were clustered in modules, with a minimum number of 8 OTUs per

module. In the two individual networks, only the OTUs (i.e., the nodes) linked by correlations (i.e., the edges) greater than 0.5 were kept for the visualisation and calculation of network topological indices (average degree, diameter, density, weighted modularity, clustering coefficient and average path length) utilizing the Gephi software (Bastian & Heymann 2009). To reveal strong positive or negative correlations between environmental factor changes and the temporal dynamics of modules, the eigengene values of modules were calculated for the modules from the two individual networks and the consensus modules. Topological roles were attributed to each OTU based on their intra- and inter- connectivity values. The threshold was arbitrarily defined in the present study as: (i) peripherals (intra < 10, inter < 2), (ii) connectors (intra < 10, inter > 2), (iii) modules hubs (intra > 10, inter < 2) and (iv) network hubs (intra > 10, inter > 2). To detect modules present in both networks, a procedure from the WGCNA package, based on the construction of a consensus network from the adjacency matrices from the two individual networks, was used. The OTUs from the consensus modules originated from either one or several modules from the individual networks.

*GAM analysis:* Generalized additive models (GAMs) (Hastie & Tibshirani 1986) analysis was performed to investigate the relationships between biological indices (richness, diversity, relative abundance and structure (i.e., based on Bray-Curtis dissimilarities)) and the two forcing factors (phosphorus concentration and air temperature) (see Figure **I.16**). The functions *gam* and *predict.gam* from the “mgcv” package in R (Wood 2015) were used and GAM parameterization was performed following Simpson and Anderson (2009)’s technical recommendations. In the present study, the predictor variables were the Daphnia-inferred phosphorus concentration and air temperature anomalies (annual mean values), while the responses variables were the chao1’s richness estimator, the Shannon diversity index and the values of the nMDS axes for the whole communities or for specific taxonomic groups. GAM analysis was also applied to investigate the relationships between environmental change and the temporal dynamics of modules in relative abundance (i.e., the sum of the DNA sequences from all the OTUs within a module). For all GAM analyses, only the correlations with p-values less than 0.01 were considered significant. Diagnostic plots (i.e., residuals vs. fitted values plots, residuals vs. time plots and observed values vs. fitted values plots) were visualized to check the robustness of performed GAMs.

Table V.1. The table presents for each lake the number of DNA sequences and OTUs within each microbial eukaryotic group and, the numbers of common OTUs and total OTUs in the two lakes.

Microbial eukaryote group		Lake Annecy		Lake Bourget		# Common OTUs	# Total OTUs
		#DNA sequences	#OTUs	#DNA sequences	#OTUs		
Alveolata	<b>Apicomplexa</b>	6925	46	13783	124	37	133
	<b>Ciliophora</b>	139132	784	363227	1121	530	1375
	Colpodellidae	1	1	80	2	1	2
	<b>Dinophyceae</b>	49459	582	43341	702	344	940
	<b>Perkinsea</b>	9128	94	5518	101	46	149
	Voromonas	1	1	8	4	1	4
	<b>unclassified_Alveolata</b>	267186	533	140973	527	281	779
Amoebozoa	Discosea			86	2		2
	Tubulinea	150	4	2	1	1	4
Cryptophyta	Cryptomonadales	28	4	167	9	2	11
	Cryptophyta_3			1	1		1
	Cryptophyta_4	19	4	129	6	3	7
	Pyrenomonadales	13	2	106	4	2	4
	<b>unclassified_Cryptophyta</b>			37	4		4
Haptophyceae	Coccosphaerales	3	2	3	2	1	3
	Pavloales			4	1		1
	Phaeocystales			23	5		5
	Prymnesiales			4	2		2
	<b>unclassified_Haptophyceae</b>	8599	11	377	14	6	19
	Zygodiscales			2	1		1
Heterolobosea	Schizopyrenida	1694	4	70	2		6
Nucleariidae and Fenticula group	Nucleariidae	15	1	47	1	1	1
Opisthokonta	Choanoflagellida	98	8	55	12		20
	<b>Fungi</b>	113962	954	222598	1990	599	2345
	Opisthokonta_incertae_sedis	1	1	27	1	1	1
Rhizaria	<b>Cercozoa</b>	27053	251	32756	362	141	472
	<b>unclassified_Rhizaria</b>	12375	93	7474	123	52	164
Rhodophyta	Florideophyceae			2	1		1
Stramenopiles	<b>Bacillariophyta</b>	9807	82	30906	150	55	177
	Bicosoecida	1089	20	2902	68	13	75
	Bolidophyceae			2	1		1
	<b>Chrysophyceae</b>	27322	94	34076	125	49	170
	Dictyochophyceae	4	2	55	4	2	4
	Eustigmatophyceae	996	12	8729	39	11	40
	Labyrinthulomycetes	107	11	159	13	4	20
	MAST	1073	13	9433	26	7	32
	Oomycetes	1465	38	4544	71	27	82
	Pelagophyceae			6	2		2
	Phaeothamniophyceae			2	1		1
	PX_clade	42	2	37	4	1	5
	Raphidophyceae	5	3	9	1	1	3
	Synurophyceae	1088	32	1602	49	17	64
	<b>unclassified_Stramenopiles</b>	10993	97	30136	162	56	203
	Viridiplantae	<b>Chlorophyta</b>	24770	226	87966	607	173
Streptophyta		857	28	1619	58	11	75
<b>unclassified_Viridiplantae</b>		52	4	372	9	2	11
Total		715512	4044	1043455	6515	2478	8081

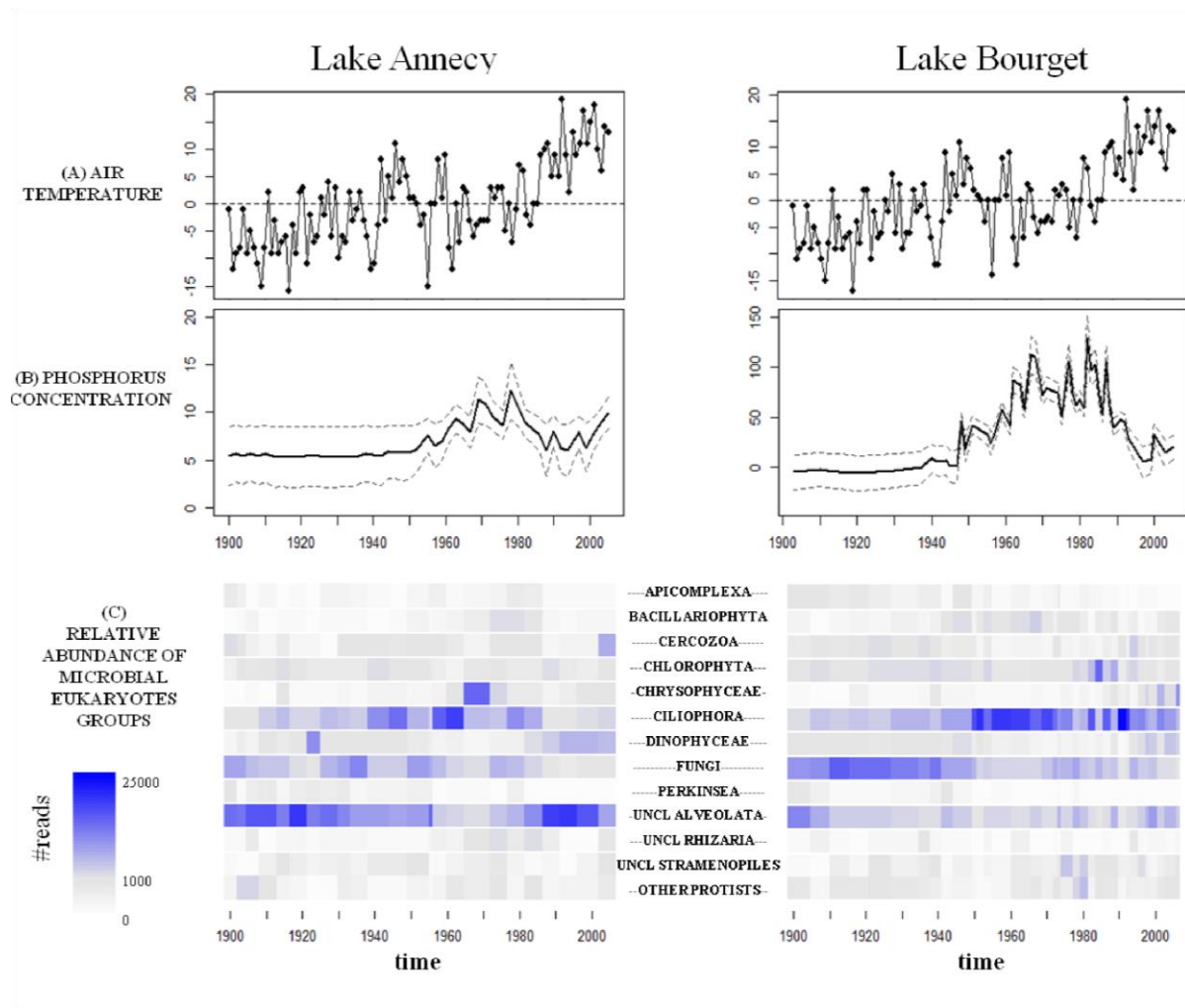


Figure V.1. Temporal changes of (A) air temperature anomalies (annual average expressed in  $10^{-1} \text{ }^{\circ}\text{C}$ ; data from Auer et al., 2007) (B) Daphnia-inferred total phosphorus concentrations (annual mean concentration over the 0-20 m water depths  $\pm 95$  confidence interval expressed in  $\mu\text{gP L}^{-1}$ ; data from Alric et al., 2013, Berthon et al., 2014) (C) relative abundance (in terms of number of DNA sequences) of each microbial eukaryotes group, illustrated by heatmaps. Data cover the last century and are presented for lakes Annecy and Bourget.



## V.4. Results

### V.4.1. Lake settings and environmental forcing factors

Lakes Bourget and Annecy are two deep monomictic lakes located at the northwest edge of the French Alps. The historical data of the following two major environmental forcings were considered (i) the anomalies in mean annual air temperature (airT), which increased in the areas of these two lakes, especially from the mid-1980s to 2000s, (Figure V.1.A; data from Auer *et al.* 2007) and (ii) the Daphnia-inferred total phosphorus concentrations ([P]) in water (Figure V.1.B; data from Alric *et al.* 2013, Berthon *et al.* 2014), which illustrates the contrasting levels of phosphorus enrichment for these two lakes.

### V.4.2. Richness, composition and structure of MECs

The molecular inventories of the 59 sediment layers (24 from Lake Annecy and 35 from Lake Bourget, covering a 100-year-long period) were all standardized at 29,813 DNA sequences and revealed a total of 8,081 OTUs. The estimated richness (Chao1 index) of microbial eukaryote (OTUs per sediment layer) varied from 188 to 940 OTUs in Lake Annecy and from 458 to 1,098 OTUs in Lake Bourget, which exhibited a significantly higher estimated richness (ANOVA p-value < 0.001; Table V.S1). The total richness was 4,044 and 6,515 OTUs for Lakes Annecy and Bourget, respectively, with 2,478 common OTUs found in both lakes. The Shannon index ranged from 1.9 to 5.1 ( $3.9 \pm 0.6$ ) in Lake Annecy and from 2.6 to 4.5 ( $3.5 \pm 0.6$ ) in Lake Bourget (Table V.S1).

The OTUs were distributed within 46 microbial eukaryotic groups. Among them, 12 groups represented more than 97 % of the total number of DNA sequences (1,710,866 DNA sequences), including strictly heterotrophic (Apicomplexa, Cercozoa, Ciliophora, Fungi and Perkinsea) or photosynthetic groups (Bacillariophyta and Chlorophyta) and groups containing heterotrophic, photosynthetic and mixotrophic taxa, such as Chrysophyceae and Dinophyceae (Figure V.1.C, Table V.1). Hence, for further statistical analyses we considered mainly the temporal dynamics of these 12 dominant groups. For the two lakes, only a few OTUs were highly abundant (14 and 15 OTUs in Lake Annecy and Lake Bourget, respectively) which together accounted for more than 50 % of the total number of DNA sequences (Table V.S2). A large portion of the highly abundant OTUs were also persistent through time (i.e., OTUs detected in all samples), with 7 and 12 persistent OTUs in Lake Annecy and Lake Bourget, respectively. Almost all persistent OTUs from Lake Annecy (6 out of the 7 OTUs), mainly

affiliated with unclassified Alveolata and Ciliophora, were also persistent OTUs in Lake Bourget. In both lakes, the majority of OTUs never reached more than 0.1 % of the total number of DNA sequences and were defined as permanently rare. These OTUs accounted for 3,571 OTUs (88 % of the total richness) in Lake Annecy and 5,916 OTUs in Lake Bourget (90 % of the total richness).

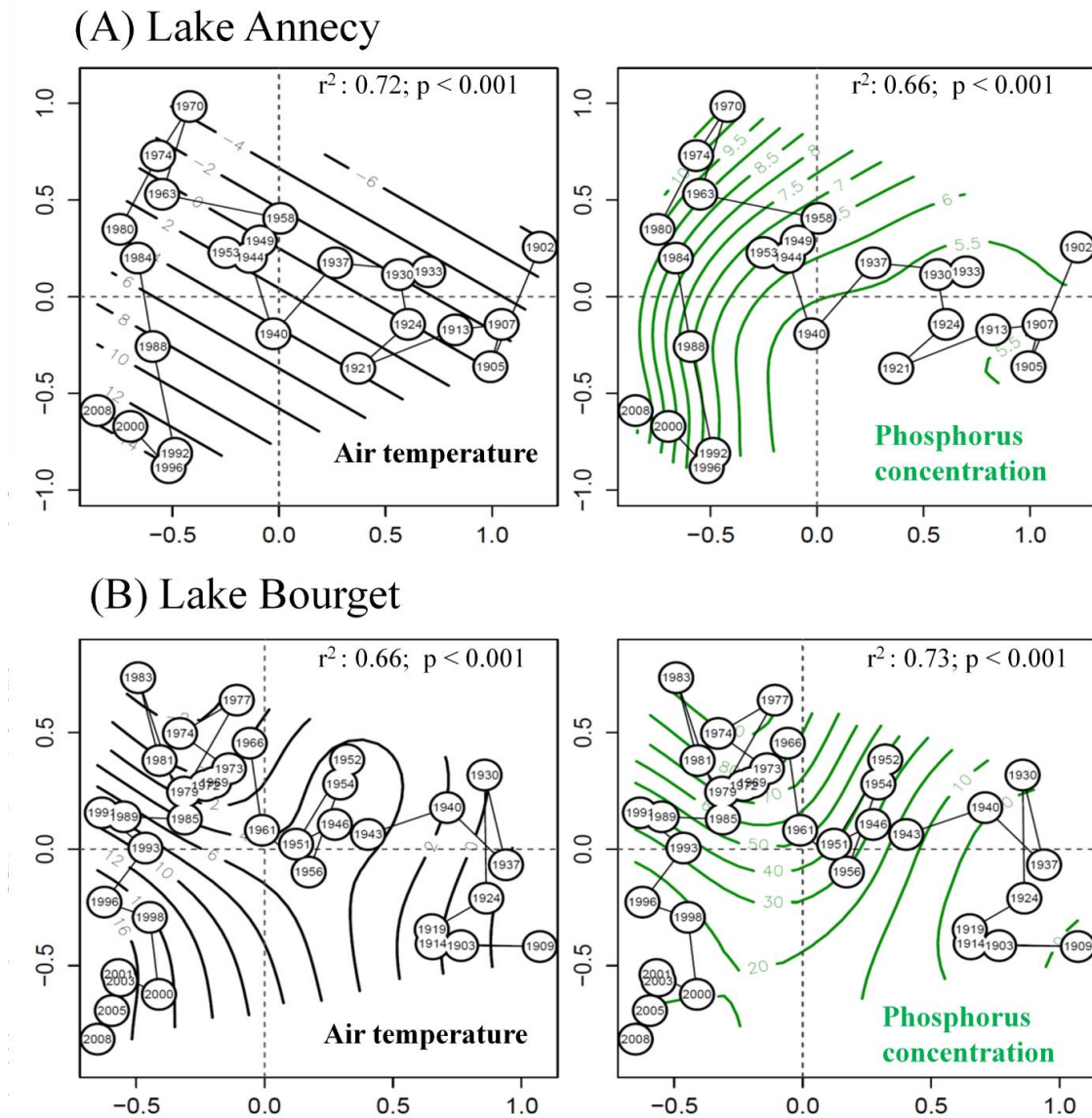


Figure V.2. Non-metric multidimensional scaling analysis of microbial eukaryotes communities based on Bray-Curtis dissimilarity obtained for Lake Annecy (A) and Lake Bourget (B) time series. Gradient of air temperature and phosphorus concentration were plotted with community structure data. The results of the analysis to evaluate the significance of the relationships between MECs structure and the two forcing factors (env.fit function) are presented on the top-right corner of each plot.

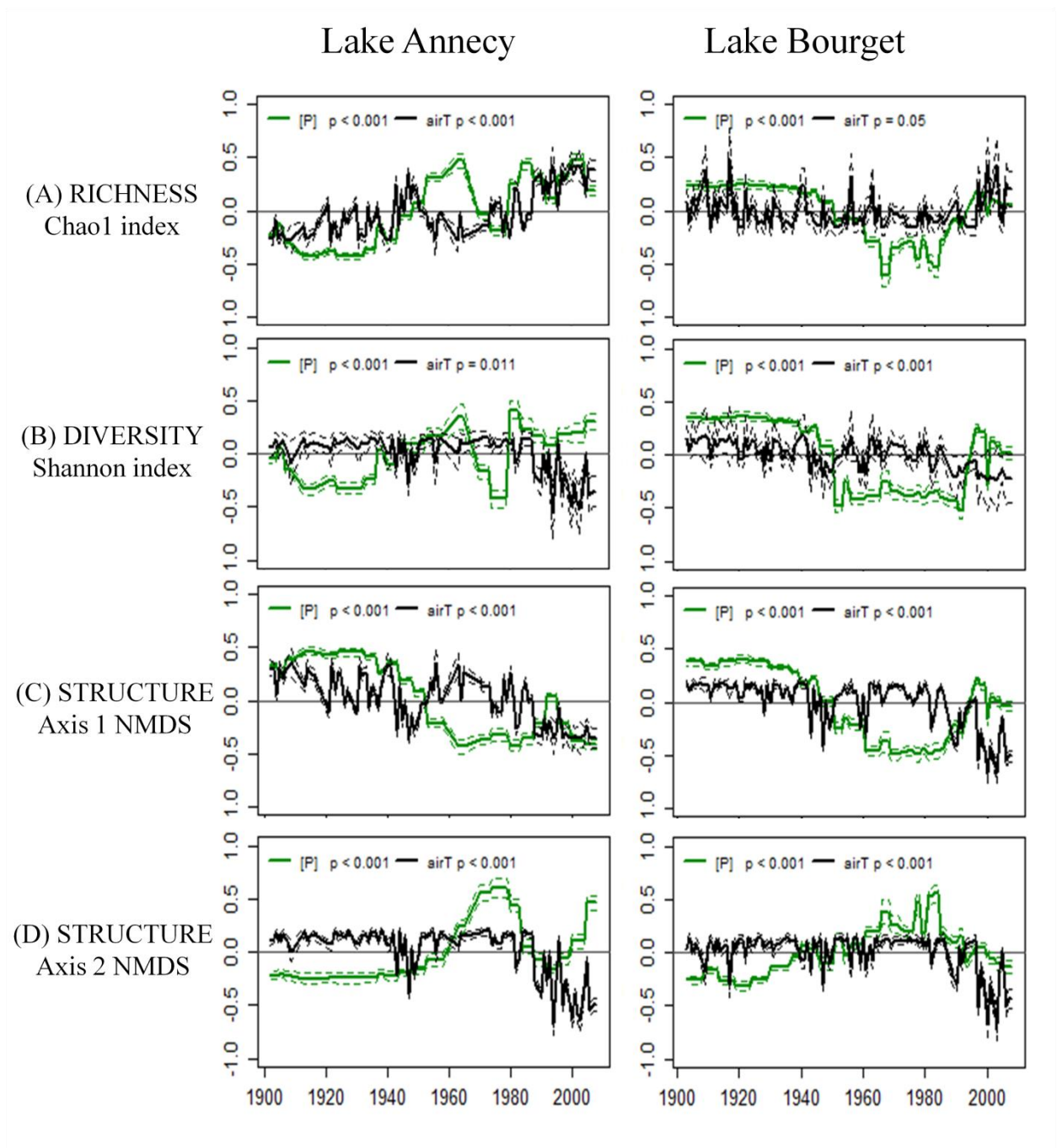


Figure V.3. GAMs analysis results for Lake Annecy and Lake Bourget. The figure presents the contribution of phosphorus concentration (green line) and air temperature (black line) to the MECs temporal changes considering (A) chao1 richness estimator (B) Shannon diversity index (C-D) the structure of the MECs (based on axis1 and axis2 of the nMDS plot). The model summary obtained for the additive model fitted to each index and the fitted smooth function for [P] and airT are presented in the Figure V.S2.

### V.4.3. Temporal rearrangements of MEC related to environmental change

As shown by the pseudo-auto Bray-Curtis correlation analysis (i.e., the correlation of a series with its own lagged values) for each lake, the similarity between samples tended to decrease with the increase in time-lag (time period that separates the 2 samples), illustrating the continual changes in MECs through time without any detection of a return to former conditions (Figure V.S1). The env.fit analysis was used to evaluate the significance of relationships between the MECs structure and the forcing factors and showed that temporal modifications in the MECs' structure (Bray-Curtis indices) were significantly correlated with changes in phosphorus concentrations and air temperature in both lakes (p-values < 0.001; Figure V.2).

GAMs (generalized additive models) analysis was applied to estimate when and how much each of the predictors contributed to the observed changes in MECs for the 2 lakes (see M&M for details). The contributions of the two predictors ([P] and airT) to the fitted values of the Chao1 estimator index, Shannon diversity, and nMDS axis scores are shown in Figure V.3 and Figure V.S2. For both lakes, the contribution of [P] is much larger than that of airT before the 80s. From the mid-80s, the contribution of airT increased and sometimes even surpassed the [P] contribution as a predictor of the MEC structures for both lakes (i.e., nMDS axes scores; Figure V.3). All metrics (i.e., richness, diversity and structure of the MECs) were significantly impacted by changes in [P] in both lakes (p-values < 0.0001), while the increased air temperature from the late 1980s primarily caused shifts in the MEC structures in the two lakes. The smooth functions clearly illustrated a moderate response of Chao1 and Shannon indices to airT and a stronger response to [P] (i.e., y-scale axis of the fitted functions; Figure V.S2). More precisely, we noticed that the fitted relationship between richness (Chao1 index) and [P] was nonlinear for Annecy lake, while a smooth linear function was found for Lake Bourget, showing that the richness increased with P enrichment as long as [P] did not exceed 10-15  $\mu\text{gP.L}^{-1}$  and decreased at the highest levels of P enrichment (> 15-20  $\mu\text{gP.L}^{-1}$ ).

Considering each taxonomic group independently, the GAM results showed that, in terms of structure, almost all microbial eukaryotic groups were impacted by [P] changes in both lakes (Figure V.S2). Modifications in their richness and relative abundance were observed concomitantly. The moderate [P] increase in Lake Annecy (up to 10-15  $\mu\text{gP.L}^{-1}$ ; from the 50s to mid- 60s) induced significant positive effects on the richness and relative abundances of Apicomplexa, Cercozoa, Chrysophyceae, Ciliophora and Dinophyceae, while the higher [P]

found in Lake Bourget had negative effects on these metrics (time period: early 50s to early 90s) (Figure V.S2). More precisely, the highest richness and relative abundances of Ciliophora were observed when the [P] ranged between 20 and 80  $\mu\text{g.L}^{-1}$  in Lake Bourget (early 50s to the early 90s). For Dinophyceae, Apicomplexa and Cercozoa, this type of “optimal” [P] was around 8  $\mu\text{g.L}^{-1}$  (in Lake Annecy), while the richness and the relative abundance of Chlorophyta seemed to favour at higher [P] conditions when the [P] reached 80  $\mu\text{g.L}^{-1}$  (Figure V.S2 Chlorophyta section). Air temperature changes strongly impacted structure of most of the taxonomic groups, with significant effects detected mainly since the late 1980s (Figure V.S2). These changes were concomitant with changes in the richness and/or relative abundances of these taxonomic groups. Some groups showed positive responses in terms of richness and relative abundance, which was the case for Chrysophyceae and Dinophyceae when the airT anomalies were greater than 0.5 °C. Other groups, such as Apicomplexa, Fungi and Perkinsea, had a negative response (i.e., a decrease in the metrics as airT increased) (Figure V.S2). The simultaneous effects of the changing phosphorus concentrations and air temperatures, which were particularly noticeable since the late 1980s, were detected for Chrysophyceae, Ciliophora, Dinophyceae and Fungi for all 3 metrics (richness, relative abundance and structure). In Lake Annecy, the airT and [P] changes appeared to induce mostly antagonistic responses (i.e., opposite effects of the two regulatory factors). In contrast, the majority of the coupled effects in the Lake Bourget MECs were cumulative with simultaneous positive (e.g., on Chrysophyceae richness) and negative effects (e.g., on Fungi relative abundance). Similar cumulative effects were detected in the two lakes for Dinophyceae richness (positive impact of [P] and airT on richness and relative abundance) and Perkinsea relative abundance (negative impact) only.

Table V.2. Properties of the two individual networks built for Lake Annecy and Lake Bourget MECs. The topological indices were calculated considering only the edges values superior to 0.5.

<b>Metrics</b>	<b>Annecy</b>	<b>Bourget</b>
Number of nodes	473	599
Number of edges	8014	13143
Average degree	16.9	21.9
Diameter	7	8
Density	0.072	0.073
Weighted modularity	0.657	0.598
Number of modules	20	20
Clustering coefficient	0.622	0.674
Average path length	3.112	3.539



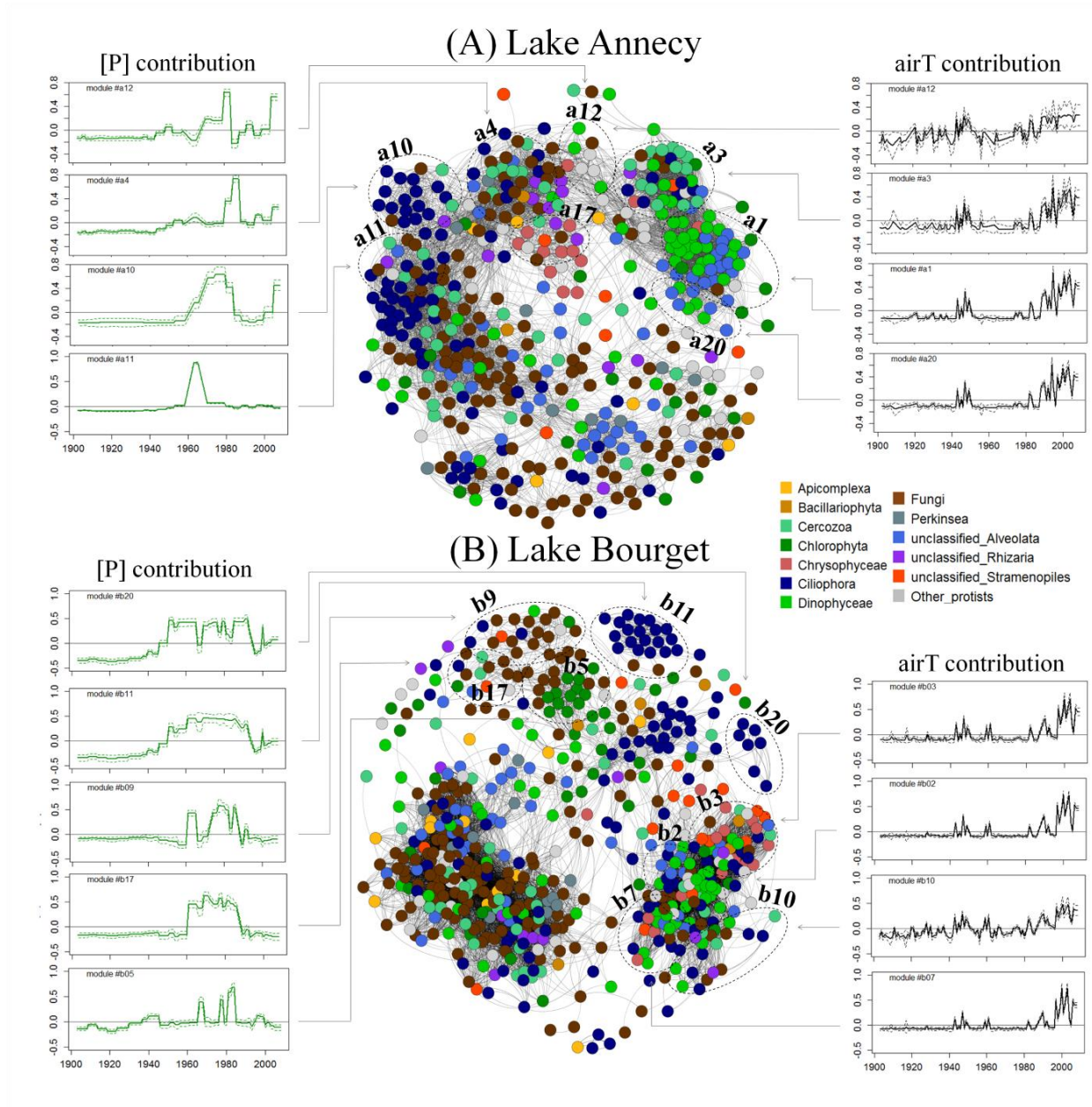


Figure V.4. Co-occurrence networks for Lake Annecy and Lake Bourget MECs. Each node represents one OTU. Different colors correspond to different taxonomic affiliation. The edges correspond to significant Pearson correlation values calculated between each pair of OTUs, only the edges with values superior to 0.5 were illustrated here. In each network, the modules found to be related to environmental factors were surrounded with dotted lines. The temporal contribution of the environmental factors to the relative abundance of the modules was estimated by GAMs analysis. Only the contribution from modules which had module eigengene values positively correlated to environmental changes are presented here (see Figure V.S4 for details).

#### **V.4.4. Network structure, modularity, connectivity and links with environmental changes**

Co-occurrence networks were built for each lake by considering all OTUs that represented more than 0.1 % of the DNA sequences in at least one sample to remove any OTUs that were defined as permanently rare. Only positive Pearson's correlations  $> 0.5$  were kept to determine the co-occurrence patterns within MECs (see M&M for details). The resulting networks consisted of 473 nodes linked by 8,014 edges for Lake Annecy and 599 nodes linked by 13,143 edges for Lake Bourget (Table **V.2**, Figure **V.4**). The diameter, showing the distance between the two most distant nodes, was similar for both networks. The two networks were constituted by highly connected OTUs (mean number of edges per node greater than 15) structured in 20 modules, with a clustering coefficient of approximately 0.6 (Table **V.2**).

Based on their values of connectivity (intra- and inter-modularity connectivity), topological roles were attributed to identify OTUs acting as connectors, module hubs, network hubs, or peripherals (see M&M for details). Most of the OTUs were peripherals ( $> 60$  % of the constitute OTUs; Figure **V.S3**). The connectors (19 % and 28 % of the constitutive OTUs in Bourget and Annecy, respectively) were represented predominantly by Fungi, Ciliophora and Dinophyceae. Moreover, 29 OTUs were found to be common connectors in both lakes and mainly belonged to heterotrophs (Fungi, Cercozoa, Ciliophora and unclassified Alveolata) but also included Chlorophyta, Chrysophyceae and pigmented Dinophyceae taxa. OTUs acting as module hubs or network hubs represented less than 10 % of the constitutive OTUs and were specific to each network. Dinophyceae and Cercozoa were, for instance, identified as network hubs in both lakes, but no common OTUs were detected between them.

Most modules exhibited temporal dynamics with changes in abundances (in terms of the number of DNA sequences of their constitutive OTUs) during specific periods. To investigate the modules' response to environmental conditions, correlations between the module eigengene values and the environmental factors were calculated, and GAM analysis was applied to the module's relative abundance (Figure **V.S4**). Additionally, the existence of common responses in both lakes was investigated by constructing a consensus network (see M&M for details), which included 5 consensus modules constituted by 86 OTUs (out of 814 total OTUs) (Figure **V.S5**). Nine modules from the 2 individual networks were positively ( $p < 0.05$ ) correlated with phosphorus concentration (Annecy: modules #a4, #a10, #a11, #a17; Bourget: modules #b5, #b9, #b11, #b17, #b20; Figure **V.S4**). The contributions of [P] to

changes in the relative abundances in the 9 modules were mainly detected when [P] increased moderately (range: 10 to 20  $\mu\text{g.L}^{-1}$ ) (Figure V.4) in the 1960-70s in Lake Annecy and in the 1950s in Lake Bourget. The 14 Alveolate\_1 OTUs from modules #a10 and #b11 formed a consensus module (module #blue) that was driven by the increase of phosphorus concentration ( $> 15 \mu\text{g.L}^{-1}$ ; Figure V.S5). When an intermediate level of [P] was reached (20-60  $\mu\text{g.L}^{-1}$ ), the module #b20, composed of Ciliophora OTUs affiliated to *Epispathidium sp.*, became abundant, while the module #b5, composed mainly of Chlorophyta OTUs assigned to *Chlamydomonas sp.*, was correlated with the highest [P] (i.e., greater than 80  $\mu\text{g.L}^{-1}$ ) and was detected only in the eutrophicated Lake Bourget. From the individual networks, six modules appeared to be positively driven by the increase in air temperature since the 1980s (modules #a1, #a3 and #a20 in Annecy and modules #b2, #b3 and #b7 in Bourget; Figure V.4). From these modules, 34 OTUs constituted the module #turquoise and were strongly correlated to air temperature ( $r^2$  0.77,  $p < 0.001$ ) and mainly composed of Dinophyceae and unclassified Alveolata.

## V.5. Discussion

### V.5.1. Characterization of MECs from sedimentary archives

Various methodological precautions are required to achieve robust findings using paleo-environmental DNA (Boere *et al.* 2011a, Rizzi *et al.* 2012, Torti *et al.* 2015). As presented in the Material & Methods section, we applied strict laboratory protocols and stringent analytic and bioinformatics procedures to ensure the validity of our molecular data. These choices theoretically minimized the artificial inflation of the OTU numbers.

A relatively high richness in the MECs from Lake Bourget was revealed over the last 100 years (6,515 OTUs) in comparison with the richness detected for the last 2,100 years in the same lake applying the same methodology (11,344 OTUs; Capo *et al.* in press). Studies investigating lacustrine MEC diversity generally differ in their experimental procedures (choice of barcodes, depth of sequencing and the stringency of bioinformatics treatments). Comparisons between studies should thus be considered cautiously. From the only other comparable study, which explored MEC richness by applying HTS to sed-DNA (from the Black Sea sediments), 2,700 eukaryotic OTUs were reported for a period of 10,000 years. This moderate richness is partly explained by the lower sequencing depth obtained in this study (around 1,300-3,300 DNA sequences per sample; Coolen *et al.* 2013). Comparing the



taxonomic composition (proportion of the main phylogenetic groups within MECs) is less problematic than comparing only the taxa number, and when considering the assemblage composition from paleo-reconstructions (few references are available Coolen *et al.* 2013, Capo *et al.* 2015) and limnological survey (Lepère *et al.* 2013, Debroas *et al.* 2015, Filker *et al.* 2016), a general pattern is observed. Two groups generally dominate lacustrine microbial eukaryotes: Alveolata, which are mainly represented by Dinophyceae and Ciliophora, and Fungi, which are mainly represented by Dikarya, Chytridiomycota and Cryptomycota. Stramenopiles, Viridiplantae and Rhizaria are 3 other dominant groups within these assemblages, with Chrysophyceae, Chlorophyta and Cercozoa being typical representatives of these 3 groups, respectively (Taib *et al.* 2013, Kammerlander *et al.* 2015, Debroas *et al.* 2015, Capo *et al.* 2015, Oikonomou *et al.* 2015, Filker *et al.* 2016). From limnological and paleolimnological molecular inventories, the heterotrophic taxa (mostly Ciliophora and Fungi and, to a lesser extent, Cercozoa) are generally predominant in terms of relative abundance of DNA reads in comparison to phototrophs. These conserved patterns tend to support the idea that sedimentary DNA do not give a distorted view of eukaryotic assemblages. As previously reported for MECs in Lake Bourget (Capo *et al.* 2015), most of the microbial eukaryote taxa were well archived in this deep calcareous lake (with the exception of Cryptophyta and Haptophyta).

The exploration of network co-occurrence offers an extension to DNA sequences analyses but has rarely been used to characterize MECs in lacustrine systems (exceptions: Vick-Majors *et al.* 2014, Kammerlander *et al.* 2015), and MEC co-occurrence patterns have never been constructed from long term data. The centennial view of microbial eukaryote co-occurrences, provided here for both lakes, confirmed that the microbial eukaryote community was non-randomly structured through time. The structures of the 2 networks (Bourget and Annecy) were similar in terms of diameter, clustering coefficient and average path length, even though the number of nodes and edges was higher for Bourget. Both networks were characterized by rather high clustering coefficient values ( $> 0.6$ ), which was consistent with another time-varying network that included protists (Vick-Majors *et al.* 2014). The clustering coefficients found for bacterial networks were systematically less than or equal to 0.5 (Chaffron *et al.* 2010, Steele *et al.* 2011, Lima-Mendez *et al.* 2015), showing stronger positive links between the MECs than bacterial communities.

Based on the attribution of topological roles to each OTU, we were able to verify the existence of common connectors in the two networks (OTUs with high inter-modular

connectivity in both lakes). This suggests a key role of these taxa in network structuration. Noticeably, among these strongly interactive (and generally abundant) OTUs, some could not be precisely taxonomically assigned. This is probably due to the incompleteness in the reference database (for instance, OTUs 23615 & 20906, two persistent OTUs and connectors, were assigned to unclassified Alveolata) and illustrates the need to continue updating barcode reference databases to obtain more complete coverage of the eukaryotic diversity. For all metrics, (richness, composition and topological structure of networks), the significant role of environmental changes ([P] and airT) was detected, at least during specific temporal periods over the last century.

### **V.5.2. The [P] range as a major driver of MECs in both lakes**

Increased nutrient loading (especially phosphorus) into lakes is known to induce shifts in the composition of specific biological compartments, such as fish, zooplankton and phytoplankton assemblages (Jeppesen *et al.* 2000, Paerl 2006, Ibelings *et al.* 2007). Here, our results showed that the effects of this bottom-up pressure were also detected when considering the whole microbial eukaryote community, with impacts on richness, structure and the relative abundances of various groups within the MECs. MEC richness increased with P enrichment as long as [P] did not exceed  $15 \mu\text{gP.L}^{-1}$  and decreased at the highest levels of P enrichment. These observations suggest that the global richness of MECs may follow an unimodal response to [P] enrichment, as already detected for phytoplanktonic species richness in several studies using microscopic analyses (Reynolds *et al.* 1993, Sommer *et al.* 1993, Jeppesen *et al.* 2000, Soares *et al.* 2013).

Even moderate changes in [P], as observed in Lake Annecy, induced responses at multiple trophic levels within the microbial eukaryote assemblage, leading to significant modifications within the structures of both heterotrophic and pigmented groups. Among the purely heterotrophic groups found to be driven by [P] conditions, Ciliophora exhibited a very clear response to [P] changes, increasing (% of reads) when [P] was greater than  $15 \mu\text{g.L}^{-1}$ . This response was consistent with previous relationships reported between the abundance of ciliates and nutrient enrichment in the lakes (Beaver & Crisman 1982, Kisand & Zingel 2000, Wetzel 2001, Triadó-Margarit & Casamayor, 2012). The predominance of those taxa, known as bacterivores and algae grazers (Pace 1982, Sommer *et al.* 1986, Posch *et al.* 2015), in productive ecosystems (meso- to eutrophic status) can be explained by an increase in the abundance of their preys following eutrophication. An interesting parallel can be drawn, for

instance, with other recent paleolimnological data obtained in Lake Bourget, which shows a concomitant increase of picocyanobacteria (from the 50s), a type of cyanobacteria that is a food resource for ciliophoran (Domaizon *et al.* 2013, Callieri & Stockner 2002). The response of ciliophoran to [P] changes was also detected at the network level; ciliophoran OTUs (mostly affiliated with Alveolate\_1) constituted a consensus module and were predominant in both lakes with moderate [P] enrichment (between 15 and 50  $\mu\text{g.L}^{-1}$ ). The observed changes in their relative proportions, diversity and structure suggest that Ciliophora could be an interesting group to focus on in future investigations of paleolimnology, particularly for exploring the relevance of this group as an indicator of [P] changes in paleorecords. Ciliophora have already been identified as potential relevant indicators of water quality (eutrophication level) in coastal waters and rivers (Madoni & Bassanini 1999, Jiang *et al.* 2013, Xu *et al.* 2014, Pawlowski *et al.* 2016); however, the transposition to a paleolimnological approach must be verified.

Conversely, negative relationships were reported between [P] increases and other purely heterotrophic groups, specifically, parasitic and/or saprotrophs microbes, such as Apicomplexa, Perkinsea and Fungi groups. Changes in the quality of organic matter with eutrophication, which are characterized by high phytoplanktonic productivity and consequently the modification of the balance between allochthonous and autochthonous carbon, as reported in Lake Bourget (Giguët-Covex *et al.* 2010), may have contributed to the decrease of some Fungi known to preferentially degrade complex allochthonous organic matter (Stelzer *et al.* 2003, Pascoal *et al.* 2005, Tank *et al.* 2010). In regards to the parasitic groups Perkinsea, Apicomplexa and Chytrids (Kagami *et al.* 2007, Mangot *et al.* 2011, Rasconi *et al.* 2012), their decline might be linked to the temporal dynamics of their hosts, probably themselves driven by environmental change. The lack of knowledge regarding specific host-prey interactions does not allow for any conclusions to be made about their potential synchronically temporal dynamics. For instance, little is known about the host specificities of Perkinsea, a recently revealed parasitic group in lacustrine systems that could be a parasite of phytoplanktonic taxa (Chlorophyceae, Chrysophyceae) (Mangot *et al.* 2011, 2013). Overall, our data provide new material for generating hypotheses regarding the impact of [P] changes on heterotrophic microbial eukaryotes.

Among pigmented groups, the response of Chlorophyta to [P] was quite explicit, with Chlorophyta becoming particularly abundant (% of reads) when the lake water column reached a certain [P] threshold in Lake Bourget (around 80-100  $\mu\text{gP.L}^{-1}$ ). This observation is

consistent with the knowledge regarding the abundance of these algae in eutrophic conditions, as revealed by both limnological (Jensen *et al.* 1994, Naselli-Flores & Barone 2000, Moss *et al.* 2003, Soares *et al.* 2013) and paleolimnological records (Fee 1979, Engstrom *et al.* 2006). Diatoms are the other pigmented group significantly affected by [P] changes (i.e., Bacillariophyta). Those algae, which are broadly used as a bioindicator in both limnology and paleolimnology (Battarbee 2000, 2002, Davidson & Jeppesen 2013), are also sensitive to [P] in both lakes but are not dominant in microbial eukaryotes assemblages (< 0.05 % of the total number DNA sequences). Although the correspondence between molecular data and fossil remains from diatoms has not been validated, shifts in diatoms abundance were detected concomitantly with the [P] increase (from different families in both lakes), as previously reported by microscopic analysis (Berthon *et al.* 2014). The other phototrophic groups (Dinophyceae and Chrysophyceae) were simultaneously positively impacted by [P] changes and air temperature as discussed below.

### **V.5.3. No return to pre-eutrophication conditions**

As previously observed for various biological proxies from the sediment records of both lakes, the re-oligotrophication of those lakes (i.e., decrease in [P]) did not result in reverse trajectories of the biological communities (including cladoceran and chironomid assemblages) (Alric *et al.* 2013, Frossard *et al.* 2013, Perga *et al.* 2015). This pattern was also reported for other European lakes when considering various biological compartments (based on phytoplankton and zooplankton biomass and amounts of pigments) (Jeppesen *et al.* 2007, Sondergaard *et al.* 2007). Here, we confirm this pattern for the whole microbial eukaryotes community (i.e., phytoplanktonic and microbial heterotrophic compartments). The non-return of MEC to pre-eutrophication conditions was revealed by the high OTUs replacement (gain and loss of taxa) through time and the gradual and inescapable temporal changes in the structure of MECs, as shown by the Bray-Curtis analysis pseudo-autocorrelation.

The non-return trajectories in both lakes, especially since the late 1980s, can be partly explained by the role of climatic warming when re-oligotrophication occurred. Climate warming is known to alter the richness and composition of biological communities in aquatic systems (Moss *et al.* 2003, Mooij *et al.* 2005, Anneville *et al.* 2005, Posch *et al.* 2012, Dokulil 2014), inducing, for example, peaks in plankton biomass, the dominance of cyanobacteria and smaller-sized zooplankton, and higher bacterivory activities (Moss *et al.* 2003, Jeppesen *et al.* 2005, Paerl & Huisman, 2008, Posch *et al.* 2012). This forcing factor could have both direct

and indirect (e.g., by altering the thermic structure of lakes) effects on biological communities.

Here, we observed the emergence of certain microbial OTUs (in terms of richness and relative abundance increases) and modifications in ecological networks, which were linked to concomitant climate warming and re-oligotrophication, and contributed to the non-return trajectory observed for both lakes. In particular, Dinophyceae, a taxonomic group known to be sensitive to climatic fluctuations (Vernal *et al.* 2001, Londeix *et al.* 2009, Chen *et al.* 2011) and nutrients inputs (in marine systems) (Klouch *et al.* 2016), has shown increased richness and relative abundance since the late 1980s in both lakes.

Chrysophyceae also became highly abundant during the recent warming period, but only in Lake Bourget. Dinophyceae, Chrysophyceae was previously found to be favoured by favoured reductions of nutrient loading in deep lakes (observations made using microscopy counts of phytoplankton) (Jeppesen *et al.* 2005, 2007), and our data showed that the simultaneous effects of [P] and airT were factors in explaining their recent dynamics.

Climate warming could have curbed the effects of the re-oligotrophication process. However, it is still difficult to disentangle the interplay between the two forcing factors (potential antagonist/synergetic interactions) and their relative influences on biological community responses (Anneville *et al.* 2005, Battarbee *et al.* 2010, Moss 2012, De Senerpont Domis *et al.* 2013, Caron & Hutchins, 2013, Perga *et al.* 2015).

#### **V.5.4. Simultaneous effects of re-oligotrophication and warming since the late 1980s**

We applied GAMs analysis to determine the potential synchronic effects of the two regulatory factors on biological groups over time. In both lakes, simultaneous effects of [P] and airT were detected in 6 microbial groups (Apicomplexa, Chrysophyceae, Ciliophora, Dinophyceae, Fungi and Perkinsea), while no concomitant significant effects were observed for the Bacillariophyta, Cercozoa and Chlorophyta, groups most impacted by phosphorus concentration changes.

Synchronic effects were mainly detected after the late 1980s, when marked positive anomalies of temperature and re-oligotrophication in the lakes occurred. Several studies previously reported the interactions between the effects of climate warming and nutrient enrichment, but were mostly conducted in eutrophic or hyper-trophic lakes and mainly focused on phytoplanktonic communities. Generally based on *in situ* observations and modelling, these

studies pointed out synergistic effects, and nutrient enrichment was identified as the predominant structuring factor for phytoplankton (Jeppesen *et al.* 2005, Elliott & May, 2008, Elliott 2012, Rigosi *et al.* 2014). In some cases, however, the temperature was found to have greater influence than nutrients, for instance as a driver of the formation of certain cyanobacterial blooms (Paerl & Huisman 2008, Kosten *et al.* 2012, Posch *et al.* 2012).

The rare studies investigating these potential interactive effects in oligotrophic lakes have led to contrasting results, in which water temperature or nutrient inputs have been alternatively identified as the primarily driver of phytoplanktonic community composition (Blenckner *et al.* 2007, Shimoda *et al.* 2011, Rigosi *et al.* 2014). A meta-analysis (> 1000 lakes) by Rigosi *et al.* (2014) illustrated that the relative importance of these two factors could depend on the lake trophic state and phytoplanktonic taxon. For cyanobacterial blooms, for instance, nutrients were supposed to play a larger role in oligotrophic lakes, while temperature was more important in mesotrophic lakes. Our long term data confirm the significant synchronic effect of phosphorus concentration and air temperature changes (since the late 1980s) in both pigmented and non-pigmented groups (Chrysophyceae, Ciliophora, Dinophyceae, Perkinsea and Fungi). Interestingly, only in the highly eutrophicated Lake Bourget were these synchronic effects mostly cumulative effects, consistent with observations reported by Rigosi *et al.* (2014) about cyanobacteria. This result also suggests that the past trajectory/history of the ecosystem (e.g., level of eutrophication) influences the response of ecosystems to recent climate warming.

## **V.6. Conclusion**

Our results illustrate the great potential of the application of HTS to sedimentary DNA for the reconstruction of past lacustrine biodiversity and the exploration of the relationships between environmental changes and biodiversity dynamics. The temporal framework (over 100 years) provides new insights into the long-term dynamics of microbial eukaryote communities in response to long-term changing environmental factors. Such temporal depth is extremely rare compared to other ecological approaches dealing with the temporal dynamics of microbial communities. Although these approaches are promising in further revolutionizing our understanding of long-term ecological dynamics, careful calibration studies are still needed, as with any paleolimnological proxy. In particular, the potential differential preservation of DNA in sediments according to environmental conditions and taxa is a key issue in paleogenetics and needs further investigations. In this study, we checked to see whether the decrease of some taxa in the oldest sediment layers could be due to a low preservation capacity (i.e., preferential decrease of taxa unable to produce resistant cells allowing physical protection for their DNA). Although not the case here, studies dealing with longer time scale reconstruction (> 100 yrs) are more likely to encounter this potential problem. Furthermore, because the response of lacustrine biological communities to climate warming and nutrient enrichment depends on various intrinsic and extrinsic properties of the systems, the generalization of our results must be tested using a sufficient number of lakes selected for their specific typologies and ecological histories (multi-lakes approach).

## V.7. Additional analysis

### V.7.1. Detection of change points in MECs richness/diversity (bcp analysis)

Richness (Chao1), diversity (Shannon, Simpson) and evenness (Pielou) indices were computed from the OTUs table separately for each sediment layer using the R package “vegan” (Oksanen *et al.* 2015). Bayesian change points (bcp) analysis was used to reveal timing changes in these indices. The bcp analysis provides estimates of posterior probabilities for the presence of a change point between consecutive dates and of posterior mean values of each index. The R package “bcp” (Erdman & Emerson 2007) was used for this analysis.

In Lake Annecy, a first change point was detected in the 1930s associated with the increase of the Chao1 and Simpson indices through all the time series while the second change point was associated to an increase of Simpson index values in the late 1980s (Figure V.5). The changes of those metrics was found to be favored by the weak increase of phosphorus concentration from the 1940s to the 1970s (Figure V.3) while here, the bcp analysis revealed that a major change was detected before. The changes in richness and diversity metrics were in accordance with the previous results that linked the air temperature increase with biological metrics in the late 1980s as described in GAMs analysis (Figure V.3).

In Lake Bourget, a decrease of the diversity and evenness indices was observed from the mid 1940s to the 1990s. In the 1990s, major change points in the diversity and evenness indices were detected (associated with the abrupt decrease and return to baseline on a 3-5-years time period). The number of OTUs, Shannon index and Pielou changes appeared to be negatively driven by the strong increase of phosphorus concentration from the 1940s to the 1990 (ranged from 10 to 100  $\mu\text{g.L}^{-1}$ ) as revealed by GAMs analysis (Figure V.3). In the 1990s, the abrupt change in diversity and dominance detected by bcp analysis was related to a Ciliophora OTU that became highly abundant sporadically (21 542 reads in the 1991-1992 sediment layer; 16 030 reads in the 1993 sediment layer; 72 and 54 % of the reads in each sample respectively).

Altogether, bcp analysis provide contrasted results with the changes detected by GAMs analysis for Lake Annecy (in particular about the increase of richness before [P] increase) and for Lake Bourget (no change points detected in the 1940s while the strong effects of [P] was revealed in the results from the paper (GAMs analysis, network analysis).



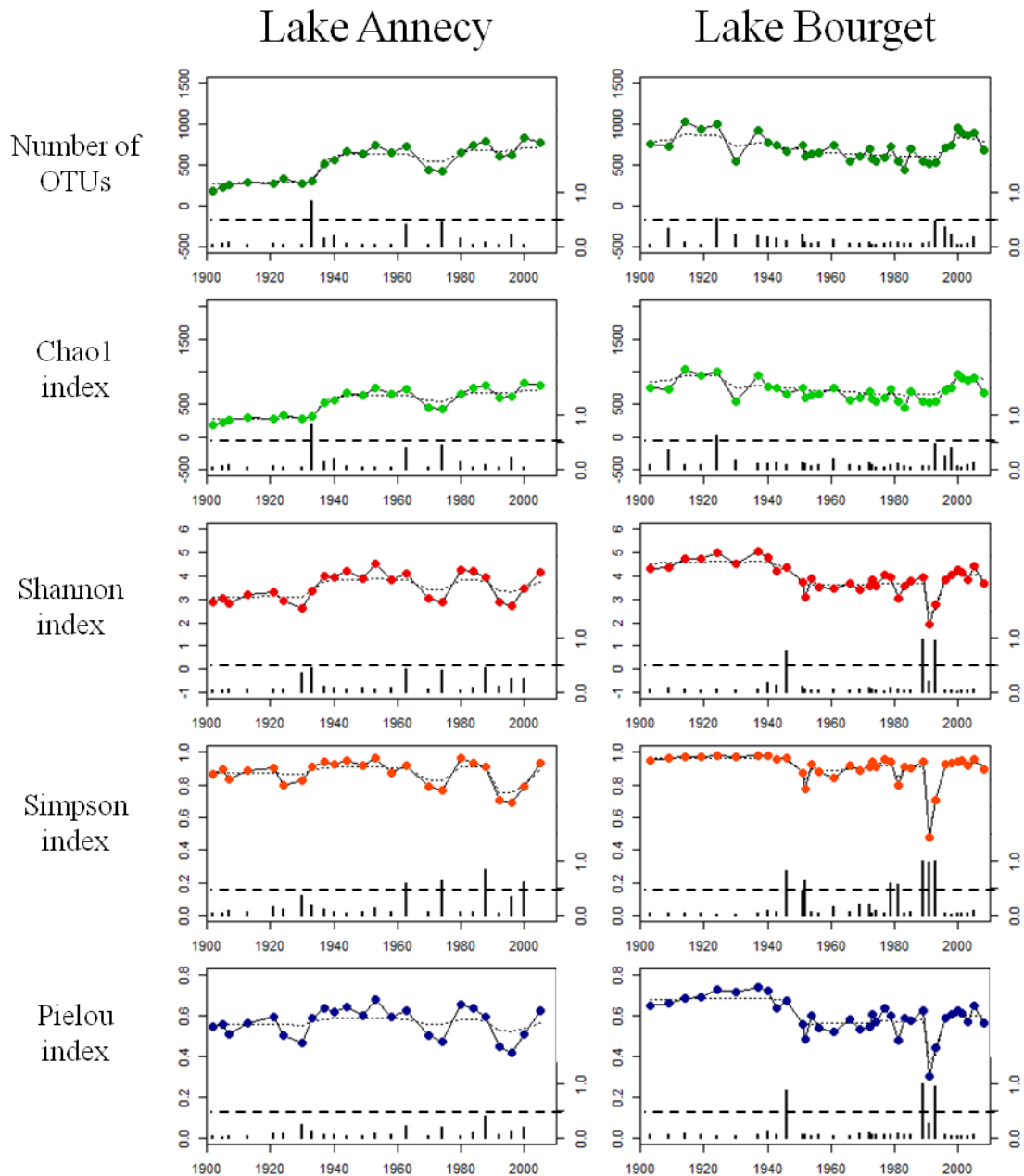


Figure V.5. Time series of richness, diversity and evenness indices for Lake Annecy and Lake Bourget. In the upper part of the graph, the colored line shows the metrics values and the dashed line shows the sample mean for the time series. The results of bcp analysis are plotted on the lower part of each graph showing the posterior probability of a change point at a given time. Timing changes are considered as relevant if the posterior probability values are equal or superior to 0.5.

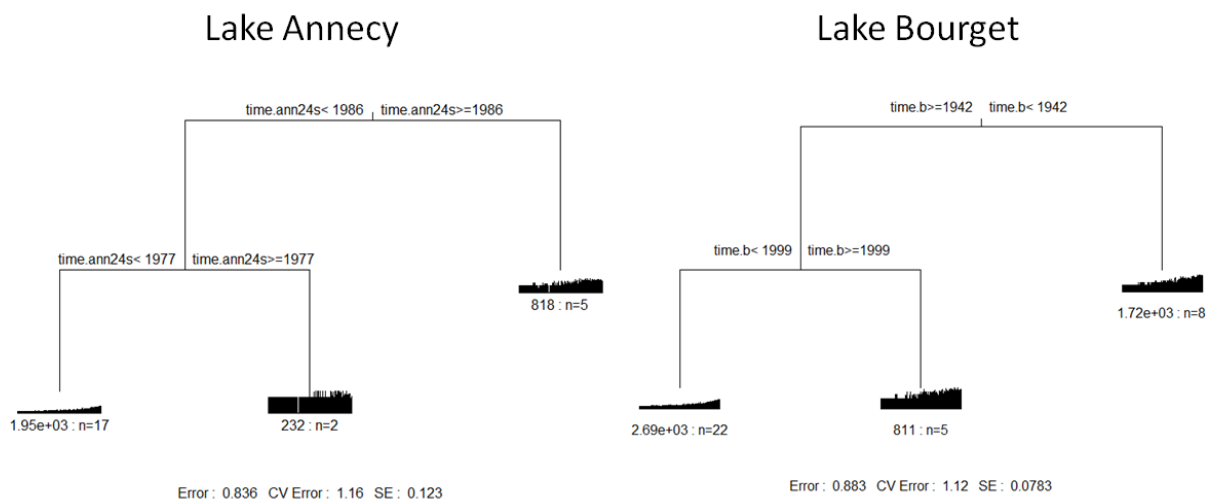


Figure V.6. Breakpoints detected in the structure of MECs by MRT analysis

### V.7.2. Detection of break points in the structure of MECs (MRT analysis)

A chronological clustering was used to detect the timing and hierarchy of marked breakpoints in the community structure. For this purpose, a Multivariate Regression Tree (MRT) analysis was performed considering the overall OTUs table or using separately the non-pigmented and the pigmented OTUs tables. This time-constrained analysis was performed according to the instructions of Borcard *et al.* (2011). The size of each tree was selected using the graphs of the relative error. The R package “mvpart” (De’ath 2002) was used for this analysis.

MRT analysis revealed breakpoints in the structure of Annecy MEC mainly related to the recent period (primarily in the mid-1980s and a secondarily in the late 1970s). Those results revealed tipping points probably linked to the changes in phosphorus concentration in the late 1970s (decrease) and air temperature increase since the 1980s (Figure V.6). The GAMs results showed that the increase in phosphorus concentration in the 1940s induced changes in MEC structure (with GAMs analysis) in the 1940s, but also that temperature had an increased impact on the response of MECs since the mid 80s; those additional results suggest that changes in recent conditions induced the major shift detected in the temporal modifications of MECs ‘structure in Lake Annecy. In Lake Bourget, the MRT analysis showed that the main changes in MEC structure were detected in 1940s and late 1990s (Figure V.6), consistently with the changes highlighted by GAMs analysis (see V.4.3), revealing that the modification in nutrients concentration was the main factor driving the microbial eukaryotes communities in Lake Bourget.

Table V.3. Distribution of total OTUs and oscillating OTUs within each microbial eukaryote group for each lake independently

Microbial eukaryote group	Lake Annecy		Lake Bourget		
	#OTUs	#Oscillating OTUs	#OTUs	#Oscillating OTUs	
Alveolata	<b>Apicomplexa</b>	46	1	124	4
	<b>Ciliophora</b>	784	4	1121	22
	Colpodellidae	1		2	
	<b>Dinophyceae</b>	582	17	702	21
	<b>Perkinsea</b>	94	2	101	3
	Voromonas	1		4	
	<b>unclassified_Alveolata</b>	533	16	527	8
Amoebozoa	Discosea			2	
	Tubulinea	4		1	
Cryptophyta	Cryptomonadales	4		9	
	Cryptophyta_3			1	
	Cryptophyta_4	4		6	
	Pyrenomonadales	2		4	
	unclassified_Cryptophyta			4	
Haptophyceae	Coccosphaerales	2		2	
	Pavloales			1	
	Phaeocystales			5	
	Prymnesiales			2	
	unclassified_Haptophyceae	11	2	14	
	Zygodiscales			1	
Heterolobosea	Schizopyrenida	4		2	
Nucleariidae and Fonticula group	Nucleariidae	1		1	
Opisthokonta	Choanoflagellida	8		12	
	<b>Fungi</b>	954	22	1990	50
	Opisthokonta_incertae_sedis	1		1	
Rhizaria	<b>Cercozoa</b>	251	5	362	14
	<b>unclassified_Rhizaria</b>	93	2	123	6
Rhodophyta	Florideophyceae			1	
Stramenopiles	<b>Bacillariophyta</b>	82		150	1
	Bicosoecida	20		68	2
	Bolidophyceae			1	
	<b>Chrysophyceae</b>	94	2	125	1
	Dictyochophyceae	2		4	
	Eustigmatophyceae	12		39	
	Labyrinthulomycetes	11		13	
	MAST	13	1	26	2
	Oomycetes	38		71	1
	Pelagophyceae			2	
	Phaeothamniophyceae			1	
	PX_clade	2		4	
	Raphidophyceae	3		1	
	Synurophyceae	32		49	3
	<b>unclassified_Stramenopiles</b>	97	1	162	3
	Viridiplantae	<b>Chlorophyta</b>	226	4	607
Streptophyta		28		58	
unclassified_Viridiplantae		4		9	1
Total	4044	79	6515	149	

### **V.7.3. Focus on the temporal dynamics of rare OTUs**

Oscillating taxa (i.e OTUs that exhibit large temporal oscillations between abundant and rare) were considered as rare during a period when, on average, their number of DNA sequence was inferior or equal to 0.1 %.

In both lakes, the majority of OTUs never reached more than 30 DNA sequences (0.1 % of the number of DNA sequences per sample). These taxa, named here permanently rares accounted for 3571 OTUs (88 % of the total richness) in Lake Annecy and 5916 OTUs in Lake Bourget (90 % of the total richness). Some taxa from the rare biosphere displayed contrasted temporal dynamics. Approximately 2 % of the OTUs were considered as oscillating taxa, being abundant or rare for long time periods. Most of them were found within Fungi, Dinophyceae and unclassified Alveolata groups in both MECs.

In Lake Annecy, three time periods were defined based on bray-curtis dissimilarity: before the mid 1940s, from the mid 1940s to the late 1980s and post late 1980s. For Lake Bourget MECs, the samples were also discriminated in 3 time periods in: before the mid-1940s, from the mid-1940s to the mid-1990s and post-mid-1990s. These periods were characterized by shifts in abundant taxa as observed in the dynamics of oscillating OTUs. In Lake Bourget, a high number of taxa (76 OTUs) shifted from an abundant to a rare state in the 1940s (Figure V.7). In this lake, the main environmental change in the 1940s was the increase of phosphorus concentration; this forcing was showed to impact negatively the richness and the diversity of MECs (Figure V.3) and could explain the decrease in abundance of numerous OTUs (Figure V.7). The numbers of OTUs that shifted from a rare to an abundant state were found to be the highest in the late 1980s for Lake Annecy (36 OTUs) and in the mid 1990s for Lake Bourget (46 OTUs). This time period is characterized by the combined effects of nutrients concentration and air temperature changes on the MECs. The present analysis based on the rare biosphere confirmed the pattern revealed in V.4 about the emergence of numerous taxa at this time period.

Those additional analyses provide the first view about the temporal dynamics of the rare microbial eukaryotic biosphere on such time scale (1 century). They confirmed the strong effects of phosphorus concentration on the temporal dynamics of the rare biosphere within the Bourget MEC but also that air temperature increase can induce the emergence of microbial eukaryotes in terms of abundance.

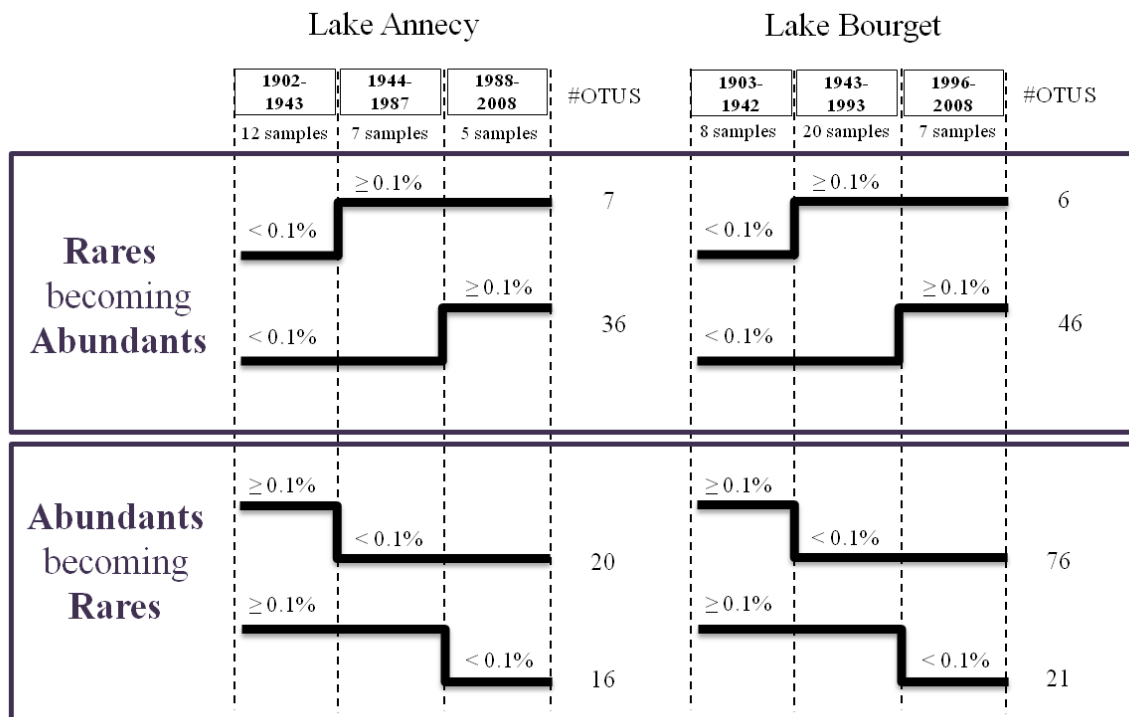


Figure V.7. Temporal dynamics of oscillating OTUs in Lake Annecy and Lake Bourget. Oscillating OTUs shifted between abundant ( $> 0.1\%$ ) and rare states ( $< 0.1\%$  of the number of DNA sequences, on average for each period). The selected time periods were defined using hierarchical clustering tree analysis and ANOSIM analysis (Oksanen *et al.* 2015).

## V.8. Supporting Information

Figure V.S1. Bray-Curtis analysis pseudo-autocorrelation. The plot showed the similarity values (1- Bray-Curtis dissimilarity values) related to the time lag between each pair of OTUs. MECs sampled more than 10 years apart were more dissimilar than those with lower time-lag in both lakes (ANOVA p-values < 0.001). The ANOVA results, also confirmed the higher dissimilarity between MECs with longer time lags (20/30/40/50 years; ANOVA p-values < 0.001).

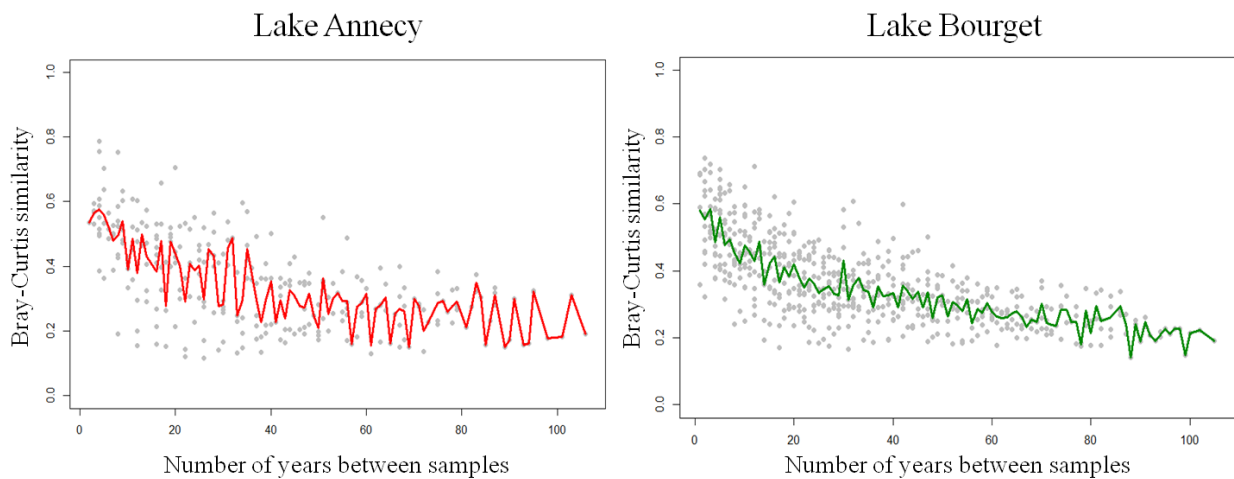
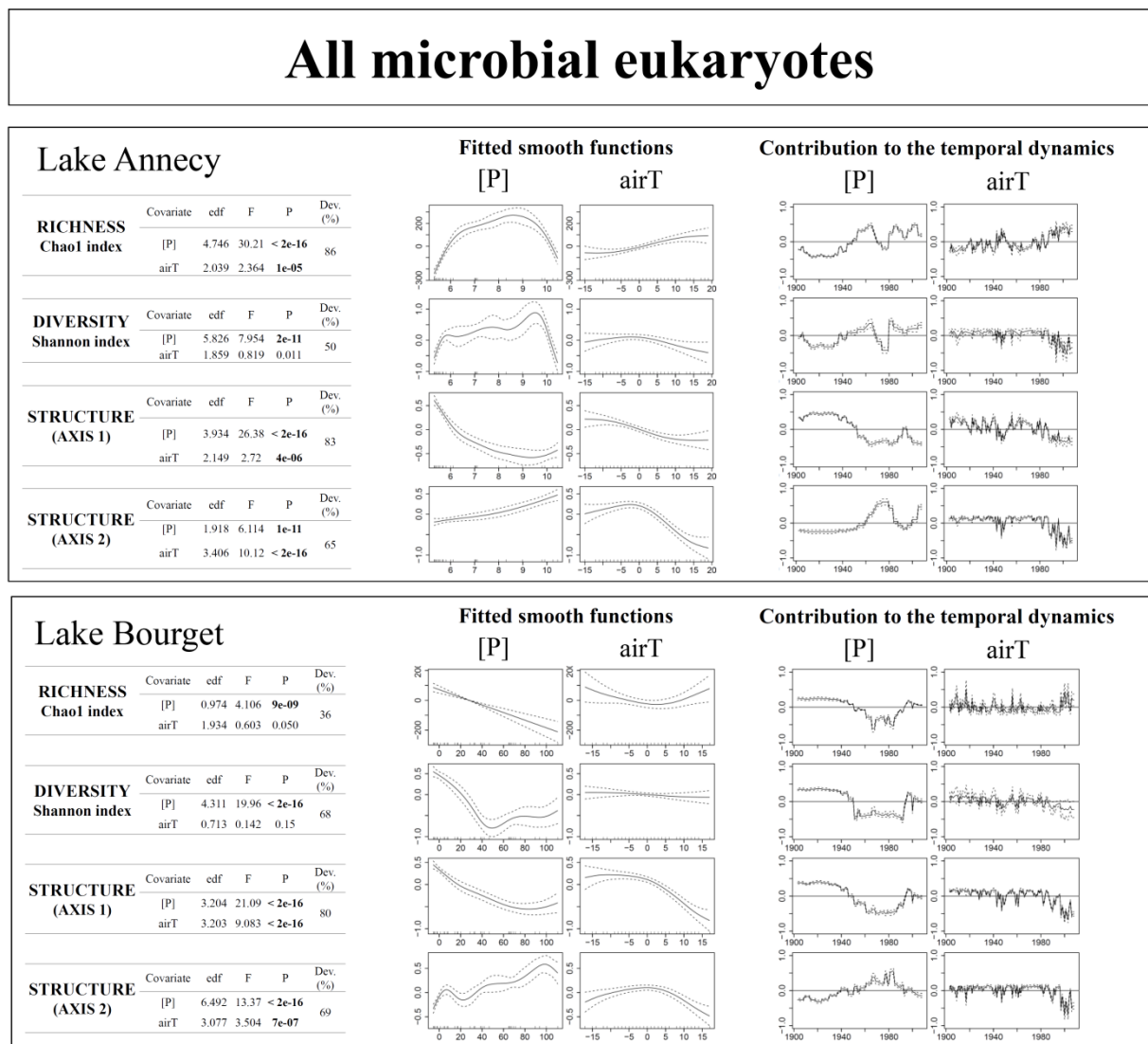
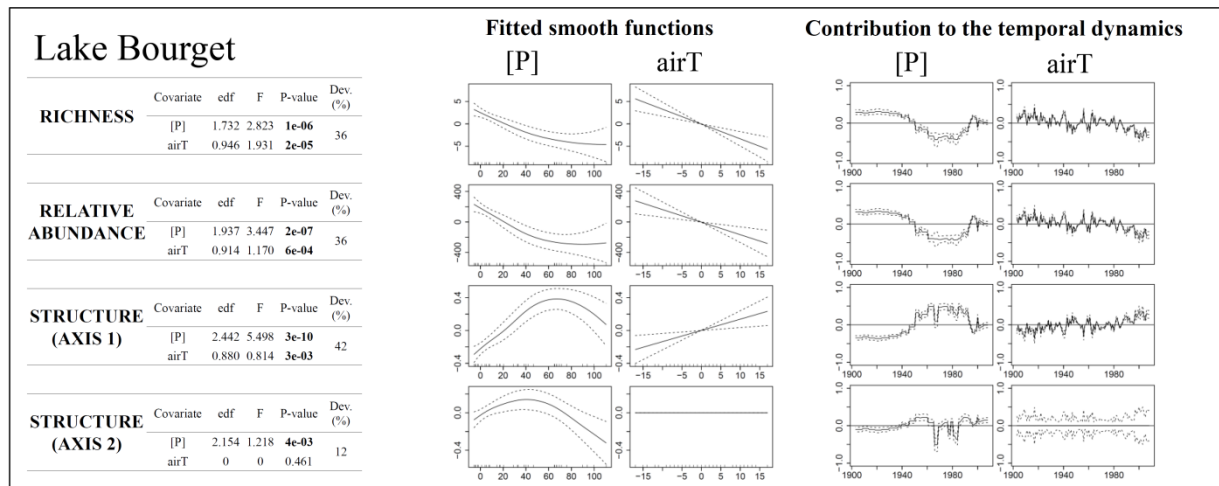
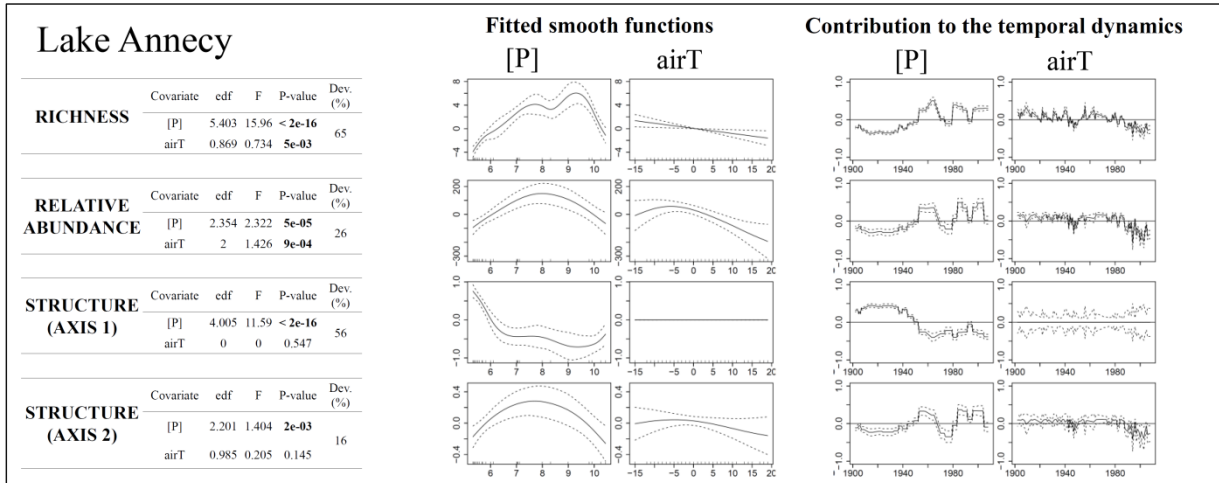
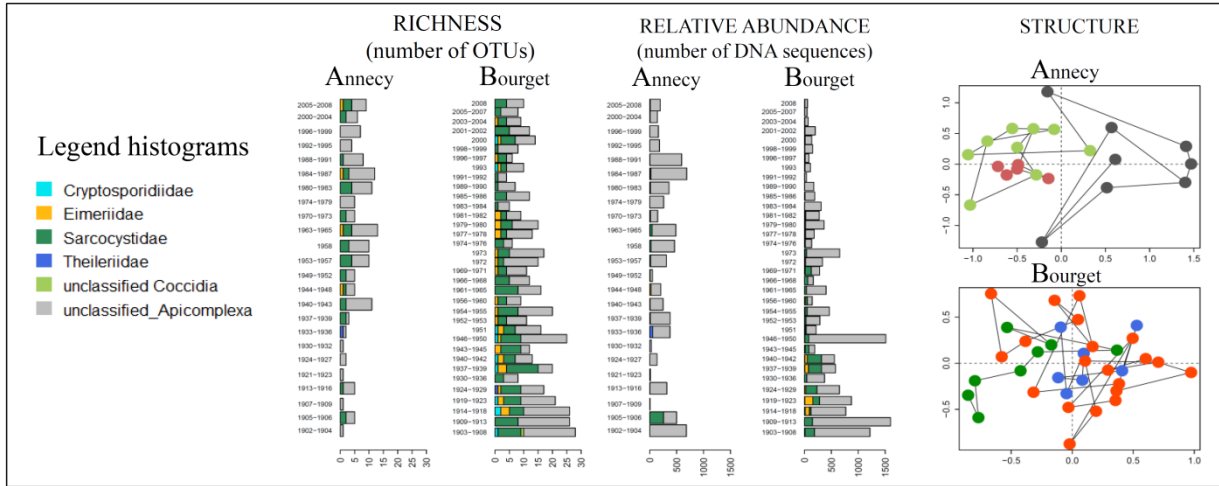


Figure V.S2. The figure presents GAMs analysis results for Lake Annecy and Lake Bourget for the 4 indices describing the whole MECs and the 9 most abundant groups. The results of GAMs analysis were reported as (i) the model summary (ii) the fitted smooth functions for [P] and airT (iii) the contribution of phosphorus concentration and air temperature. For Annecy, the colors within nMDS plots correspond to the following time periods: from 1902--1943 (red), from 1944--1987 (green), from 1988-2008 (gray). For Bourget, the colors within nMDS plots correspond to the following time periods: from 1903-1945 (green), from 1946-1992 (green), from 1993 to 2008 (blue).

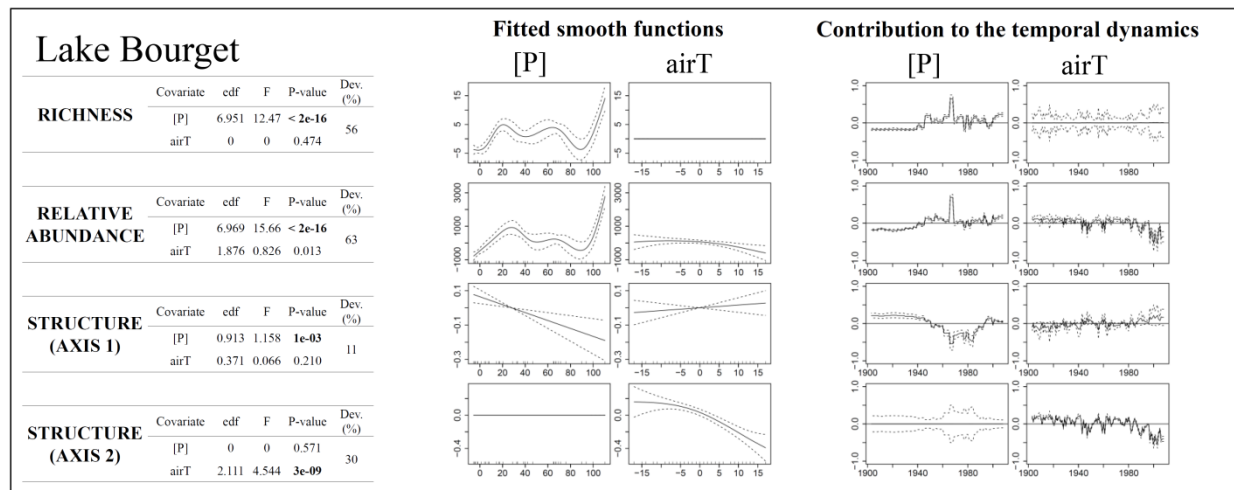
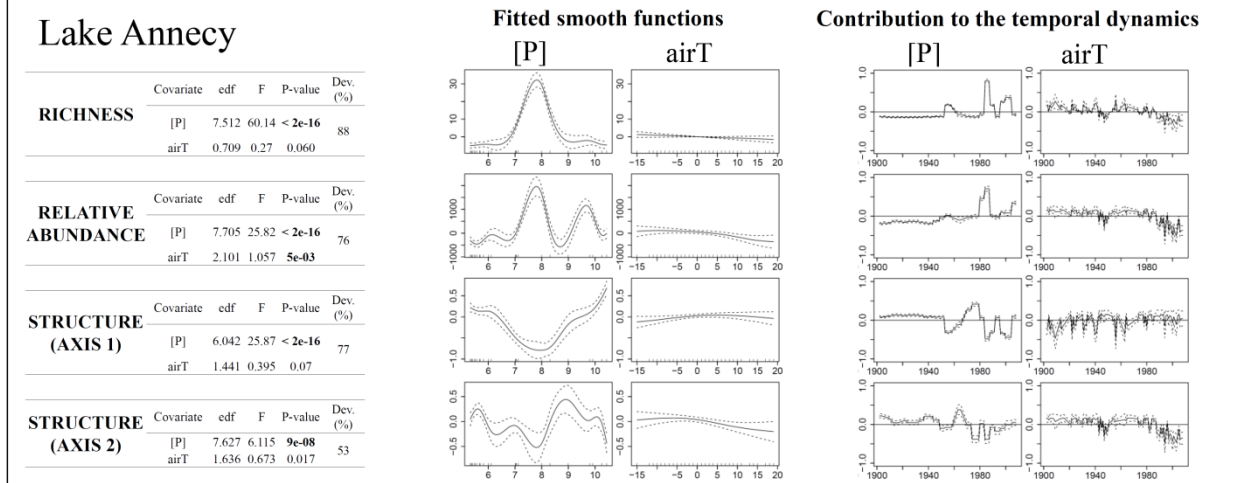
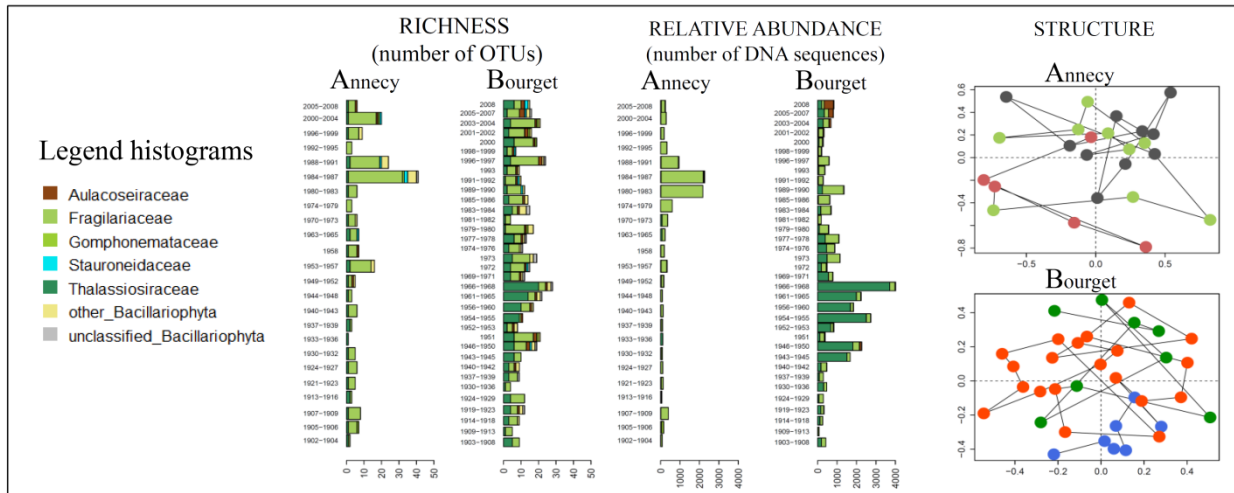


# Apicomplexa

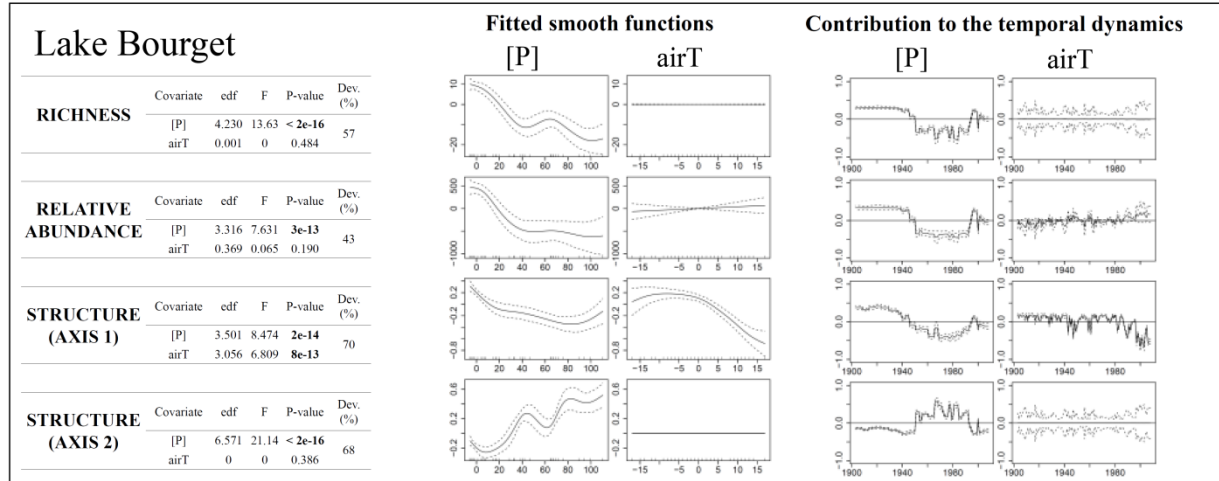
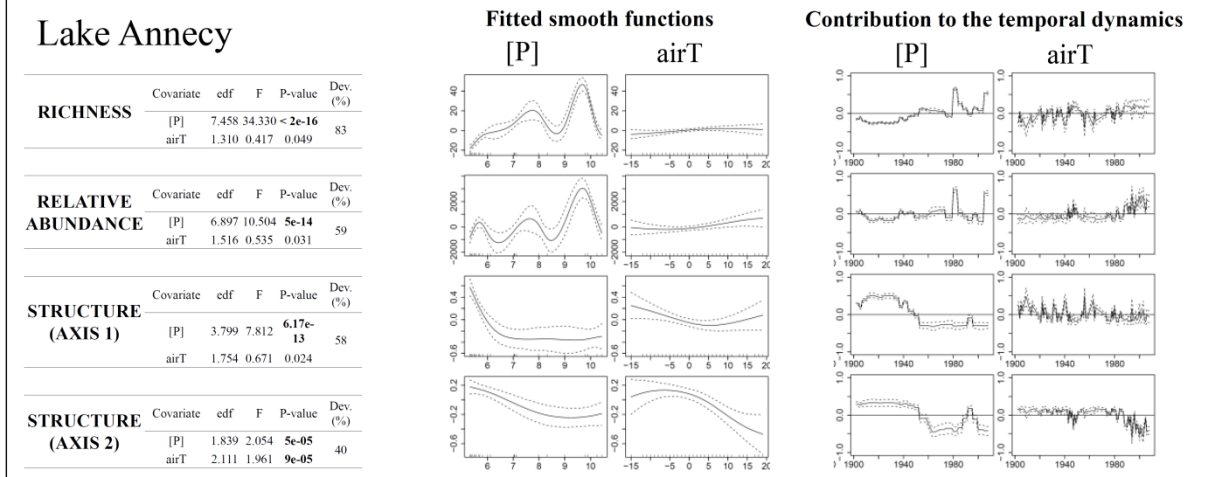
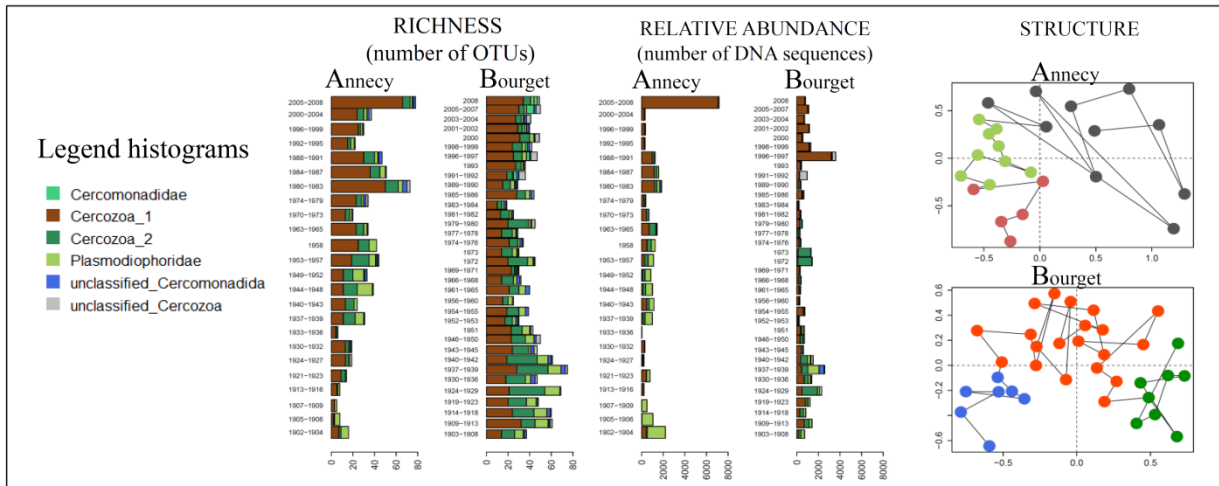




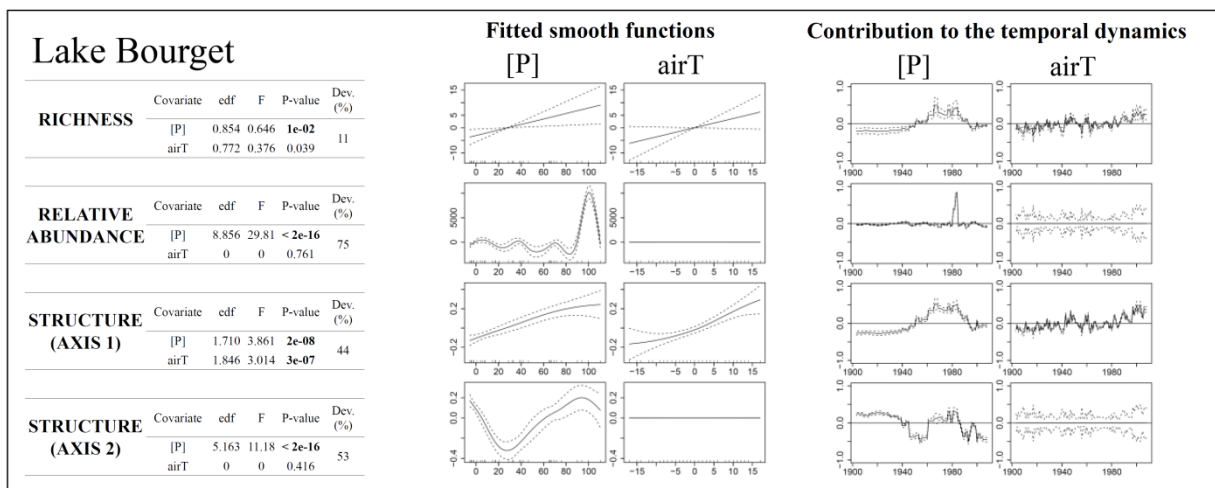
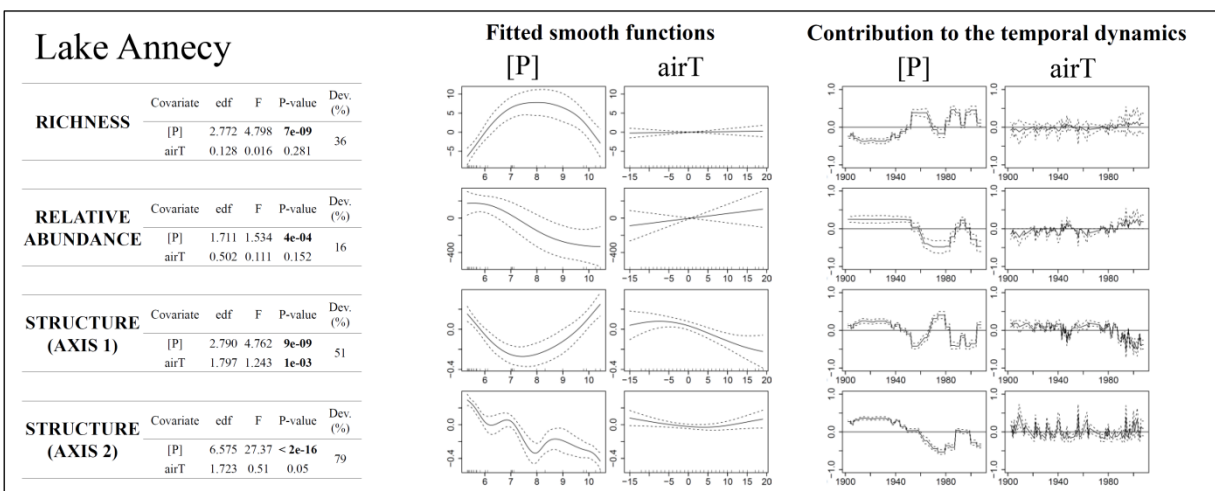
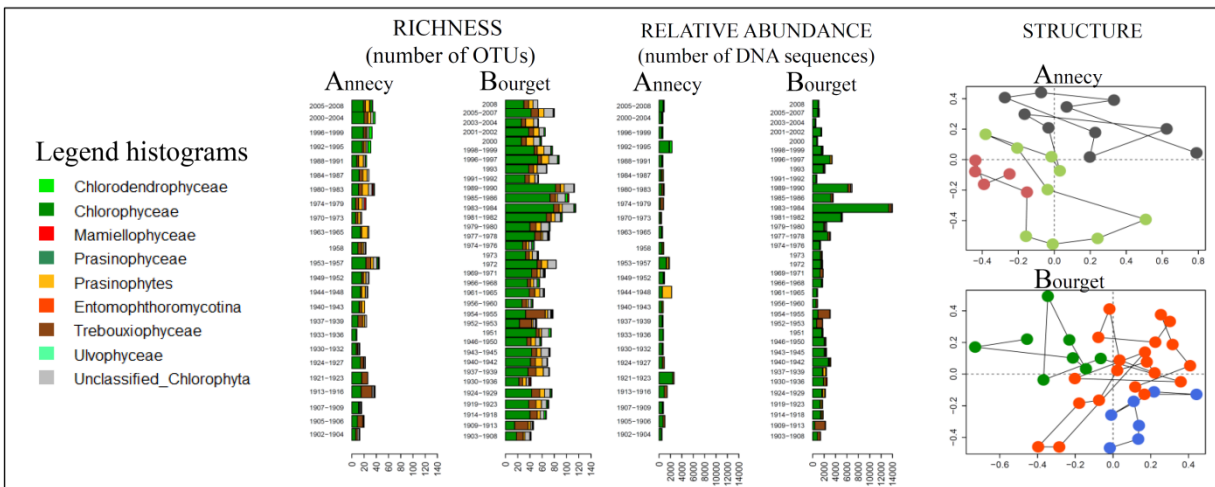
# Bacillariophyta



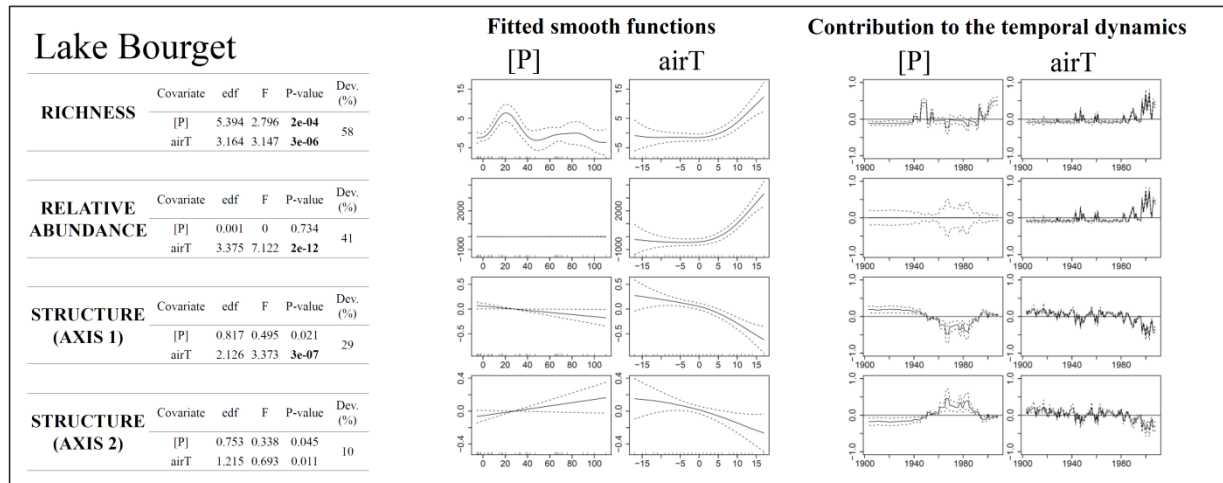
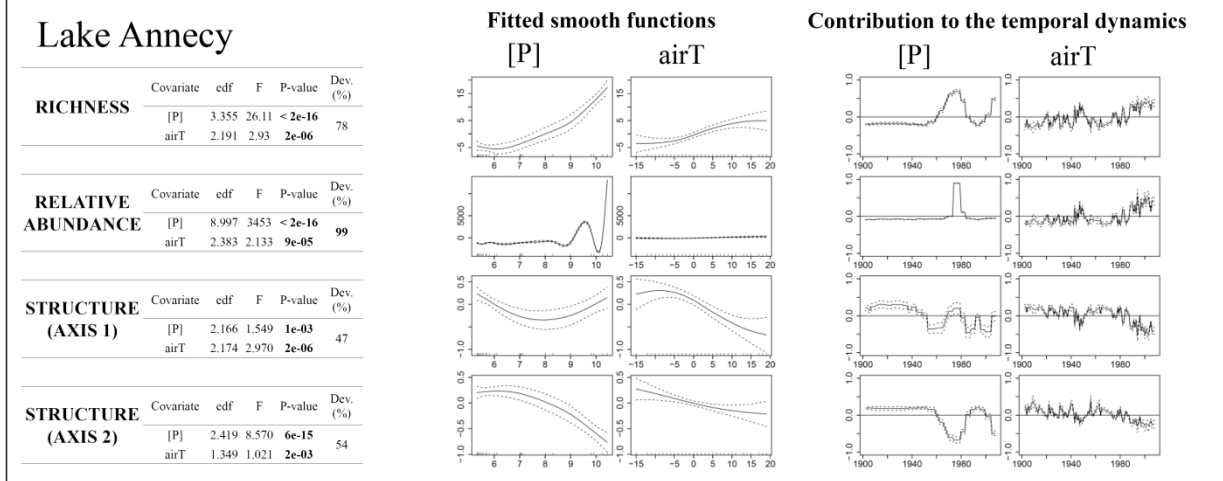
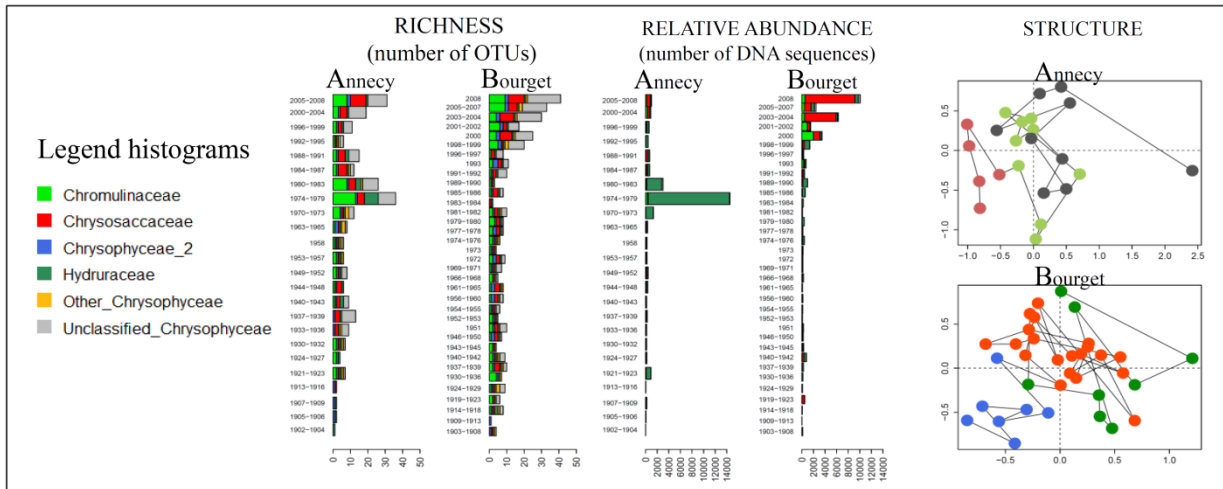
# Cercozoa



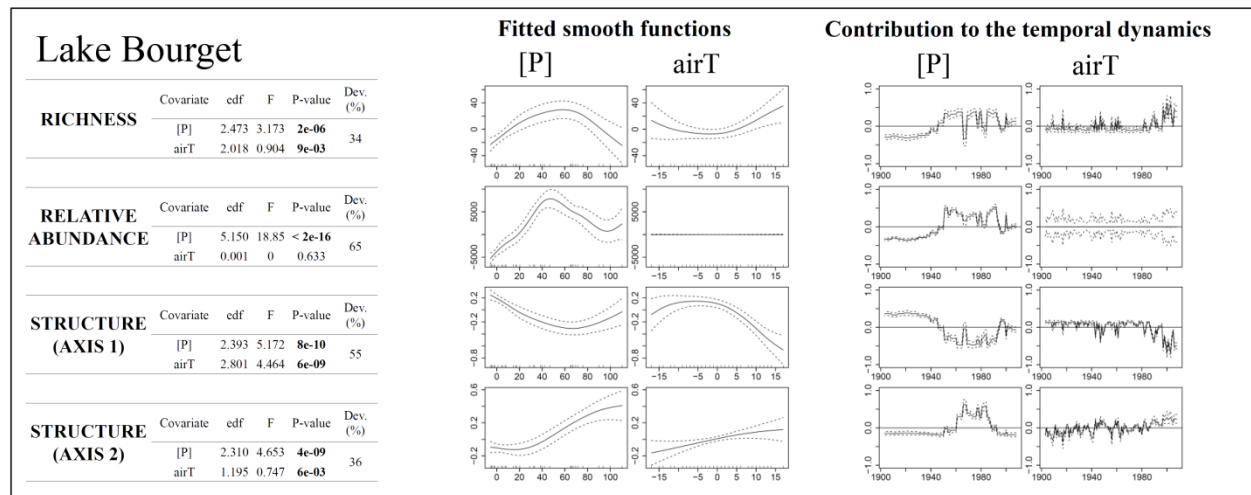
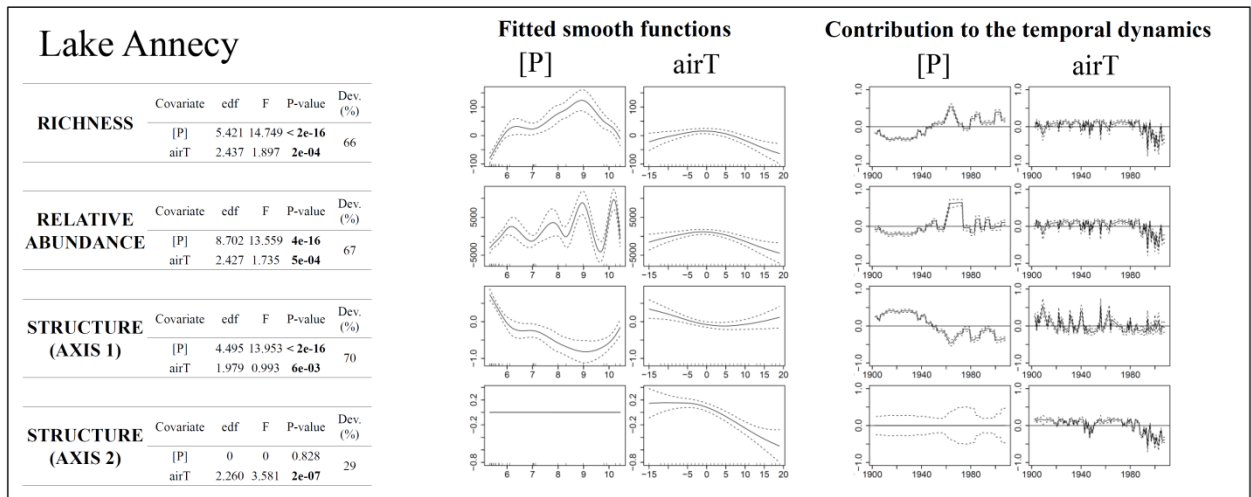
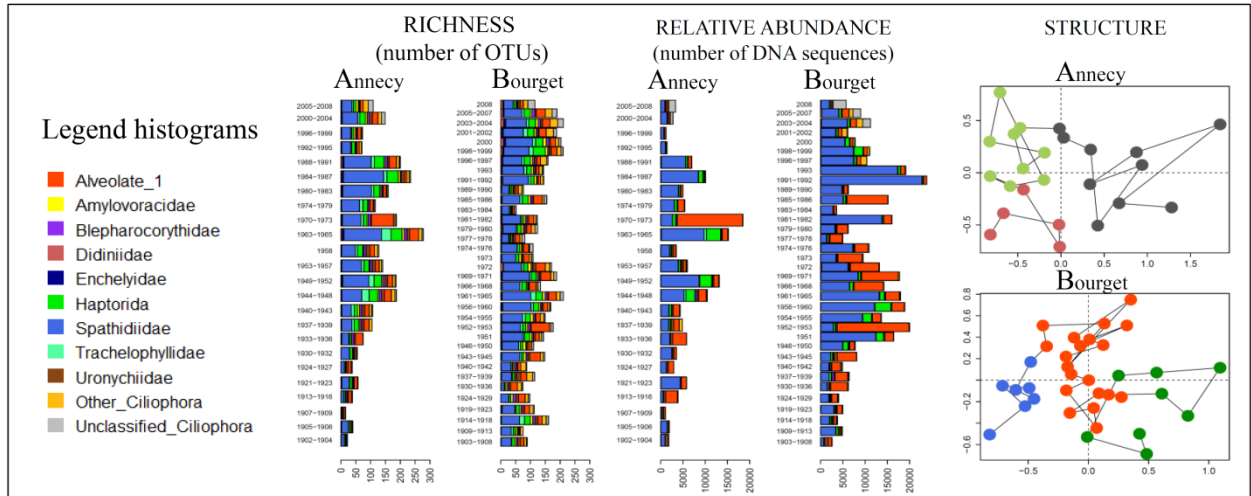
# Chlorophyta



# Chrysophyceae

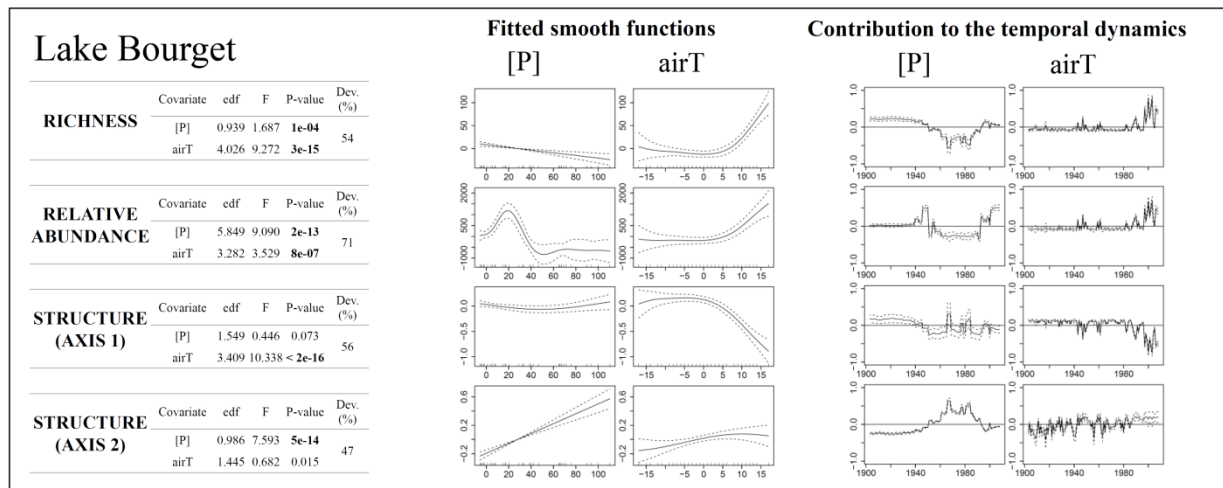
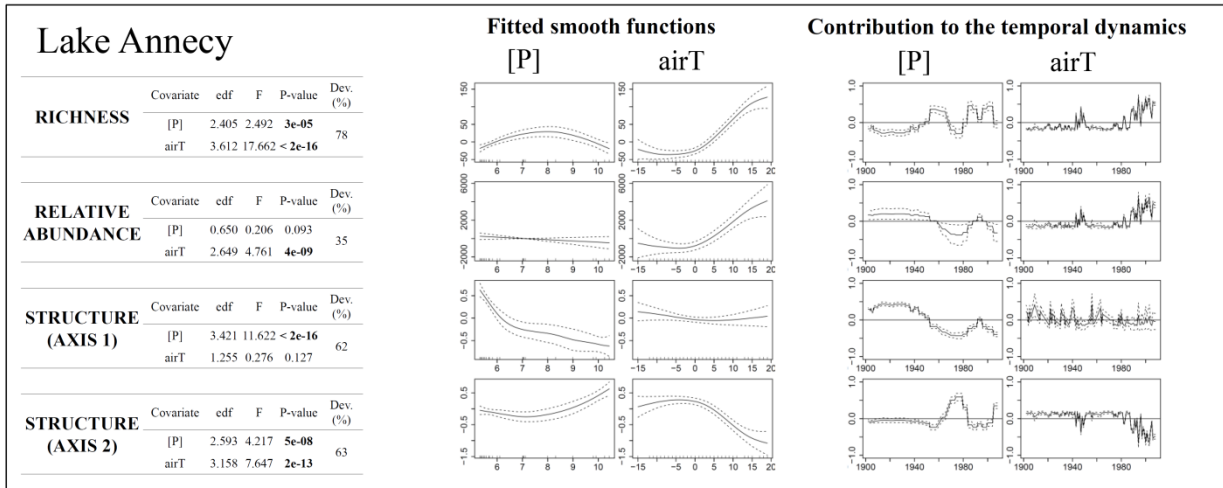
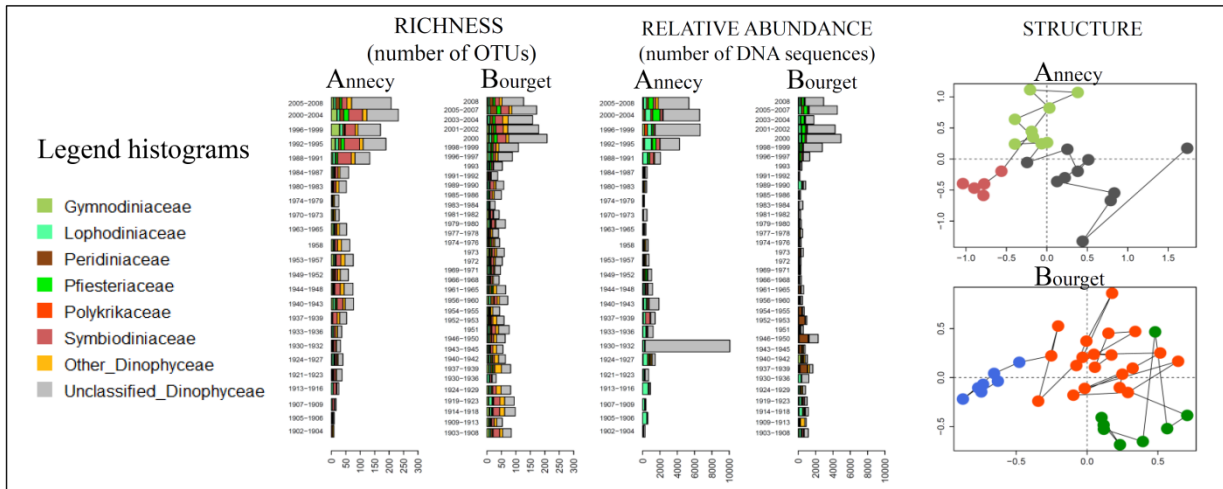


# Ciliophora

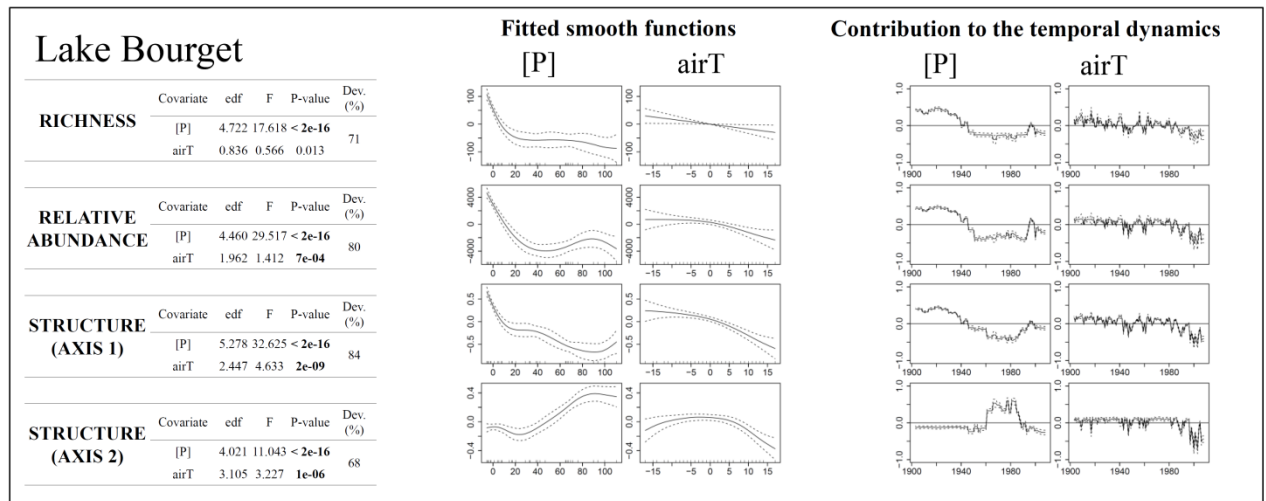
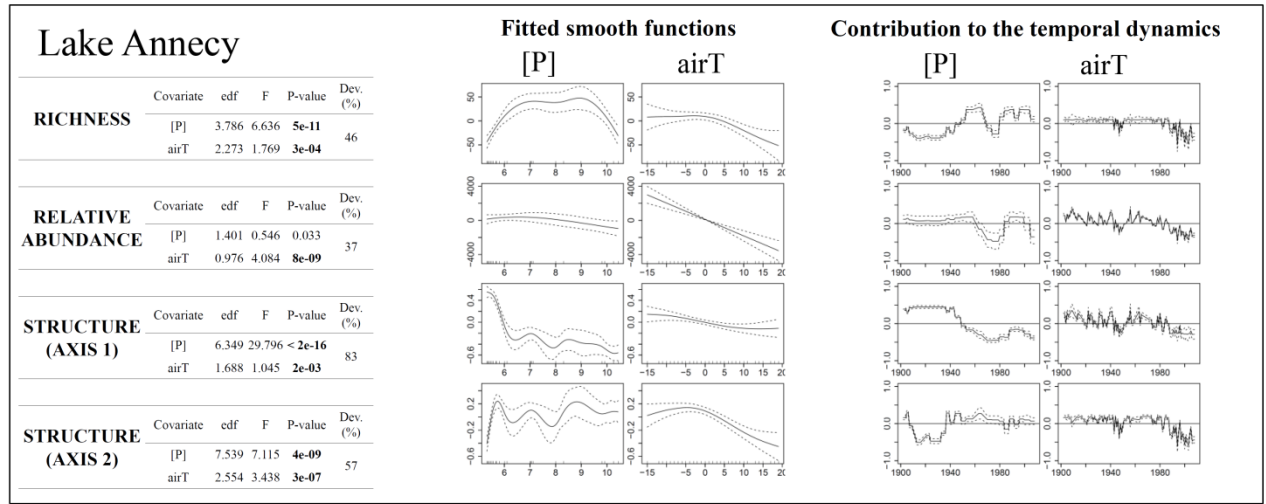
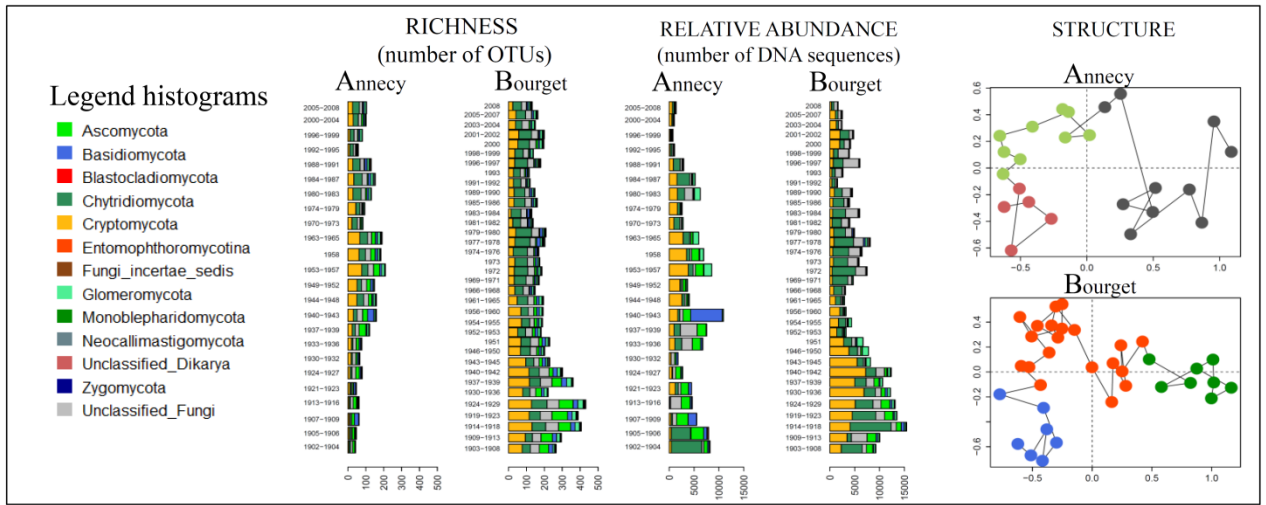




# Dinophyceae



# Fungi



# Perkinsea

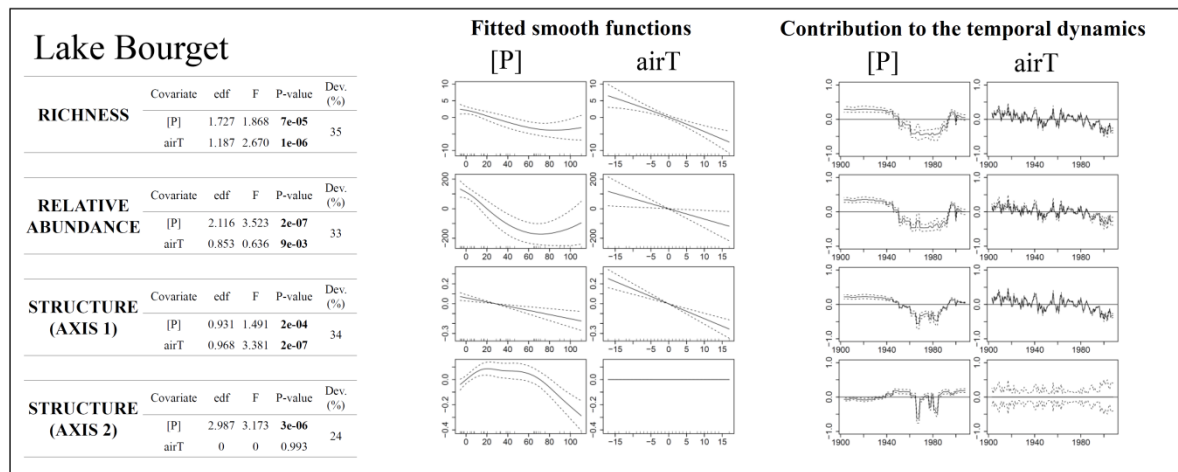
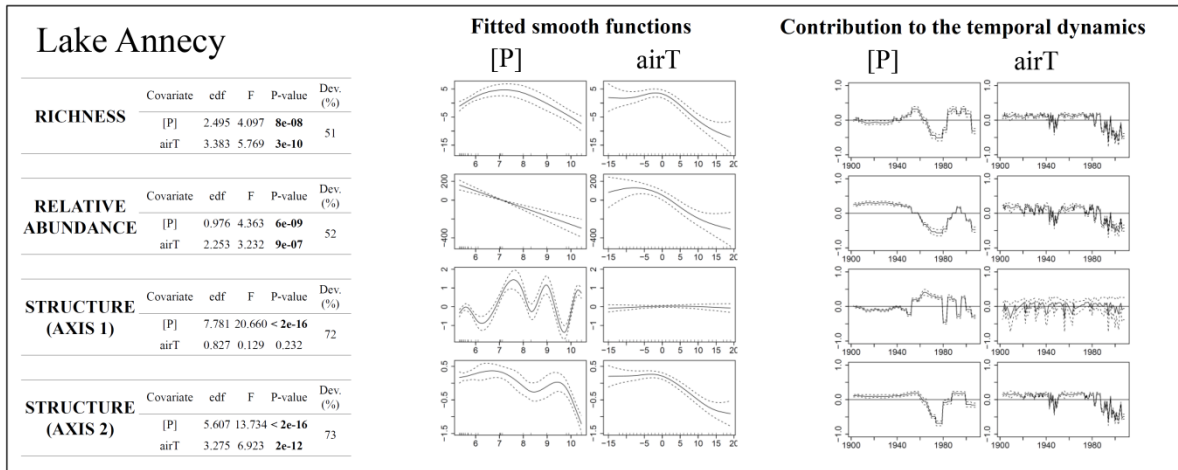
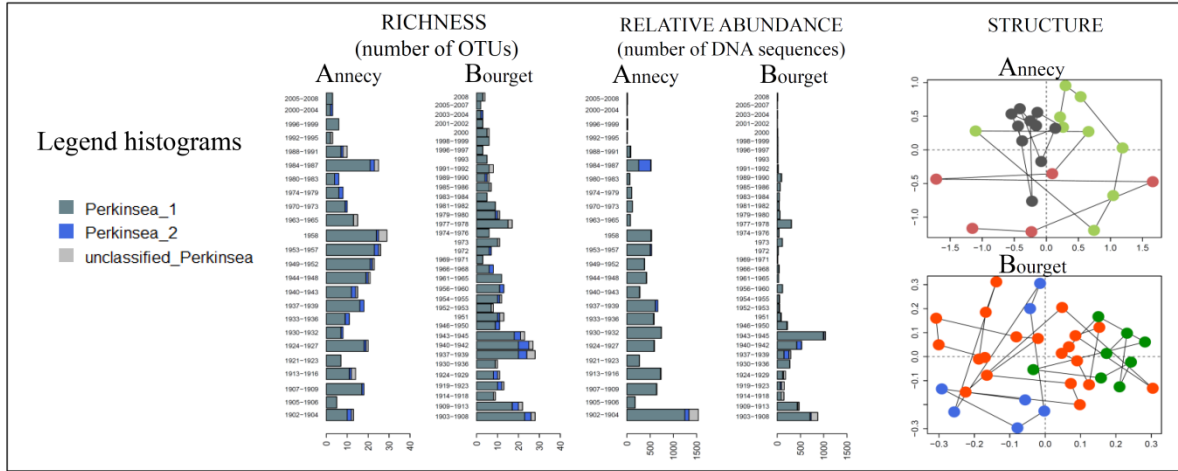




Figure V.S3. Topological roles of OTUs within the Anney and Bourget network (A & B). Plot of intra vs inter-modular connectivity values illustrating the proportion of connectors, module hubs, network hubs and peripherals in each network. Topological roles were attributed to each OTU based on their intra- and inter- connectivity values. The threshold was arbitrarily defined in the present study as: (i) peripherals (intra < 10, inter < 2), (ii) connectors (intra < 10, inter > 2), (iii) modules hubs (intra > 10, inter < 2) and (iv) network hubs (intra > 10, inter > 2).

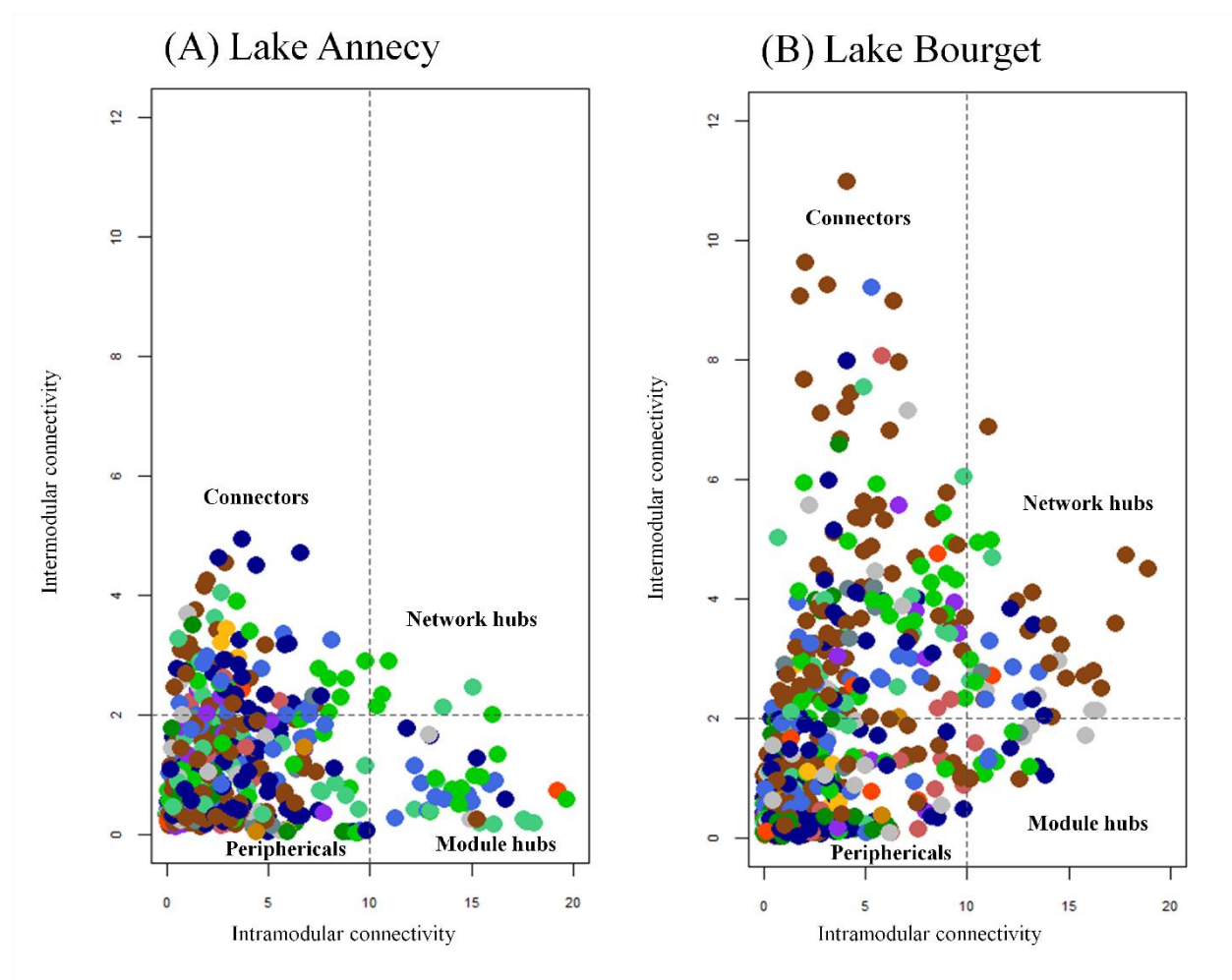


Figure V.S4. This figure illustrates the correlations between modules eigengene values and the environmental factors (phosphorus concentration and air temperature) in Lake Annecy and Lake Bourget. Red color means highly positive correlation and blue color highly negative correlation. The correlation coefficient (r) and significance (p) in parentheses are displayed within each plot. [P] corresponds to phosphorus concentration and airT corresponds to air temperature.

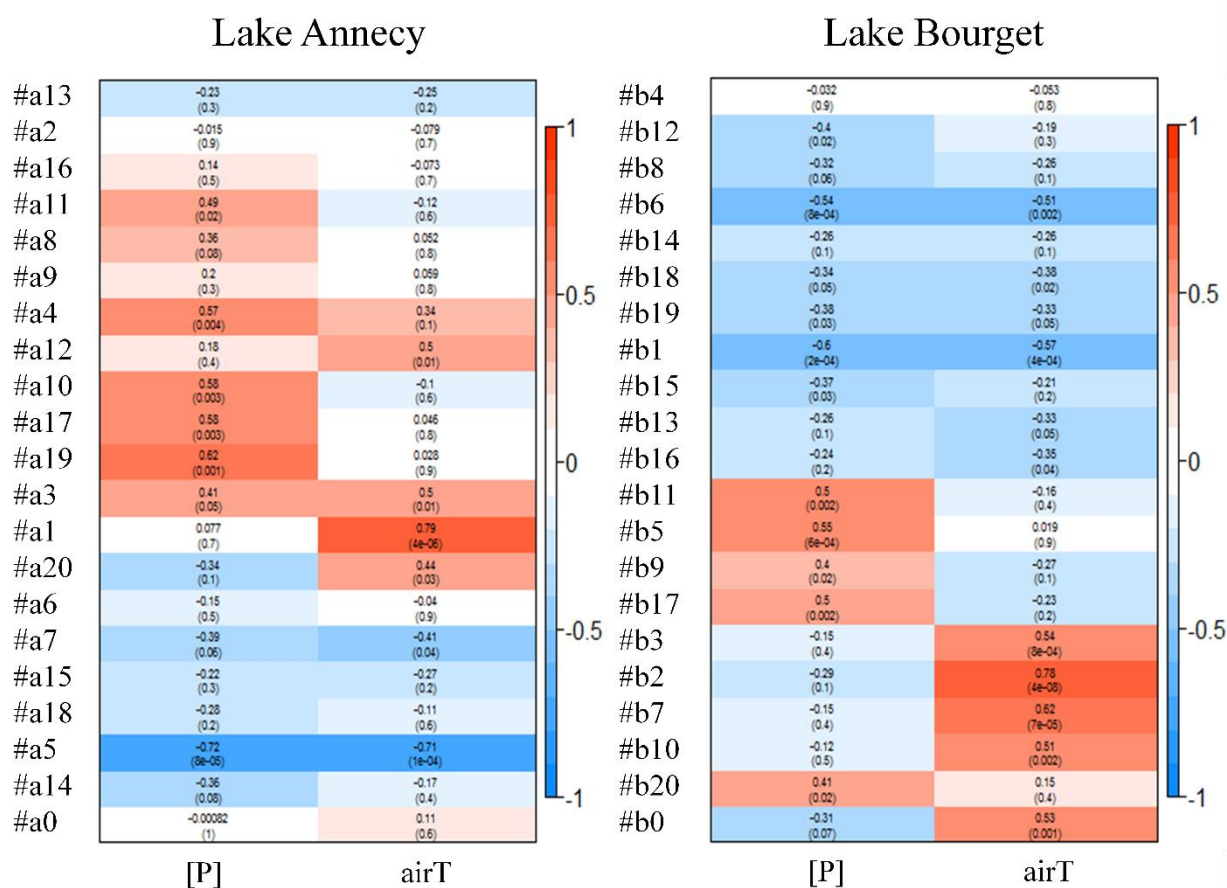
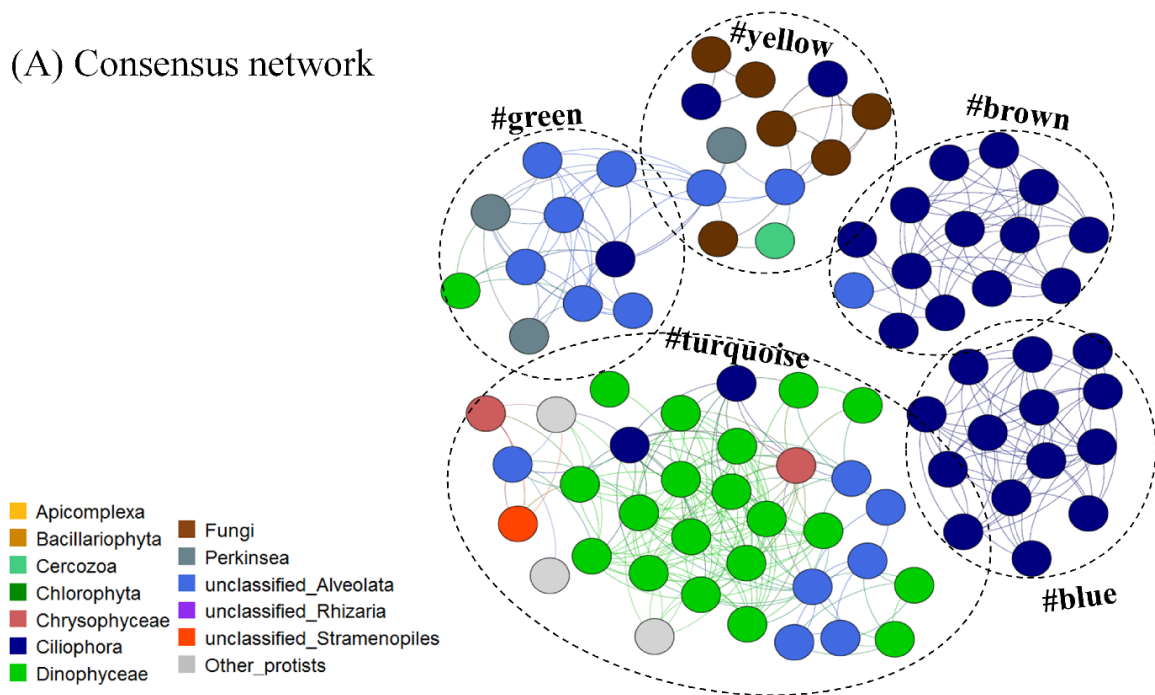
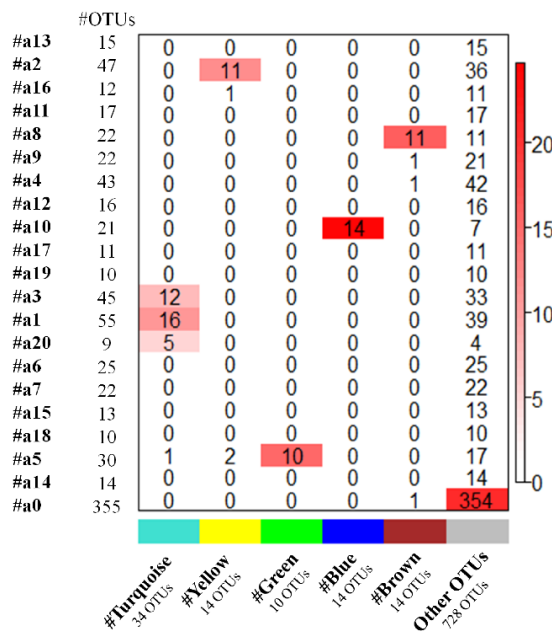


Figure V.S5. Consensus network. (A) Only the OTUs retrieved within the 5 consensus modules were represented. Correspondence of the (B) Annecy set-specific and (C) Bourget set-specific with the consensus modules. Each row of the table corresponds to one module from the individual networks, and each column corresponds to one consensus module. Numbers in the table indicate the number of OTUs in the intersection of the corresponding modules. Coloring of the table encodes  $-\log(p)$ , with  $p$  being the Fisher's exact test  $p$ -value for the overlap of the two modules. The stronger the red color, the more significant the overlap is.

(A) Consensus network



(B) Correspondence with Annecy network



(C) Correspondence with Bourget network

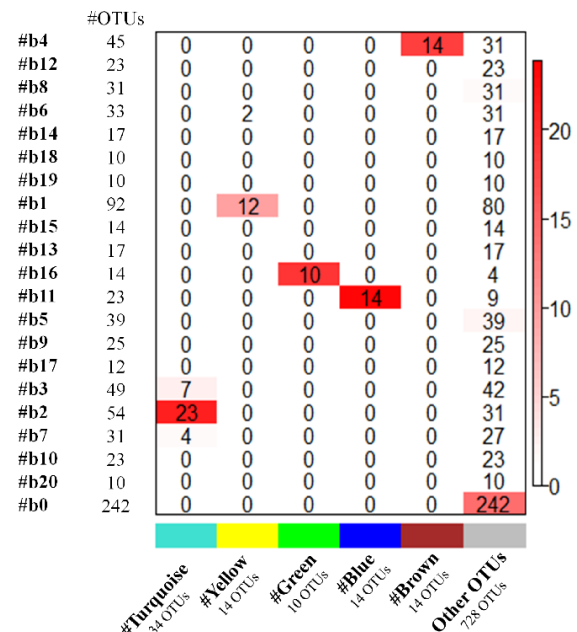


Table V.S1. Richness, diversity and evenness indices obtained for each sample.

Lake	Sample	Composite depth (mm)	Dating / Time (AD)	# all OTUs	Chao1 richness estimator	Shannon diversity index	Simpson diversity index	Pielou evenness index
Bourget	B1	0-12	2008	682	777	3.67	0.89	0.56
Bourget	B2	12-17	2005-2007	904	976	4.42	0.95	0.65
Bourget	B3	17-25	2003-2004	864	1013	3.86	0.92	0.57
Bourget	B4	25-35	2001-2002	898	1037	4.17	0.95	0.61
Bourget	B5	35-45	2000	956	1037	4.29	0.94	0.63
Bourget	B6	45-54	1998-1999	752	779	4.05	0.94	0.61
Bourget	B7	54-65	1996-1997	721	813	3.86	0.93	0.59
Bourget	B9	72-80	1993	537	586	2.77	0.70	0.44
Bourget	B10	80-90	1991-1992	516	601	1.91	0.48	0.31
Bourget	B11	90-100	1989-1990	551	560	3.96	0.94	0.63
Bourget	B13	105-112	1985-1986	697	734	3.79	0.90	0.58
Bourget	B14	112-120	1983-1984	444	458	3.59	0.91	0.59
Bourget	B15	121-129	1981-1982	551	598	3.05	0.80	0.48
Bourget	B16	129-138	1979-1980	729	761	3.96	0.95	0.60
Bourget	B17	138-150	1977-1978	594	624	4.08	0.96	0.64
Bourget	B18	150-157	1974-1976	541	554	3.58	0.91	0.57
Bourget	B19	158-165	1973	573	596	3.86	0.94	0.61
Bourget	B20	165-185	1972	693	788	3.58	0.91	0.55
Bourget	B21	185-198	1969-1971	606	649	3.44	0.89	0.54
Bourget	B22	198-213	1966-1968	555	568	3.68	0.92	0.58
Bourget	B23	213-228	1961-1965	747	857	3.44	0.84	0.52
Bourget	B24	228-240	1956-1960	651	720	3.51	0.88	0.54
Bourget	B25	240-245	1954-1955	642	662	3.88	0.92	0.60
Bourget	B26	245-255	1952-1953	608	706	3.11	0.78	0.48
Bourget	B27	255-265	1951	745	797	3.71	0.88	0.56
Bourget	B28	265-275	1946-1950	665	691	4.37	0.97	0.67
Bourget	B29	275-290	1943-1945	745	799	4.23	0.96	0.64
Bourget	B30	290-300	1940-1942	775	780	4.81	0.98	0.72
Bourget	B31	300-309	1937-1939	935	962	5.06	0.98	0.74
Bourget	B32	309-315	1930-1936	548	574	4.52	0.97	0.72
Bourget	B33	315-325	1924-1929	997	1028	5.04	0.98	0.73
Bourget	B34	325-335	1919-1923	939	1059	4.76	0.97	0.69
Bourget	B35	335-345	1914-1918	1035	1098	4.75	0.97	0.68
Bourget	B36	345-355	1909-1913	734	808	4.36	0.96	0.66
Bourget	B37	355-365	1903-1908	756	787	4.32	0.95	0.65
Annecy	A1	0-10	2005-2008	782	814	4.18	0.93	0.63
Annecy	A2	10-20	2000-2004	837	940	3.45	0.79	0.51
Annecy	A3	20-30	1996-1999	618	725	2.71	0.69	0.42
Annecy	A4	30-39	1992-1995	602	655	2.89	0.71	0.45
Annecy	A5	39-47	1988-1991	793	869	3.97	0.91	0.59
Annecy	A6	47-55	1984-1987	748	781	4.23	0.94	0.64
Annecy	A7	55-64	1980-1983	658	693	4.24	0.96	0.65
Annecy	A8	64-79	1974-1979	430	472	2.88	0.77	0.47
Annecy	A9	79-87	1970-1973	441	486	3.06	0.79	0.50
Annecy	A11	96-104	1963-1965	728	782	4.11	0.92	0.62
Annecy	A13	112-115	1958	660	691	3.85	0.88	0.59
Annecy	A14	115-125	1953-1957	744	744	4.51	0.96	0.68
Annecy	A15	125-135	1949-1952	634	650	3.87	0.92	0.60
Annecy	A16	135-145	1944-1948	676	692	4.19	0.95	0.64
Annecy	A17	145-155	1940-1943	564	573	3.93	0.93	0.62
Annecy	A18	155-165	1937-1939	516	551	3.98	0.94	0.64
Annecy	A19	165-175	1933-1936	308	320	3.38	0.91	0.59
Annecy	A20	175-183	1930-1932	276	285	2.63	0.83	0.47
Annecy	A22	190-200	1924-1927	330	332	2.92	0.80	0.50
Annecy	A23	200-210	1921-1923	272	272	3.33	0.90	0.59
Annecy	A25	220-230	1913-1916	289	293	3.20	0.89	0.56
Annecy	A27	240-248	1907-1909	254	286	2.83	0.84	0.51
Annecy	A28	248-255	1905-1906	222	248	3.03	0.90	0.56
Annecy	A29	255-262	1902-1904	187	188	2.87	0.86	0.55

Table V.S2. For each lake independently, the table presents for each OTU, the percentage of DNA sequences, the occurrence through time, the modularity features (intraM = intramodular connectivity, interM = intermodular connectivity), the topological role within network and the taxonomic affiliation.

(A) Lake Annecy

OTUs	% DNA sequences	% Occurrence	Modules	intraM	interM	Topological_role	Best_hit_identity	Identity	Taxonomic groups
OTU1027	0.2	96	a8	1.95	1.11	peripheral	<i>Semispathidium sp.</i>	98.5	Ciliophora
OTU10347	0.03	8	a0	1.04	0.5	peripheral	<i>Glomus sp.</i>	87.6	Fungi
OTU10796	0.01	13	a14	2.41	0.43	peripheral	Eurotiales	99.6	Fungi
OTU1094	0.01	8	a14	1.99	0.18	peripheral	Tricholomataceae	86	Fungi
OTU10949	0.33	75	a3	1.24	2.23	connector	<i>Chrysosaccus sp.</i>	95.9	Chrysophyceae
OTU11204	0.02	8	a18	2.86	0.54	peripheral	unclassified Fungi	99.2	Fungi
OTU11259	0.01	8	a3	13.00	1.65	module_hub	unclassified Ciliophora	97.3	Ciliophora
OTU11365	0.02	29	a2	3.74	1.16	peripheral	unclassified Alveolata	94.7	unclassified Alveolata
OTU11867	0.02	33	a14	0.41	0.81	peripheral	unclassified Fungi	96.6	Fungi
OTU12713	0.03	58	a2	0.8	0.74	peripheral	<i>Geosiphon sp.</i>	89.5	Fungi
OTU12745	0.02	21	a1	14.09	0.74	module_hub	unclassified Dinophyceae	91.4	Dinophyceae
OTU12882	0.09	54	a8	1.65	1.33	peripheral	Cryptomycota	92.8	Fungi
OTU12893	0.14	71	a1	3.3	1.81	peripheral	unclassified Dinophyceae	95.5	Dinophyceae
OTU13054	0.01	13	a3	12.23	0.41	module_hub	Cercozoa 1	97	Cercozoa
OTU1320	0.11	71	a3	0.38	0.86	peripheral	unclassified Dinophyceae	94.7	Dinophyceae
OTU13432	0.04	54	a5	1.65	1.01	peripheral	unclassified Dinophyceae	93.2	Dinophyceae
OTU13557	0.22	25	a15	1.49	0.5	peripheral	<i>Plasmodiophora sp.</i>	87.5	Cercozoa
OTU13708	0.03	63	a4	1.53	2.41	connector	unclassified Rhizaria	92.3	unclassified Rhizaria
OTU13974	0.01	4	a3	18.11	0.2	module_hub	Hypocreales	85	Fungi
OTU14370	0.02	50	a8	3.69	4.94	connector	<i>Trachelophyllum sp.</i>	97	Ciliophora
OTU14527	0.02	13	a19	4.67	1.85	peripheral	<i>Ochromonas sp.</i>	97.4	Chrysophyceae
OTU14707	0.04	67	a9	3.43	2.17	connector	Haptorida	97.7	Ciliophora
OTU1474	0.02	38	a9	3.61	3.24	connector	<i>Protospathidium sp.</i>	98.5	Ciliophora
OTU14743	0.01	29	a9	4.46	1.86	peripheral	<i>Enchelyodon sp.</i>	98.1	Ciliophora
OTU14970	0.01	50	a10	3.42	0.47	peripheral	<i>Blepharocorys sp.</i>	92.4	Ciliophora
OTU15351	0.07	17	a15	2.53	0.29	peripheral	unclassified Fungi	97.4	Fungi
OTU15587	0.01	4	a18	3.63	0.37	peripheral	<i>Theileria sp.</i>	71.6	Apicomplexa
OTU15648	0.03	33	a1	4.07	3.4	connector	unclassified Dinophyceae	95.5	Dinophyceae
OTU15803	1.56	100	a7	1.09	3.18	connector	Leotiomycetes	99.6	Fungi
OTU15881	0.79	83	a1	0.51	0.37	peripheral	Chlamydomonadaceae	99.6	Chlorophyta
OTU16585	0.01	17	a1	16.15	0.9	module_hub	unclassified Alveolata	98.5	unclassified Alveolata
OTU16591	0.01	38	a9	5.55	1.66	peripheral	Haptorida	96.6	Ciliophora
OTU17065	0.01	25	a12	2.00	1.5	peripheral	unclassified Fungi	98.5	Fungi
OTU17098	0.03	13	a7	5.69	0.05	peripheral	Dikarya	99.6	Fungi
OTU17118	0.03	58	a9	5.54	1.77	peripheral	<i>Trachelophyllum sp.</i>	97	Ciliophora
OTU17212	0.01	29	a16	0.88	1.89	peripheral	unclassified Alveolata	94.3	unclassified Alveolata
OTU17329	0.05	96	a8	1.28	0.97	peripheral	<i>Semispathidium sp.</i>	98.1	Ciliophora
OTU17407	0.01	8	a11	2.03	0.9	peripheral	Cercozoa 1	94	Cercozoa
OTU17593	0.5	42	a3	17.7	0.2	module_hub	Cercozoa 1	93.7	Cercozoa
OTU17762	0.45	79	a12	3.4	2.45	connector	unclassified Rhizaria	91.4	unclassified Rhizaria
OTU18016	0.12	54	a2	2.23	0.57	peripheral	Cryptomycota	93.9	Fungi
OTU18296	0.09	75	a8	0.71	3.08	connector	Cryptomycota CM1	97.7	Fungi
OTU18307	0.02	17	a1	16.28	1.33	module_hub	<i>Gymnodinium sp.</i>	98.1	Dinophyceae
OTU1844	0.25	79	a2	1.72	1.58	peripheral	Cryptomycota	94.7	Fungi
OTU18581	0.00	8	a3	15.39	0.18	module_hub	Cercozoa 1	97.7	Cercozoa
OTU18725	0.02	33	a2	3.19	1.15	peripheral	Cercozoa 2	92.5	Cercozoa
OTU18889	0.02	13	a10	5.83	0.63	peripheral	unclassified Fungi	88.8	Fungi
OTU19134	0.03	38	a2	0.86	0.95	peripheral	Alveolate 1	93.1	Ciliophora
OTU19247	0.02	38	a4	2.12	1.66	peripheral	MAST-7C	90.2	MAST-7
OTU1931	0.02	38	a13	0.07	0.42	peripheral	unclassified Alveolata	94.4	unclassified Alveolata
OTU19423	0.07	54	a4	3.49	2.96	connector	unclassified Apicomplexa	98.1	Apicomplexa
OTU19456	0.04	58	a9	6.2	2.13	connector	<i>Trachelophyllum sp.</i>	97	Ciliophora
OTU19523	0.02	58	a4	2.15	1.02	peripheral	<i>Enchelys sp.</i>	88.9	Ciliophora
OTU19758	0.01	13	a3	16.14	0.17	module_hub	Cercozoa 1	96.6	Cercozoa
OTU19838	0.01	42	a9	0.72	0.56	peripheral	unclassified Dinophyceae	90.6	Dinophyceae
OTU20074	0.02	38	a13	0.03	0.33	peripheral	<i>Chlamydomonas sp.</i>	99.6	Chlorophyta
OTU2018	0.06	58	a11	2.97	2.93	connector	Cryptomycota LKM11	95.1	Fungi
OTU20216	0.00	8	a19	2.68	0.15	peripheral	<i>Chrysonobula sp.</i>	94.1	Chrysophyceae
OTU20786	0.03	33	a2	0.38	0.77	peripheral	<i>Condylostoma sp.</i>	85.2	Ciliophora
OTU20906	1.83	100	a1	2.66	2.03	connector	unclassified Alveolata	96.2	unclassified Alveolata
OTU20958	0.02	38	a2	2.57	2.04	connector	Cryptomycota	94	Fungi
OTU20981	0.01	21	a2	1.94	0.6	peripheral	<i>Metschnikovia sp.</i>	98.8	Fungi
OTU21318	0.02	42	a1	4.79	1.93	peripheral	unclassified Dinophyceae	97	Dinophyceae
OTU21450	0.52	42	a2	3.64	1.65	peripheral	<i>Metschnikovia sp.</i>	99.2	Fungi
OTU21844	0.05	63	a20	1.45	2.00	connector	unclassified Alveolata	98.1	unclassified Alveolata
OTU21978	0.01	25	a1	14.81	0.66	module_hub	unclassified Alveolata	98.1	unclassified Alveolata
OTU22140	0.03	25	a17	2.38	3.63	connector	<i>Metopion sp.</i>	92	Cercozoa
OTU22398	0.01	13	a6	0.59	0.91	peripheral	<i>Plasmodiophora sp.</i>	90.4	Cercozoa
OTU22929	0.03	21	a7	5.00	0.04	peripheral	<i>Diophrys sp.</i>	85.9	Ciliophora
OTU22949	0.02	8	a3	14.97	0.25	module_hub	<i>Amoeba sp.</i>	96.6	Tubulinea
OTU23235	0.01	33	a3	4.62	0.38	peripheral	Oocystaceae	98.9	Chlorophyta
OTU23383	0.1	67	a20	1.4	2.87	connector	<i>Woloszynskia sp.</i>	99.6	Dinophyceae
OTU23615	15.67	100	a1	5.77	3.34	connector	unclassified Alveolata	99.2	unclassified Alveolata
OTU2382	0.07	71	a11	3.74	1.64	peripheral	Cercozoa 2	95.9	Cercozoa
OTU2385	0.01	13	a1	12.82	0.4	module_hub	unclassified Alveolata	98.1	unclassified Alveolata
OTU24204	0.05	63	a1	2.56	1.76	peripheral	unclassified Alveolata	97.4	unclassified Alveolata
OTU24265	0.04	33	a4	2.23	1.36	peripheral	Chytridiomycota	98.1	Fungi
OTU24457	0.02	13	a15	5.03	0.19	peripheral	unclassified Bicosoecida	82.2	Bicosoecida
OTU24583	0.02	63	a10	2.76	0.83	peripheral	Alveolate 1	94.6	Ciliophora
OTU24991	0.04	50	a10	5.97	0.42	peripheral	Alveolate 1	96.9	Ciliophora
OTU25048	0.02	29	a17	2.19	1.13	peripheral	unclassified Chrysophyceae	99.3	Chrysophyceae
OTU25061	0.01	17	a2	1.15	0.21	peripheral	<i>Prorocentrum sp.</i>	89.8	Dinophyceae
OTU25130	0.04	58	a1	2.31	1.15	peripheral	unclassified Alveolata	98.5	unclassified Alveolata

OTU25184	0.5	71	a1	14.77	0.53	module_hub	unclassified Dinophyceae	99.6	Dinophyceae
OTU25267	0.03	17	a11	3.18	1.84	peripheral	unclassified Alveolata	95.8	unclassified Alveolata
OTU25317	0.44	83	a1	9.00	0.76	peripheral	unclassified Dinophyceae	98.1	Dinophyceae
OTU25493	0.11	75	a20	1.94	1.55	peripheral	unclassified Alveolata	98.1	unclassified Alveolata
OTU25613	0.01	17	a6	9.39	0.03	peripheral	<i>Rhizophlyctis sp.</i>	91.3	Fungi
OTU25904	0.01	13	a1	6.36	0.2	peripheral	Trebouxiophyceae	99.6	Chlorophyta
OTU26648	0.02	63	a4	2.68	1.5	peripheral	<i>Epispathidium sp.</i>	97.7	Ciliophora
OTU26841	0.04	8	a7	5.97	0.05	peripheral	Sarcocystidae	94.7	Apicomplexa
OTU26991	0.00	4	a14	3.29	0.38	peripheral	unclassified Viridiplantae	62.3	Viridiplantae
OTU27086	0.03	50	a1	1.54	0.46	peripheral	Tremellomycetes	94	Fungi
OTU2724	0.34	71	a1	8.01	2.04	connector	<i>Pfiesteria sp.</i>	97.4	Dinophyceae
OTU27369	0.02	21	a4	7.56	0.73	peripheral	Cercozoa 1	96.6	Cercozoa
OTU27550	0.02	8	a6	8.93	0.04	peripheral	<i>Tintinnophagus sp.</i>	91	Dinophyceae
OTU28251	0.05	54	a18	0.1	0.15	peripheral	unclassified Fungi	99.6	Fungi
OTU28409	0.09	79	a8	0.76	1.59	peripheral	<i>Gymnodinium sp.</i>	92.1	Dinophyceae
OTU28750	0.03	29	a12	0.41	1.38	peripheral	<i>Endogone sp.</i>	94	Fungi
OTU29058	0.02	54	a5	0.62	0.34	peripheral	unclassified Alveolata	95.5	unclassified Alveolata
OTU29061	0.01	29	a15	2.02	0.4	peripheral	<i>Procoentrum sp.</i>	92.1	Dinophyceae
OTU29114	0.01	8	a6	3.96	0.26	peripheral	Cryptomycota	87.5	Fungi
OTU29298	0.03	13	a15	4.01	0.24	peripheral	Mortierellaceae	94.4	Fungi
OTU29307	0.01	13	a14	0.93	0.33	peripheral	Chlorellales	96.6	Chlorophyta
OTU29344	0.01	25	a10	4.99	0.29	peripheral	Alveolate 1	96.9	Ciliophora
OTU29563	0.04	67	a5	3.68	0.9	peripheral	<i>Symbiodinium sp.</i>	91.7	Dinophyceae
OTU30001	0.01	17	a3	2.64	0.38	peripheral	unclassified Ciliophora	84.8	Ciliophora
OTU30168	0.05	67	a9	0.88	2.73	connector	<i>Amylavorax sp.</i>	89.4	Ciliophora
OTU30268	0.01	21	a11	1.96	1.59	peripheral	Cercozoa 1	94.1	Cercozoa
OTU30350	0.45	92	a4	2.22	1.86	peripheral	Cercozoa 1	94.4	Cercozoa
OTU30415	0.01	8	a3	12.92	1.66	module_hub	Saprolegniaceae	90.6	Oomycetes
OTU30593	0.02	21	a8	1.35	3.02	connector	Chytridiomycota	95.8	Fungi
OTU30682	0.23	29	a2	7.02	1.63	peripheral	unclassified Alveolata	88.3	unclassified Alveolata
OTU30855	0.02	33	a4	5.45	0.89	peripheral	unclassified Alveolata	94	unclassified Alveolata
OTU30941	0.09	92	a1	0.3	0.89	peripheral	Chrysophyceae 2	92.9	Chrysophyceae
OTU31161	0.1	79	a7	0.81	1.73	peripheral	<i>Gremmeniella sp.</i>	97.3	Fungi
OTU31306	0.05	42	a1	7.99	2.61	connector	<i>Peridinium sp.</i>	97	Dinophyceae
OTU31405	0.02	42	a5	0.2	0.26	peripheral	unclassified Alveolata	96.6	unclassified Alveolata
OTU31421	1.14	54	a7	1.21	0.6	peripheral	unclassified Haptophyceae	61.1	Haptophyceae
OTU31959	0.15	46	a11	2.89	4.55	connector	<i>Gigaspora sp.</i>	89.8	Fungi
OTU32009	0.05	8	a15	1.6	0.89	peripheral	<i>Papiliotrema sp.</i>	89.1	Fungi
OTU32153	0.02	54	a2	0.93	0.53	peripheral	unclassified Fungi	90.5	Fungi
OTU32171	0.02	25	a1	12.5	0.86	module_hub	unclassified Alveolata	97.4	unclassified Alveolata
OTU32294	0.01	13	a0	1.00	0.51	peripheral	Syndiniales	91.8	Dinophyceae
OTU32918	0.01	8	a3	19.24	0.73	module_hub	unclassified Stramenopiles	95.1	Stramenopiles
OTU32995	0.04	46	a3	3.19	2.41	connector	<i>Pfiesteria sp.</i>	97.7	Dinophyceae
OTU33040	0.02	8	a13	1.93	0.63	peripheral	Cercozoa 2	92.2	Cercozoa
OTU33160	0.42	38	a18	0.48	1.11	peripheral	unclassified Fungi	99.2	Fungi
OTU33321	0.04	58	a18	0.06	0.73	peripheral	<i>Sabulodinium sp.</i>	95.5	Dinophyceae
OTU3343	0.02	4	a7	5.92	0.05	peripheral	<i>Scutellospora sp.</i>	94.3	Fungi
OTU33464	0.13	38	a6	9.3	0.04	peripheral	<i>Lobulomyces sp.</i>	91.7	Fungi
OTU33538	0.04	42	a3	16.03	2.00	module_hub	<i>Blastodinium sp.</i>	96.6	Dinophyceae
OTU33730	0.01	29	a17	2.61	1.69	peripheral	unclassified Chrysophyceae	99.6	Chrysophyceae
OTU33934	0.01	13	a19	3.53	1.00	peripheral	<i>Mallomonas sp.</i>	94.8	Synurophyceae
OTU3406	0.03	38	a4	3.09	1.92	peripheral	Cryptomycota CM1	92.4	Fungi
OTU34105	0.12	50	a1	0.34	0.31	peripheral	Chaetophorales	94.7	Chlorophyta
OTU34287	0.05	42	a4	5.1	1.91	peripheral	unclassified Rhizaria	92.9	unclassified Rhizaria
OTU34608	0.01	4	a13	3.44	0.49	peripheral	<i>Hyphodontia sp.</i>	99.2	Fungi
OTU34616	0.02	33	a9	6.23	2.18	connector	Haptorida	97.3	Ciliophora
OTU34633	0.82	58	a17	0.05	0.22	peripheral	unclassified Stramenopiles	99.6	Stramenopiles
OTU34674	0.32	83	a9	0.06	0.54	peripheral	<i>Spathidium sp.</i>	95.8	Ciliophora
OTU35041	1.5	100	a2	2.44	2.81	connector	Cryptomycota CM1	99.2	Fungi
OTU35241	0.01	17	a1	11.24	0.26	module_hub	unclassified Alveolata	95.9	unclassified Alveolata
OTU35462	0.02	8	a15	4.98	0.18	peripheral	unclassified Bicosoecida	83.3	Bicosoecida
OTU35796	0.06	33	a2	4.83	3.16	connector	unclassified Fungi	85.8	Fungi
OTU36330	1.29	83	a13	1.66	0.17	peripheral	Agaricomycetes	99.2	Fungi
OTU36370	0.08	29	a6	9.19	0.04	peripheral	<i>Colpodidium sp.</i>	90.8	Ciliophora
OTU36614	3.07	92	a19	4.2	0.92	peripheral	<i>Chrysonebula sp.</i>	98.1	Chrysophyceae
OTU36769	2.92	100	a10	7.48	0.4	peripheral	Alveolate 1	97.7	Ciliophora
OTU37312	0.03	63	a2	0.72	0.89	peripheral	Agaricomycetes	87.9	Fungi
OTU37709	0.02	38	a5	2.14	1.3	peripheral	unclassified Fungi	100	Fungi
OTU37719	0.01	8	a3	11.82	1.77	module_hub	unclassified Ciliophora	93.1	Ciliophora
OTU37814	0.00	8	a6	8.69	0.04	peripheral	unclassified Chlorophyta	90.3	Chlorophyta
OTU38079	0.02	21	a7	0.97	0.4	peripheral	<i>Botryococcus sp.</i>	98.9	Chlorophyta
OTU38548	0.01	4	a16	1.94	0.6	peripheral	Chytridiomycota	90.2	Fungi
OTU38930	0.02	17	a2	1.75	0.76	peripheral	Cryptomycota LKM11	95.4	Fungi
OTU39067	0.43	58	a2	1.84	1.52	peripheral	<i>Paraglomus sp.</i>	89.1	Fungi
OTU3912	0.02	25	a12	1.98	1.72	peripheral	Chytridiomycota	92.9	Fungi
OTU3929	0.03	58	a2	2.94	1.41	peripheral	Cryptomycota CM1	98.5	Fungi
OTU39420	0.05	58	a3	1.34	1.19	peripheral	<i>Cosmarium sp.</i>	90.2	Streptophyta
OTU39480	0.01	21	a6	1.39	0.51	peripheral	Alveolate 1	96.2	Ciliophora
OTU39622	0.02	21	a6	4.5	0.23	peripheral	Cryptomycota	85.6	Fungi
OTU39693	0.09	92	a10	0.17	0.25	peripheral	Thalassiosiraceae	99.6	Bacillariophyta
OTU39804	0.01	4	a13	3.44	0.49	peripheral	<i>Trichosporon sp.</i>	99.6	Fungi
OTU39918	0.03	21	a1	3.43	0.98	peripheral	<i>Chrysosphaerella sp.</i>	97	Synurophyceae
OTU40292	0.01	29	a16	0.35	1.11	peripheral	unclassified Alveolata	94.7	unclassified Alveolata
OTU40294	0.01	13	a3	1.34	0.49	peripheral	unclassified Dinophyceae	96.6	Dinophyceae
OTU40557	0.02	33	a4	2.56	1.68	peripheral	unclassified Chlorophyta	98.1	Chlorophyta
OTU40776	0.01	13	a4	2.33	0.5	peripheral	<i>Epispathidium sp.</i>	94.3	Ciliophora
OTU4105	0.01	25	a7	1.66	0.12	peripheral	unclassified Fungi	98.9	Fungi
OTU41110	0.23	83	a9	0.59	3.27	connector	Cercozoa 2	96.3	Cercozoa
OTU41196	0.83	96	a8	0.76	1.72	peripheral	<i>Arcuospathidium sp.</i>	98.1	Ciliophora
OTU41271	0.19	54	a2	0.51	0.54	peripheral	<i>Plasmodiophora sp.</i>	87.5	Cercozoa
OTU41739	0.01	8	a12	4.48	0.37	peripheral	MAST-7C	89.6	MAST-7
OTU42187	0.08	92	a8	4.35	1.3	peripheral	<i>Semispithidium sp.</i>	98.1	Ciliophora
OTU42278	0.04	38	a4	6.34	1.91	peripheral	unclassified Fungi	97	Fungi
OTU42289	0.17	42	a1	15.05	0.53	module_hub	unclassified Alveolata	96.2	unclassified Alveolata
OTU42697	0.1	54	a1	5.57	2.23	connector	<i>Cyclonexis sp.</i>	95.1	Chrysophyceae
OTU43007	0.01	29	a10	4.93	0.18	peripheral	Alveolate 1	97.7	Ciliophora
OTU4316	0.06	67	a15	0.08	0.25	peripheral	unclassified Stramenopiles	97.7	Stramenopiles
OTU43307	0.24	21	a17	2.92	2.39	connector	<i>Naegleria sp.</i>	94.1	Schizopyrenida

OTU43430	0.01	8	a3	6.72	2.04	connector	unclassified Alveolata	97.3	unclassified Alveolata
OTU43775	0.01	13	a16	2.01	4.26	connector	<i>Gigaspora sp.</i>	91	Fungi
OTU43836	0.01	13	a12	2.46	0.65	peripheral	Chytridiomycota	96.6	Fungi
OTU44283	0.11	54	a4	5.68	1.53	peripheral	Cercozoa 1	97.7	Cercozoa
OTU44378	0.13	92	a6	0.03	0.36	peripheral	<i>Nannochloropsis sp.</i>	99.6	Eustigmatophyceae
OTU44421	0.04	13	a6	9.52	0.06	peripheral	unclassified Alveolata	89.1	unclassified Alveolata
OTU44745	0.01	13	a10	4.91	0.38	peripheral	Chytridiomycota	94.4	Fungi
OTU44875	0.01	38	a10	1.94	0.16	peripheral	Alveolate 1	92.7	Ciliophora
OTU44914	0.01	38	a14	0.41	0.51	peripheral	<i>Endogone sp.</i>	89.5	Fungi
OTU44957	1.61	92	a14	1.03	1.08	peripheral	Alveolate 1	98.1	Ciliophora
OTU45351	0.03	58	a9	6.01	3.19	connector	<i>Enchelyodon sp.</i>	96.2	Ciliophora
OTU45514	1.66	58	a3	0.67	0.87	peripheral	unclassified Dinophyceae	98.9	Dinophyceae
OTU45600	0.07	75	a9	0.34	1.63	peripheral	<i>Didinium sp.</i>	100	Ciliophora
OTU45652	0.5	83	a3	9.44	0.41	peripheral	Cercozoa 1	99.2	Cercozoa
OTU45711	0.04	75	a4	1.43	1.09	peripheral	unclassified Fungi	99.2	Fungi
OTU4581	0.05	63	a0	0.32	0.64	peripheral	Cercozoa 1	88.9	Cercozoa
OTU46247	0.02	8	a18	0.93	0.41	peripheral	Sarcosmataceae	89.4	Fungi
OTU46384	0.03	50	a4	6.8	1.42	peripheral	unclassified Rhizaria	93.8	unclassified Rhizaria
OTU46767	0.81	96	a9	6.56	4.71	connector	Haptorida	97.7	Ciliophora
OTU46895	0.29	96	a6	0.15	0.29	peripheral	unclassified Stramenopiles	98.1	Stramenopiles
OTU46993	14.13	100	a5	4.68	2.76	connector	unclassified Alveolata	97	unclassified Alveolata
OTU47146	0.01	13	a3	7.76	0.35	peripheral	unclassified Rhizaria	93.3	unclassified Rhizaria
OTU47187	0.03	42	a3	7.75	1.7	peripheral	<i>Symbiodinium sp.</i>	94.4	Dinophyceae
OTU47427	0.01	17	a12	2.78	0.74	peripheral	<i>Symbiodinium sp.</i>	90	Dinophyceae
OTU47642	0.11	54	a4	4.58	1.09	peripheral	<i>Gigaspora sp.</i>	92.9	Fungi
OTU47784	0.18	88	a8	4.39	4.51	connector	<i>Semispithidium sp.</i>	94.7	Ciliophora
OTU47889	0.03	13	a2	3.82	0.45	peripheral	Chytridiomycota	87.1	Fungi
OTU47985	0.05	63	a1	3.91	0.76	peripheral	unclassified Alveolata	96.6	unclassified Alveolata
OTU48080	0.08	71	a1	6.56	1.92	peripheral	unclassified Dinophyceae	97.4	Dinophyceae
OTU48285	0.01	21	a3	13.31	0.89	module_hub	Cercozoa 1	86	Cercozoa
OTU48386	0.06	50	a1	10.6	2.33	network_hub	unclassified Dinophyceae	96.6	Dinophyceae
OTU49053	0.01	25	a2	3.8	1.82	peripheral	Cryptomycota CM1	96.2	Fungi
OTU49244	0.38	79	a7	0.37	0.46	peripheral	unclassified Alveolata	98.1	unclassified Alveolata
OTU4925	0.03	46	a5	1.61	2.48	connector	Saprolegniaceae	99.6	Oomycetes
OTU49428	0.09	38	a7	0.29	0.47	peripheral	Chlorellaceae	98.9	Chlorophyta
OTU49865	0.00	4	a15	5.15	0.16	peripheral	unclassified Bicosoecida	84.5	Bicosoecida
OTU49866	0.03	25	a10	6.46	0.31	peripheral	Alveolate 1	96.9	Ciliophora
OTU49961	0.01	8	a7	5.57	0.05	peripheral	unclassified Bicosoecida	80.2	Bicosoecida
OTU50049	0.03	29	a1	13.7	0.59	module_hub	unclassified Alveolata	96.6	unclassified Alveolata
OTU50309	0.03	33	a15	4.25	0.14	peripheral	unclassified Perkinsea	85.7	Perkinsea
OTU50400	0.08	54	a1	0.39	0.44	peripheral	<i>Peridinium sp.</i>	98.5	Dinophyceae
OTU50665	0.02	38	a16	1.79	1.89	peripheral	<i>Kofofidinium sp.</i>	89.5	Dinophyceae
OTU50690	0.02	25	a16	1.06	3.04	connector	<i>Geosiphon sp.</i>	87	Fungi
OTU51006	0.02	38	a2	2.88	0.69	peripheral	Cryptomycota	92.4	Fungi
OTU51149	0.01	4	a15	5.15	0.16	peripheral	Perkinsea 2	83.1	Perkinsea
OTU51237	0.05	58	a5	6.64	1.6	peripheral	unclassified Alveolata	93.2	unclassified Alveolata
OTU51326	0.02	38	a2	3.37	1.42	peripheral	<i>Raabena sp.</i>	90.9	Ciliophora
OTU51389	0.5	67	a5	1.41	1.1	peripheral	unclassified Fungi	99.6	Fungi
OTU51490	0.01	29	a0	0.6	0.18	peripheral	unclassified Rhizaria	95.5	unclassified Rhizaria
OTU51536	0.07	67	a3	0.07	0.25	peripheral	unclassified Stramenopiles	97.7	Stramenopiles
OTU51678	0.07	71	a1	2.28	0.89	peripheral	<i>Symbiodinium sp.</i>	97.8	Dinophyceae
OTU5173	0.14	46	a13	0.16	0.65	peripheral	unclassified Stramenopiles	97.8	Stramenopiles
OTU51758	0.01	29	a3	8.15	3.26	connector	unclassified Alveolata	98.1	unclassified Alveolata
OTU52195	0.03	4	a0	1.04	0.57	peripheral	Cryptomycota	94.3	Fungi
OTU52304	0.04	38	a2	7.27	2.16	connector	Cryptomycota CM1	88.5	Fungi
OTU52553	0.02	33	a12	0.72	1.19	peripheral	Syndiniales	92.5	Dinophyceae
OTU52631	0.11	75	a8	0.65	2.53	connector	Lagenidiales	99.6	Oomycetes
OTU52649	0.02	58	a10	2.29	2.68	connector	<i>Blepharocorys sp.</i>	92.4	Ciliophora
OTU52687	0.09	33	a8	1.85	4.14	connector	Chytridiomycota	95.1	Fungi
OTU53141	0.15	8	a0	0.5	0.37	peripheral	Chlamydomonadales	98.9	Chlorophyta
OTU53153	0.01	29	a12	0.87	1.72	peripheral	unclassified Rhizaria	92.9	unclassified Rhizaria
OTU53445	0.05	50	a1	10.37	2.14	network_hub	unclassified Dinophyceae	97	Dinophyceae
OTU53489	0.42	96	a4	3.88	2.56	connector	Chytridiomycota	99.6	Fungi
OTU53534	0.01	38	a11	2.54	4.62	connector	<i>Balantidium sp.</i>	94.7	Ciliophora
OTU53880	0.02	29	a2	1.08	0.46	peripheral	<i>Lobulomyces sp.</i>	91	Fungi
OTU53968	0.06	67	a20	0.91	1.68	peripheral	unclassified Alveolata	96.3	unclassified Alveolata
OTU54049	0.1	50	a1	14.51	0.77	module_hub	<i>Polykrikos sp.</i>	98.1	Dinophyceae
OTU54068	0.03	38	a4	4.31	1.45	peripheral	<i>Sistotrema sp.</i>	93.6	Fungi
OTU54240	0.11	71	a2	4.1	1.5	peripheral	unclassified Alveolata	94.3	unclassified Alveolata
OTU54358	0.05	71	a20	1.69	1.82	peripheral	unclassified Alveolata	98.1	unclassified Alveolata
OTU54811	0.05	92	a0	0.24	1.44	peripheral	<i>Balantidium sp.</i>	96.6	Ciliophora
OTU5512	0.01	4	a5	0.1	0.35	peripheral	Ascomycota	99.6	Fungi
OTU55663	0.24	79	a5	0.53	1.28	peripheral	unclassified Alveolata	99.2	unclassified Alveolata
OTU55868	0.05	21	a6	9.35	0.03	peripheral	<i>Symbiodinium sp.</i>	93.6	Dinophyceae
OTU56299	0.03	46	a4	1.69	2.9	connector	Cercozoa 1	93.3	Cercozoa
OTU56391	0.15	88	a5	7.06	2.14	connector	unclassified Alveolata	95.5	unclassified Alveolata
OTU56479	0.01	21	a4	4.85	0.23	peripheral	Chytridiales	91.4	Fungi
OTU56583	0.01	8	a11	2.83	1.7	peripheral	unclassified Rhizaria	92.7	unclassified Rhizaria
OTU56729	0.01	17	a12	2.12	1.2	peripheral	<i>Plasmodiophora sp.</i>	90.1	Cercozoa
OTU56795	0.03	46	a11	2.68	4.04	connector	Cercozoa 2	93	Cercozoa
OTU57254	0.4	79	a4	6.71	1.2	peripheral	Chytridiomycota	93.5	Fungi
OTU57450	0.87	96	a5	2.95	2.07	connector	unclassified Alveolata	95.5	unclassified Alveolata
OTU57556	0.06	50	a4	4.08	2.6	connector	unclassified Fungi	99.6	Fungi
OTU57607	0.01	25	a13	1.14	0.53	peripheral	unclassified Fungi	98.1	Fungi
OTU57921	0.01	17	a19	3.72	2.42	connector	unclassified Stramenopiles	96.6	Stramenopiles
OTU58232	0.01	21	a2	7.8	1.83	peripheral	unclassified Alveolata	87.1	unclassified Alveolata
OTU58447	0.03	25	a4	1.94	1.17	peripheral	<i>Ochromonas sp.</i>	97.4	Chrysophyceae
OTU58701	0.02	21	a2	3.94	0.74	peripheral	<i>Plasmodiophora sp.</i>	87.3	Cercozoa
OTU58707	0.15	38	a1	1.99	0.69	peripheral	<i>Scherffelia sp.</i>	100	Chlorophyta
OTU58945	0.01	4	a3	18.11	0.2	module_hub	<i>Oxytrichidae sp.</i>	95.5	Ciliophora
OTU59130	0.44	71	a11	2.52	3.41	connector	Cryptomycota LKM11	97.7	Fungi
OTU59140	0.03	42	a8	4.01	2.11	connector	<i>Semispithidium sp.</i>	95.1	Ciliophora
OTU59296	0.02	13	a12	1.43	3.74	connector	Chytridiomycota	95.8	Fungi
OTU59597	0.03	54	a5	5.85	2.00	peripheral	unclassified Alveolata	95.1	unclassified Alveolata
OTU60288	0.06	67	a5	6.39	1.08	peripheral	Perkinsea 1	92.5	Perkinsea
OTU60340	0.03	38	a14	0.89	0.64	peripheral	Alveolate 1	97.7	Ciliophora
OTU61533	0.03	25	a1	15.15	0.99	module_hub	unclassified Dinophyceae	95.9	Dinophyceae
OTU61585	0.03	54	a2	2.18	2.06	connector	Cercozoa 2	92.2	Cercozoa

OTU61621	0.03	4	a4	3.35	0.47	peripheral	Perkinsea 2	87.1	Perkinsea
OTU62093	0.02	25	a16	0.87	0.97	peripheral	<i>Candida</i> sp.	94.7	Fungi
OTU62122	0.01	25	a14	0.85	0.16	peripheral	unclassified Rhizaria	93.3	unclassified Rhizaria
OTU62190	0.01	29	a14	1.34	0.21	peripheral	Chlorellaceae	98.1	Chlorophyta
OTU62195	0.04	71	a2	3.38	2.33	connector	<i>Amylovoracidae</i>	87.8	Ciliophora
OTU62202	0.01	4	a18	3.63	0.37	peripheral	<i>Tetracladium</i> sp.	99.6	Fungi
OTU62641	0.04	25	a11	4.84	1.63	peripheral	<i>Cosmarium</i> sp.	92.2	Streptophyta
OTU62964	0.03	38	a1	9.81	2.91	connector	unclassified Dinophyceae	97	Dinophyceae
OTU63100	0.01	4	a3	18.11	0.2	module_hub	unclassified Fungi	90.7	Fungi
OTU63222	0.01	4	a19	2.86	0.17	peripheral	<i>Ochromonas</i> sp.	96.2	Chrysophyceae
OTU63251	0.06	83	a8	4.22	1.9	peripheral	<i>Semispithidium</i> sp.	96.3	Ciliophora
OTU63381	0.01	21	a16	1.38	1.82	peripheral	Cryptomycota	95.1	Fungi
OTU63382	0.09	33	a16	0.41	1.5	peripheral	<i>Triparticalcar</i> sp.	89.8	Fungi
OTU63696	0.02	17	a13	2.16	0.57	peripheral	unclassified Alveolata	95.1	unclassified Alveolata
OTU6405	0.03	75	a5	0.6	0.47	peripheral	Exobasidiomycetes	95.1	Fungi
OTU64074	0.5	83	a5	0.3	0.38	peripheral	unclassified Apicomplexa	99.6	Apicomplexa
OTU64186	0.02	50	a5	2.7	3.22	connector	unclassified Apicomplexa	98.1	Apicomplexa
OTU64276	0.00	4	a11	4.03	1.05	peripheral	<i>Bitricha</i> sp.	89.7	Ciliophora
OTU64284	0.15	58	a7	0.9	0.43	peripheral	Pleosporales	100	Fungi
OTU64463	0.02	42	a5	3.29	1.43	peripheral	unclassified Alveolata	95.8	unclassified Alveolata
OTU64653	0.02	58	a0	0.43	0.13	peripheral	unclassified Rhizaria	94.8	unclassified Rhizaria
OTU64686	0.02	21	a1	0.55	0.49	peripheral	unclassified Ciliophora	99.2	Ciliophora
OTU64740	0.09	83	a10	5.03	0.89	peripheral	Alveolate 1	92	Ciliophora
OTU64774	0.12	92	a11	1.96	1.46	peripheral	Chytridiomycota	89.8	Fungi
OTU64841	0.01	29	a2	1.63	0.65	peripheral	Cryptomycota	92.4	Fungi
OTU64848	0.06	38	a13	1.13	0.18	peripheral	Microbotryomycetes	99.6	Fungi
OTU64878	0.13	79	a9	5.81	3.17	connector	<i>Balantidium</i> sp.	97.3	Ciliophora
OTU64900	1.16	96	a4	6.77	1.46	peripheral	Fragilariaceae	99.3	Bacillariophyta
OTU6493	0.01	29	a10	1.53	1.28	peripheral	Perkinsea 1	94.7	Perkinsea
OTU64958	0.01	8	a3	17.59	0.26	module_hub	Cercozoa 1	94.8	Cercozoa
OTU6522	1.67	79	a15	1.1	0.85	peripheral	Chytridiomycota	81	Fungi
OTU65552	0.16	54	a2	1.8	2.31	connector	<i>Plasmodiophora</i> sp.	89.7	Cercozoa
OTU65685	0.01	21	a17	2.99	1.67	peripheral	unclassified Fungi	85.7	Fungi
OTU6584	0.04	71	a10	4.45	1.44	peripheral	Spathidiidae	95.4	Ciliophora
OTU66161	0.02	38	a1	8.8	2.61	connector	unclassified Dinophyceae	97.4	Dinophyceae
OTU66244	0.05	63	a1	2.5	0.94	peripheral	unclassified Alveolata	97.4	unclassified Alveolata
OTU66524	0.01	29	a6	1.43	0.22	peripheral	Perkinsea 2	85.2	Perkinsea
OTU66826	0.04	33	a12	2.08	2.15	connector	unclassified Fungi	95.1	Fungi
OTU66846	0.01	25	a13	1.36	0.34	peripheral	Boletales	97	Fungi
OTU67220	0.01	33	a4	3.25	1.46	peripheral	Haptorida	94.3	Ciliophora
OTU67227	0.01	29	a10	0.49	2.76	connector	<i>Arcuospithidium</i> sp.	97.3	Ciliophora
OTU67365	0.29	83	a4	2.95	3.45	connector	unclassified Apicomplexa	99.2	Apicomplexa
OTU67531	0.01	17	a1	12.98	0.37	module_hub	unclassified Dinophyceae	95.9	Dinophyceae
OTU67806	0.05	42	a1	12.2	1.15	module_hub	unclassified Alveolata	96.7	unclassified Alveolata
OTU68062	0.04	63	a1	0.12	1.02	peripheral	unclassified Alveolata	94.7	unclassified Alveolata
OTU68386	0.01	8	a14	1.92	0.25	peripheral	unclassified Fungi	95.1	Fungi
OTU68538	0.01	4	a6	9.83	0.06	peripheral	<i>Colpodidium</i> sp.	90.5	Ciliophora
OTU68950	0.01	8	a2	2.23	0.64	peripheral	unclassified Chlorophyta	97.4	Chlorophyta
OTU68999	0.01	17	a10	6.66	0.32	peripheral	Alveolate 1	97.3	Ciliophora
OTU69179	0.23	67	a9	1.00	1.75	peripheral	Cryptomycota	93.2	Fungi
OTU69251	0.02	50	a8	0.22	1.16	peripheral	Cryptomycota CM1	95.2	Fungi
OTU69686	0.05	63	a5	2.31	1.2	peripheral	unclassified Alveolata	94.7	unclassified Alveolata
OTU69713	0.04	50	a20	1.54	2.56	connector	<i>Woloszynskia</i> sp.	98.1	Dinophyceae
OTU69715	0.01	8	a2	3.37	0.17	peripheral	Cryptomycota LKM11	98.1	Fungi
OTU69909	0.08	67	a17	0.94	0.83	peripheral	unclassified Bicosoecida	95.4	Bicosoecida
OTU70016	0.04	88	a8	2.79	1.2	peripheral	<i>Semispithidium</i> sp.	97.7	Ciliophora
OTU70291	0.01	21	a12	4.05	0.63	peripheral	unclassified Dinophyceae	92.9	Dinophyceae
OTU70335	0.01	38	a9	0.9	1.16	peripheral	<i>Triplumaria</i> sp.	91.6	Ciliophora
OTU70434	0.02	21	a4	2.67	2.31	connector	unclassified Rhizaria	96.6	unclassified Rhizaria
OTU70857	0.01	42	a9	3.71	2.63	connector	Haptorida	96.6	Ciliophora
OTU70903	0.01	21	a5	1.13	1.05	peripheral	<i>Penicillium</i> sp.	99.2	Fungi
OTU71401	0.01	21	a16	2.08	1.51	peripheral	Chytridiomycota	95.5	Fungi
OTU7163	0.77	96	a8	4.84	2.34	connector	Haptorida	98.5	Ciliophora
OTU71740	0.01	4	a2	2.8	0.2	peripheral	<i>Condylostoma</i> sp.	60.3	Ciliophora
OTU718	0.02	29	a4	6.77	1.11	peripheral	Chytridiomycota	97.7	Fungi
OTU72074	0.02	46	a14	1.39	0.61	peripheral	Alveolate 1	97.3	Ciliophora
OTU72265	0.02	17	a1	15.89	0.84	module_hub	unclassified Alveolata	98.5	unclassified Alveolata
OTU7229	0.02	46	a7	0.22	0.62	peripheral	<i>Plasmodiophora</i> sp.	89.7	Cercozoa
OTU72656	0.02	46	a8	2.9	2.41	connector	<i>Balantidium</i> sp.	97.7	Ciliophora
OTU7291	0.01	21	a4	7.37	1.05	peripheral	Chytridiomycota	97.4	Fungi
OTU72947	0.02	42	a1	7.49	2.78	connector	<i>Pfiesteria</i> sp.	97.4	Dinophyceae
OTU7328	0.01	17	a3	16.72	0.58	module_hub	Trachelidae	99.6	Ciliophora
OTU73320	0.02	13	a14	3.13	0.3	peripheral	<i>Condylostoma</i> sp.	60.4	Ciliophora
OTU7342	0.08	21	a12	4.09	0.37	peripheral	MAST-4	88.4	MAST-4
OTU73478	0.02	25	a2	1.1	0.99	peripheral	Cercozoa 1	91.4	Cercozoa
OTU74002	0.01	8	a10	0.82	0.8	peripheral	Cercozoa 2	88.8	Cercozoa
OTU74223	0.01	25	a17	2.79	2.62	connector	unclassified Chrysophyceae	99.3	Chrysophyceae
OTU74356	0.01	17	a11	1.68	0.78	peripheral	<i>Gymnophrys</i> sp.	80.6	Cercozoa
OTU74521	0.01	17	a3	13.25	0.94	module_hub	unclassified Dinophyceae	97.8	Dinophyceae
OTU74646	0.04	50	a0	0.18	0.32	peripheral	unclassified Fungi	98.9	Fungi
OTU74648	0.76	100	a1	2.33	1.03	peripheral	<i>Desmodesmus</i> sp.	99.6	Chlorophyta
OTU74714	0.01	13	a3	13.6	2.12	network_hub	Cercozoa 1	97.4	Cercozoa
OTU75188	0.00	4	a3	18.11	0.2	module_hub	unclassified Dinophyceae	93.6	Dinophyceae
OTU75649	0.05	42	a0	0.37	0.85	peripheral	Chytridiomycota	95.5	Fungi
OTU76015	0.04	71	a5	1.32	1.93	peripheral	Perkinsea 1	97.4	Perkinsea
OTU76139	0.05	92	a2	0.41	0.54	peripheral	Chytridiomycota	98.1	Fungi
OTU76236	0.05	58	a3	9.8	1.15	peripheral	Cercozoa 1	96.6	Cercozoa
OTU76545	0.1	54	a6	0.38	2.46	connector	Schizosaccharomyces	86.5	Fungi
OTU76581	0.05	46	a7	0.4	0.18	peripheral	Chytridiomycota	95.8	Fungi
OTU76774	0.19	92	a5	7.02	1.99	peripheral	unclassified Alveolata	93.2	unclassified Alveolata
OTU76947	0.69	92	a20	2.59	2.52	connector	<i>Woloszynskia</i> sp.	98.9	Dinophyceae
OTU77154	0.01	4	a3	0.4	0.7	peripheral	Chlamydomonadaceae	97.3	Chlorophyta
OTU77249	0.04	42	a4	7.66	0.7	peripheral	Cercozoa 2	96.6	Cercozoa
OTU77506	0.01	13	a4	8.34	0.85	peripheral	Cercozoa 1	88.5	Cercozoa
OTU77560	0.01	29	a8	0.33	0.72	peripheral	unclassified Alveolata	95.1	unclassified Alveolata
OTU77575	0.05	63	a4	1.77	0.7	peripheral	Chytridiomycota	95.9	Fungi
OTU77667	0.2	92	a20	2.66	2.57	connector	unclassified Alveolata	96.3	unclassified Alveolata
OTU77687	0.03	50	a1	0.76	1.08	peripheral	<i>Symbiodinium</i> sp.	96.6	Dinophyceae



OTU77782	0.06	67	a15	0.15	0.56	peripheral	Cercozoa 1	91.7	Cercozoa
OTU77847	0.03	29	a2	4.35	1.86	peripheral	Alveolate 1	95.8	Ciliophora
OTU78036	0.01	25	a4	4.7	0.33	peripheral	Cercozoa 1	93.6	Cercozoa
OTU7819	0.87	100	a5	7.25	2.31	connector	Perkinsea 1	91.7	Perkinsea
OTU78212	0.06	25	a6	1.4	0.73	peripheral	Cryptomycota	92.4	Fungi
OTU78302	0.04	58	a1	6.3	1.16	peripheral	unclassified Dinophyceae	99.2	Dinophyceae
OTU78509	0.01	17	a7	1.64	0.3	peripheral	Pleosporales	96.2	Fungi
OTU78580	0.02	25	a13	0.92	0.5	peripheral	unclassified Rhizaria	95.6	unclassified Rhizaria
OTU78615	0.01	17	a3	15.07	2.47	network_hub	Cercozoa 1	95.9	Cercozoa
OTU79410	0.14	63	a7	1.41	1.83	peripheral	Ascomycota	99.6	Fungi
OTU7964	0.05	54	a4	8.24	1.11	peripheral	<i>Epispathidium</i> sp.	97.3	Ciliophora
OTU79742	0.01	8	a6	3.79	0.47	peripheral	<i>Diophrys</i> sp.	91.3	Ciliophora
OTU79931	0.01	8	a16	1.85	0.5	peripheral	Cercozoa 1	91.1	Cercozoa
OTU79981	0.03	33	a2	2.11	1.54	peripheral	unclassified Fungi	97.3	Fungi
OTU8005	0.04	46	a9	0.94	2.68	connector	Cryptomycota	90.9	Fungi
OTU80070	0.02	17	a17	2.01	2.02	connector	unclassified Rhizaria	92.6	unclassified Rhizaria
OTU80096	0.01	29	a9	4.48	1.99	peripheral	<i>Enchelyodon</i> sp.	96.2	Ciliophora
OTU80146	0.04	67	a8	0.5	0.85	peripheral	Perkinsea 1	97.4	Perkinsea
OTU80160	0.01	4	a3	18.11	0.2	module_hub	Cercozoa 1	96.6	Cercozoa
OTU80182	0.03	46	a2	0.68	0.24	peripheral	MAST-7A	91	MAST-7
OTU80277	0.02	13	a7	4.42	0.04	peripheral	<i>Gomphonema</i> sp.	98.5	Bacillariophyta
OTU80345	0.01	21	a13	1.57	1.03	peripheral	<i>Uronychia</i> sp.	86.6	Ciliophora
OTU80434	0.01	13	a3	8.85	0.64	peripheral	Cercozoa 1	95.9	Cercozoa
OTU80462	0.31	88	a5	0.85	1.74	peripheral	unclassified Alveolata	99.2	unclassified Alveolata
OTU80493	0.02	33	a16	0.92	1.61	peripheral	Cercozoa 1	96.3	Cercozoa
OTU80720	0.01	25	a8	1.28	0.92	peripheral	unclassified Dinophyceae	89.8	Dinophyceae
OTU80731	0.93	92	a19	0.72	1.79	peripheral	unclassified Rhizaria	93.6	unclassified Rhizaria
OTU80895	0.04	46	a2	3.72	0.89	peripheral	unclassified Ciliophora	92.7	Ciliophora
OTU80967	0.46	92	a2	2.11	0.98	peripheral	Cryptomycota	97.3	Fungi
OTU80989	0.03	50	a0	0.12	0.57	peripheral	<i>Woloszynskia</i> sp.	88.1	Dinophyceae
OTU81072	0.01	21	a3	12.95	0.4	module_hub	Cercozoa 1	96.3	Cercozoa
OTU81083	0.03	33	a1	8.55	2.29	connector	unclassified Dinophyceae	98.1	Dinophyceae
OTU81089	0.01	8	a13	3.55	0.47	peripheral	Chytridiomycota	98.1	Fungi
OTU81159	0.01	17	a1	0.56	0.38	peripheral	<i>Dunaliella</i> sp.	90.6	Chlorophyta
OTU81624	0.02	29	a1	10.95	2.9	network_hub	<i>Peridinium</i> sp.	98.1	Dinophyceae
OTU81816	0.2	83	a5	0.65	2.19	connector	unclassified Alveolata	93.2	unclassified Alveolata
OTU81818	0.04	29	a11	3.14	1.38	peripheral	Cryptomycota LKM11	97	Fungi
OTU81980	0.00	4	a7	5.92	0.05	peripheral	<i>Chloroparva</i> sp.	97.3	Chlorophyta
OTU82016	0.01	4	a18	3.63	0.37	peripheral	<i>Gigaspora</i> sp.	85.5	Fungi
OTU82307	0.02	33	a14	0.33	0.31	peripheral	<i>Blepharocorys</i> sp.	92.8	Ciliophora
OTU82355	0.01	21	a3	2.03	1.04	peripheral	Ochromonadaceae	97.8	Synurophyceae
OTU82469	0.12	71	a20	2.00	2.98	connector	unclassified Alveolata	98.1	unclassified Alveolata
OTU82555	0.02	63	a10	2.74	0.79	peripheral	<i>Blepharocorys</i> sp.	92.4	Ciliophora
OTU82595	0.13	58	a2	3.51	2.83	connector	<i>Blepharocorys</i> sp.	92.4	Ciliophora
OTU82660	0.37	46	a7	0.5	0.49	peripheral	<i>Plasmodiophora</i> sp.	87.9	Cercozoa
OTU8276	0.1	71	a5	1.79	2.85	connector	unclassified Alveolata	94.9	unclassified Alveolata
OTU83052	0.02	13	a19	4.49	1.9	peripheral	Chytridiomycota	91.4	Fungi
OTU83430	0.01	25	a0	0.59	0.96	peripheral	<i>Chrysosphaerella</i> sp.	96.2	Synurophyceae
OTU83566	0.01	17	a17	1.98	1.33	peripheral	unclassified Chrysophyceae	97.8	Chrysophyceae
OTU83700	0.1	67	a11	2.07	1.46	peripheral	<i>Paraglomus</i> sp.	89.8	Fungi
OTU83732	0.04	8	a0	0.32	0.76	peripheral	Boletales	99.6	Fungi
OTU83851	0.01	8	a6	3.61	0.51	peripheral	Chytridiomycota	90.6	Fungi
OTU83880	0.01	4	a6	9.83	0.06	peripheral	<i>Raabena</i> sp.	90.9	Ciliophora
OTU8447	0.02	25	a4	1.02	1.32	peripheral	unclassified Rhizaria	93.3	unclassified Rhizaria
OTU84554	0.01	13	a3	6.34	0.31	peripheral	<i>Cyclidium</i> sp.	73.8	Ciliophora
OTU84763	0.01	13	a19	3.88	1.47	peripheral	<i>Ochromonas</i> sp.	94.7	Chrysophyceae
OTU84837	0.02	33	a5	1.00	3.68	connector	Saprogleniaceae	88.3	Oomycetes
OTU85080	0.00	4	a9	0.6	0.87	peripheral	unclassified Cercozoa	89.5	Cercozoa
OTU85107	0.03	54	a2	3.27	2.18	connector	Glomeromycota	89.1	Fungi
OTU8547	0.02	21	a2	2.74	1.53	peripheral	unclassified Dinophyceae	91.7	Dinophyceae
OTU85607	7.57	100	a8	2.84	2.92	connector	<i>Epispathidium</i> sp.	98.9	Ciliophora
OTU85789	0.02	29	a4	1.78	1.43	peripheral	Cercozoa 1	92.6	Cercozoa
OTU85848	0.06	38	a12	0.28	1.44	peripheral	unclassified Haptophyceae	99.2	Haptophyceae
OTU86081	0.31	88	a19	0.67	0.65	peripheral	Trebouxiophyceae	99.6	Chlorophyta
OTU86195	0.01	8	a4	6.3	0.53	peripheral	Chytridiomycota	92.8	Fungi
OTU86250	0.02	17	a7	0.94	0.27	peripheral	Cryptomycota	95.1	Fungi
OTU86445	0.02	21	a1	13.25	0.63	module_hub	unclassified Alveolata	98.5	unclassified Alveolata
OTU86847	0.01	8	a12	3.32	0.25	peripheral	unclassified Dinophyceae	90	Dinophyceae
OTU87028	0.03	67	a0	0.21	1.09	peripheral	<i>Arcuospathidium</i> sp.	95.2	Ciliophora
OTU87225	0.33	96	a4	1.23	1.74	peripheral	unclassified Alveolata	98.5	unclassified Alveolata
OTU87405	0.02	46	a2	0.91	0.74	peripheral	<i>Diophrys</i> sp.	91	Ciliophora
OTU8754	0.02	42	a18	0.43	0.42	peripheral	unclassified Fungi	97.3	Fungi
OTU87900	0.24	17	a3	15.31	1.28	module_hub	unclassified Ciliophora	98.5	Ciliophora
OTU88053	0.7	100	a5	7.62	2.31	connector	<i>Diophrys</i> sp.	89.4	Ciliophora
OTU88082	0.03	50	a1	3.42	3.9	connector	unclassified Dinophyceae	97.4	Dinophyceae
OTU8823	0.1	54	a13	0.7	1.44	peripheral	Cryptomycota	95.4	Fungi
OTU88514	0.02	42	a10	5.02	0.7	peripheral	<i>Blepharocorys</i> sp.	91.6	Ciliophora
OTU88547	0.6	63	a6	3.11	0.24	peripheral	unclassified Fungi	91.3	Fungi
OTU88635	0.37	54	a9	1.29	3.5	connector	<i>Pyramimonas</i> sp.	90.8	Chlorophyta
OTU89021	0.03	67	a5	5.22	1.42	peripheral	unclassified Alveolata	94.7	unclassified Alveolata
OTU8938	0.12	46	a4	5.81	0.72	peripheral	Cryptomycota CM1	92.8	Fungi
OTU89472	0.29	92	a7	0.28	1.76	peripheral	<i>Mychonastes</i> sp.	99.6	Chlorophyta
OTU89696	0.04	29	a6	2.71	0.83	peripheral	unclassified Alveolata	98.9	unclassified Alveolata
OTU90104	0.01	8	a6	3.27	0.44	peripheral	<i>Litonotus</i> sp.	88.6	Ciliophora
OTU90383	0.01	25	a10	3.55	0.4	peripheral	Alveolate 1	95.4	Ciliophora
OTU90426	0.62	75	a1	14.38	0.5	module_hub	unclassified Dinophyceae	100	Dinophyceae
OTU90540	0.03	46	a3	0.34	0.46	peripheral	Cercozoa 1	97.4	Cercozoa
OTU90577	0.01	13	a4	2.37	0.32	peripheral	<i>Hemiophrys</i> sp.	92	Ciliophora
OTU90674	0.01	17	a18	2.42	0.29	peripheral	Cryptomycota	93.9	Fungi
OTU91124	0.02	4	a6	9.83	0.06	peripheral	<i>Diophrys</i> sp.	62.8	Ciliophora
OTU91130	0.00	8	a3	19.67	0.58	module_hub	unclassified Dinophyceae	98.5	Dinophyceae
OTU91336	0.01	8	a3	15.26	0.25	module_hub	<i>Cryptococcus</i> sp.	91.7	Fungi
OTU91517	0.01	21	a1	15.45	0.96	module_hub	unclassified Dinophyceae	96.2	Dinophyceae
OTU9388	0.03	67	a17	0.79	2.01	connector	<i>Mallomonas</i> sp.	98.9	Synurophyceae
OTU9425	0.01	21	a6	1.22	0.56	peripheral	<i>Blepharocorys</i> sp.	93.6	Ciliophora
OTU9991	0.05	58	a11	1.54	1.25	peripheral	Cryptomycota LKM11	95.8	Fungi

## (B) Lake Bourget

OTUs	% DNA sequences	% Occurrence	#modules	intraM	interM	Topological_roles	Best hit identity	Identity	Taxonomic group
OTU10147	0	6	b13	9.67	1.1	peripheral	unclassified Fungi	92.5	Fungi
OTU1027	0.23	100	b4	3.54	0.24	peripheral	<i>Semiospathidium sp.</i>	98.5	Ciliophora
OTU10305	0.07	80	b4	1.82	0.08	peripheral	unclassified Dinophyceae	98.1	Dinophyceae
OTU10385	0.01	11	b3	8.75	1.32	peripheral	Chromulinaceae	91	Chrysophyceae
OTU10567	0.06	89	b5	0.77	0.31	peripheral	Chlamydomonadaceae	97.7	Chlorophyta
OTU10600	0.05	57	b1	12.47	3.95	network_hub	Saccharomycetaceae	99.6	Fungi
OTU1072	0.01	6	b18	3.62	1.98	peripheral	Chytridiomycota	91.6	Fungi
OTU10949	1.82	91	b3	8.55	2.16	connector	<i>Chrysosaccus sp.</i>	95.9	Chrysophyceae
OTU11287	0.01	17	b5	1.39	0.15	peripheral	<i>Pascherina sp.</i>	96.6	Chlorophyta
OTU11365	0.06	40	b6	5.31	9.2	connector	unclassified Alveolata	94.7	unclassified Alveolata
OTU12560	0.02	43	b0	0.3	0.7	peripheral	Haptorida	92.5	Ciliophora
OTU12882	0.11	77	b4	0.36	0.67	peripheral	Cryptomycota	92.8	Fungi
OTU12893	0.09	91	b2	2.47	1.24	peripheral	unclassified Dinophyceae	95.5	Dinophyceae
OTU1320	0.26	97	b5	0.18	0.56	peripheral	unclassified Dinophyceae	94.7	Dinophyceae
OTU13218	0.03	54	b1	11.04	6.88	network_hub	unclassified Fungi	99.6	Fungi
OTU1327	0.01	6	b18	3.58	2.15	connector	unclassified Alveolata	91.3	unclassified Alveolata
OTU13351	0.01	9	b3	13.91	1.04	module_hub	Spathidiidae	94.4	Ciliophora
OTU13557	0.02	46	b8	2.78	2.76	connector	<i>Plasmodiophora sp.</i>	87.5	Cercozoa
OTU13600	0.14	31	b8	7.5	4.68	connector	unclassified Fungi	88.3	Fungi
OTU13988	0.01	31	b3	7.6	1.39	peripheral	<i>Rhizophydium sp.</i>	97.3	Fungi
OTU14339	0.01	20	b11	0.86	0.03	peripheral	Alveolate 1	96.9	Ciliophora
OTU14347	0.02	20	b1	16.26	2.78	network_hub	Cryptomycota	93.5	Fungi
OTU14370	0.03	60	b4	2.05	0.9	peripheral	<i>Trachelophyllum sp.</i>	97	Ciliophora
OTU14707	0.04	69	b10	0.39	2.01	connector	Haptorida	97.7	Ciliophora
OTU1474	0.03	60	b7	4.75	1.8	peripheral	<i>Protospathidium sp.</i>	98.5	Ciliophora
OTU14970	0.04	71	b11	0.51	0.34	peripheral	<i>Blepharocorys sp.</i>	92.4	Ciliophora
OTU15006	0.3	91	b13	7.14	1.36	peripheral	unclassified Fungi	95.8	Fungi
OTU15175	0.01	14	b12	2.74	0.38	peripheral	<i>Oxytricha sp.</i>	97	Ciliophora
OTU15284	0.04	66	b20	1.23	0.32	peripheral	<i>Epispathidium sp.</i>	98.1	Ciliophora
OTU15648	0.03	71	b2	1.12	1.44	peripheral	unclassified Dinophyceae	95.5	Dinophyceae
OTU15803	0.63	100	b1	18.91	4.49	network_hub	Leotiomyces	99.6	Fungi
OTU15833	0.09	66	b4	1.07	0.83	peripheral	Chlorophyceae	97.4	Chlorophyta
OTU15881	3.62	100	b5	3.53	0.3	peripheral	Chlamydomonadaceae	99.6	Chlorophyta
OTU16326	0.01	26	b1	10.42	2.81	network_hub	<i>Colpodella sp.</i>	90.2	Colpodellidae
OTU16670	0.02	43	b17	0.25	0.49	peripheral	unclassified Stramenopiles	86.8	Stramenopiles
OTU17061	0.02	40	b2	8.92	3.46	connector	Cercozoa 1	92.6	Cercozoa
OTU17093	0.01	34	b5	0.77	0.64	peripheral	MAST-6	95.9	MAST-6
OTU17098	0.01	29	b1	4.03	2.69	connector	Dikarya	99.6	Fungi
OTU17329	0.04	83	b4	1.9	0.57	peripheral	<i>Semiospathidium sp.</i>	98.1	Ciliophora
OTU17495	0.02	66	b1	0.74	0.22	peripheral	Tintinnophagus sp.	90.6	Dinophyceae
OTU17553	0.02	49	b12	0.46	0.48	peripheral	Cryptomycota LKM11	94.7	Fungi
OTU17686	0.01	20	b1	11.28	2.71	network_hub	unclassified Stramenopiles	97.4	Stramenopiles
OTU17741	0.01	20	b5	4.47	0.14	peripheral	<i>Chlamydomonas sp.</i>	96.6	Chlorophyta
OTU17762	0.13	100	b0	0.02	0.43	peripheral	unclassified Rhizaria	91.4	unclassified Rhizaria
OTU17839	0.01	17	b3	4.94	0.15	peripheral	unclassified Chrysophyceae	94.8	Chrysophyceae
OTU17977	0.02	63	b20	0.25	0.15	peripheral	<i>Epispathidium sp.</i>	97	Ciliophora
OTU18003	0.01	11	b0	1.86	0.06	peripheral	<i>Sarcocystis sp.</i>	90.2	Apicomplexa
OTU18016	0.43	69	b1	14.63	3.23	network_hub	Cryptomycota	93.9	Fungi
OTU18095	0.03	14	b13	7.96	0.44	peripheral	unclassified Rhizaria	92.5	unclassified Rhizaria
OTU18296	0.14	100	b6	3.3	2.86	connector	Cryptomycota CM1	97.7	Fungi
OTU18323	0.01	11	b4	2.27	0.17	peripheral	unclassified Apicomplexa	98.5	Apicomplexa
OTU1844	0.2	100	b1	5.64	5.55	connector	Cryptomycota	94.7	Fungi
OTU18613	0.01	9	b10	3.02	0.28	peripheral	Stichotrichia	97.7	Ciliophora
OTU18660	0.01	14	b19	2.24	2	connector	<i>Pheopolykrikos sp.</i>	90.6	Dinophyceae
OTU18853	0.02	31	b4	1.69	0.11	peripheral	<i>Anylovorax sp.</i>	89.1	Ciliophora
OTU18858	0	9	b13	8.24	2.59	connector	unclassified Fungi	92.8	Fungi
OTU19134	0.03	34	b12	3.17	5.97	connector	Alveolate 1	93.1	Ciliophora
OTU19308	0	3	b8	9.81	0.47	peripheral	<i>Plasmodiophora sp.</i>	86.8	Cercozoa
OTU19368	0.03	57	b2	0.61	1.05	peripheral	unclassified Alveolata	97.7	unclassified Alveolata
OTU19423	0.04	60	b4	0.58	0.29	peripheral	unclassified Apicomplexa	98.1	Apicomplexa
OTU19456	0.04	66	b2	0.93	1.33	peripheral	<i>Trachelophyllum sp.</i>	97	Ciliophora
OTU19538	0.01	17	b6	2.71	4.56	connector	Cryptomycota CM1	95.4	Fungi
OTU19581	0.04	69	b1	9.04	5.77	connector	Cryptomycota	95.1	Fungi
OTU1976	0.04	46	b10	3.41	4.04	connector	<i>Nephroselmis sp.</i>	90.7	Chlorophyta
OTU19806	0.01	20	b7	4.22	2.06	connector	Cryptomycota LKM11	96.2	Fungi
OTU19876	0.01	20	b13	3.01	1.48	peripheral	unclassified Alveolata	88.9	unclassified Alveolata
OTU20074	0.2	97	b5	6.21	0.15	peripheral	<i>Chlamydomonas sp.</i>	99.6	Chlorophyta
OTU20365	0.02	14	b0	1.11	0.44	peripheral	unclassified Stramenopiles	97.7	Stramenopiles
OTU20786	0.01	26	b1	1.11	1.1	peripheral	<i>Condylostoma sp.</i>	85.2	Ciliophora
OTU20906	2.98	100	b16	0.45	2.16	connector	unclassified Alveolata	96.2	unclassified Alveolata
OTU20958	0.06	69	b6	4.07	10.97	connector	Cryptomycota	94	Fungi
OTU21078	0.01	37	b20	0.09	0.94	peripheral	Haptorida	96.6	Ciliophora
OTU2128	0.02	20	b7	7.59	0.59	peripheral	<i>Protospathidium sp.</i>	95.4	Ciliophora
OTU21318	0.02	49	b2	9.48	4.31	connector	unclassified Dinophyceae	97	Dinophyceae
OTU21450	0.01	37	b4	0.23	0.15	peripheral	<i>Metschnikowia sp.</i>	99.2	Fungi
OTU21704	0.01	29	b0	0	0.41	peripheral	unclassified Ciliophora	95.1	Ciliophora
OTU21967	0.02	66	b15	0.57	1.96	peripheral	<i>Nephroselmis sp.</i>	91.4	Chlorophyta
OTU22133	0.01	14	b12	3.68	0.35	peripheral	unclassified Stramenopiles	85.9	Stramenopiles
OTU22140	0.05	9	b8	10.15	2.63	network_hub	<i>Metopion sp.</i>	92	Cercozoa
OTU22364	0.01	17	b4	2.06	0.16	peripheral	unclassified Dinophyceae	96.6	Dinophyceae
OTU22544	0.01	49	b10	0.98	1.18	peripheral	Chaetophorales	95.8	Chlorophyta
OTU22695	0.02	83	b1	4.29	2.65	connector	unclassified Alveolata	92.9	unclassified Alveolata
OTU22837	0.01	54	b3	1.12	0.16	peripheral	unclassified Stramenopiles	98.5	Stramenopiles
OTU22929	0.03	34	b1	5.64	1.7	peripheral	<i>Diophrys sp.</i>	85.9	Ciliophora
OTU23235	0.39	97	b13	3.36	0.95	peripheral	<i>Lagerheimia sp.</i>	98.9	Chlorophyta
OTU23310	0.01	9	b14	5.53	3.86	connector	Perkinsea 2	94.3	Perkinsea
OTU23383	0.03	66	b2	3.15	1.03	peripheral	<i>Woloszynskia sp.</i>	99.6	Dinophyceae
OTU23403	0.01	11	b8	1.73	0.98	peripheral	<i>Geminigera sp.</i>	86	Pyrenomonadales
OTU23615	4.39	100	b2	7.2	2.99	connector	unclassified Alveolata	99.2	unclassified Alveolata
OTU2382	0.33	100	b17	1.42	0.21	peripheral	Cercozoa 2	95.9	Cercozoa

OTU24036	0.02	17	b3	12.69	2.28	network_hub	unclassified Alveolata	97.7	unclassified Alveolata
OTU24161	0.01	29	b9	1.93	0.11	peripheral	unclassified Fungi	95.1	Fungi
OTU24265	0.03	57	b10	1.67	2.29	connector	Chytridiomycota	98.1	Fungi
OTU24457	0.02	31	b1	13.5	2.38	network_hub	unclassified Bicosoecida	82.2	Bicosoecida
OTU24502	0.01	29	b4	0.79	0.11	peripheral	<i>Semispithidium sp.</i>	93.9	Ciliophora
OTU24527	0.08	51	b5	1.64	0.12	peripheral	<i>Pascherina sp.</i>	97.7	Chlorophyta
OTU2455	0.02	66	b5	1.08	0.67	peripheral	Chytridiomycota	95.1	Fungi
OTU24583	0.05	83	b11	1.12	0.23	peripheral	Alveolata 1	94.6	Ciliophora
OTU24650	0.02	54	b6	0.99	1.43	peripheral	Cryptomycota CM1	97.3	Fungi
OTU24689	0.01	23	b1	13.17	1.88	module_hub	unclassified Bicosoecida	80.8	Bicosoecida
OTU24854	0.01	17	b13	4.3	7.44	connector	<i>Hanseniastora sp.</i>	93.9	Fungi
OTU24890	0.02	40	b6	2.98	3.21	connector	unclassified Alveolata	91.3	unclassified Alveolata
OTU24948	0.04	17	b2	3.68	3.68	connector	Haptorida 1	99.6	Ciliophora
OTU24980	0.01	6	b15	4.33	0.79	peripheral	<i>Ambispora sp.</i>	74.4	Fungi
OTU24991	0.13	74	b11	4.61	0.21	peripheral	Alveolata 1	96.9	Ciliophora
OTU25093	0.03	29	b12	2.8	3.79	connector	Mamiellales	90.6	Chlorophyta
OTU25267	0.02	20	b15	2.2	2.37	connector	unclassified Alveolata	95.8	unclassified Alveolata
OTU25317	0.05	94	b11	0.08	0.31	peripheral	unclassified Dinophyceae	98.1	Dinophyceae
OTU25493	0.04	80	b2	2.95	1.29	peripheral	unclassified Alveolata	98.1	unclassified Alveolata
OTU25743	0.02	43	b1	8.62	1.54	peripheral	Cryptomycota CM1	96.2	Fungi
OTU25781	0.06	17	b4	0.63	0.1	peripheral	Chytridiomycota	92.5	Fungi
OTU26136	0.04	20	b3	3.1	1.66	peripheral	MAST-4	89	MAST-4
OTU26230	0.01	6	b19	3.87	2.37	connector	<i>Saitoella sp.</i>	79.2	Fungi
OTU26253	0.04	43	b11	3.49	0.14	peripheral	Alveolata 1	92	Ciliophora
OTU26648	0.14	94	b20	1.11	0.57	peripheral	<i>Epispithidium sp.</i>	97	Ciliophora
OTU26991	0.03	23	b15	2.25	5.56	connector	unclassified Viridiplantae	62.3	Viridiplantae
OTU27086	0.08	91	b5	0.67	1.07	peripheral	Tremellomycetes	94	Fungi
OTU27143	0	6	b13	10.17	0.99	module_hub	Dikarya	93.2	Fungi
OTU27236	0.03	60	b4	0.61	0.18	peripheral	<i>Arcuospithidium sp.</i>	96.6	Ciliophora
OTU2724	0.21	94	b2	10.53	4.94	network_hub	<i>Pfiesteria sp.</i>	97.4	Dinophyceae
OTU27369	0.02	31	b2	3.68	1.93	peripheral	Cercozoa 1	96.6	Cercozoa
OTU27551	0.03	37	b6	1.59	2.89	connector	<i>Eimeria sp.</i>	88.3	Apicomplexa
OTU27552	0.01	37	b6	2.12	3.62	connector	<i>Glomus sp.</i>	90.3	Fungi
OTU28105	0	6	b4	1.99	0.11	peripheral	unclassified Dinophyceae	94	Dinophyceae
OTU28150	0.01	20	b5	5.43	0.18	peripheral	<i>Chlamydomonas sp.</i>	96.6	Chlorophyta
OTU28251	0.03	69	b1	9.78	3.13	connector	unclassified Fungi	99.6	Fungi
OTU28409	0.07	94	b14	2.22	2.41	connector	<i>Gymnodinium sp.</i>	92.1	Dinophyceae
OTU29008	0.01	34	b5	1.66	0.13	peripheral	Tremellomycetes	94.3	Fungi
OTU29109	0.06	63	b1	1.16	1.73	peripheral	Sarcocystidae	92.1	Apicomplexa
OTU29114	0.02	23	b4	0.7	0.84	peripheral	Cryptomycota	87.5	Fungi
OTU29234	0.08	17	b0	1.79	0.07	peripheral	Alveolata 1	92	Ciliophora
OTU29329	0.01	37	b0	0	0.35	peripheral	Cryptomycota LKM11	94.3	Fungi
OTU29344	0.04	60	b11	2.83	0.07	peripheral	Alveolata 1	96.9	Ciliophora
OTU29553	0.05	51	b0	0.24	0.08	peripheral	unclassified Rhizaria	93.7	unclassified Rhizaria
OTU29563	0.01	40	b16	0.85	0.74	peripheral	<i>Symbiodinium sp.</i>	91.7	Dinophyceae
OTU29696	0.03	74	b0	0.02	0.68	peripheral	Spathidiidae	97.3	Ciliophora
OTU29949	0.01	46	b5	0.68	0.47	peripheral	Cercozoa 1	92.5	Cercozoa
OTU30147	0.03	57	b14	1.14	2.4	connector	<i>Gigaspora sp.</i>	89.8	Fungi
OTU30154	0.05	69	b5	5.34	0.15	peripheral	<i>Chlamydomonas sp.</i>	98.9	Chlorophyta
OTU30168	0.08	94	b6	4.71	4.08	connector	<i>Amylovorax sp.</i>	89.4	Ciliophora
OTU30350	0.2	100	b16	0.02	0.12	peripheral	Cercozoa 1	94.4	Cercozoa
OTU30392	0.01	9	b1	5.68	2.69	connector	unclassified Alveolata	93.6	unclassified Alveolata
OTU30415	0.02	11	b3	10.95	2.47	network_hub	Saprolegniaceae	90.6	Oomycetes
OTU30503	0.01	49	b10	0.54	1.72	peripheral	<i>Protogastrostyla sp.</i>	98.1	Ciliophora
OTU30593	0.13	29	b19	3.79	6.67	connector	Chytridiomycota	95.8	Fungi
OTU30633	0.01	43	b7	2.35	0.22	peripheral	<i>Protospithidium sp.</i>	90.7	Ciliophora
OTU30704	0.01	17	b2	1.06	2.74	connector	Haptorida	96.2	Ciliophora
OTU30855	0.02	54	b7	1.73	3.36	connector	unclassified Alveolata	94	unclassified Alveolata
OTU30866	0.02	31	b2	5.54	5.91	connector	<i>Karenia sp.</i>	97.7	Dinophyceae
OTU30906	0.01	31	b9	0.73	0.89	peripheral	<i>Rhizidium sp.</i>	90.4	Fungi
OTU30992	0.01	11	b10	1.01	0.12	peripheral	<i>Spathidium sp.</i>	90.7	Ciliophora
OTU31035	0.01	31	b0	0	0.27	peripheral	Chytridiomycota	96.6	Fungi
OTU31161	0.04	69	b1	13.26	4.11	network_hub	<i>Gremmeniella sp.</i>	97.3	Fungi
OTU31306	0.03	43	b2	9.24	4.93	connector	<i>Peridinium sp.</i>	97	Dinophyceae
OTU31347	0.01	9	b3	6.21	0.11	peripheral	Cercozoa 1	98.1	Cercozoa
OTU31536	0.15	86	b5	1.17	1.31	peripheral	Chytridiomycota	93.8	Fungi
OTU31680	0	9	b10	3.72	0.17	peripheral	unclassified Fungi	92.5	Fungi
OTU31805	0.01	17	b9	1.98	0.04	peripheral	<i>Cryptococcus sp.</i>	93.2	Fungi
OTU31959	0.07	97	b6	2.11	3.04	connector	<i>Gigaspora sp.</i>	89.8	Fungi
OTU31963	0.01	29	b14	2.22	3.18	connector	Cryptomycota CM1	92.8	Fungi
OTU32153	1.03	100	b17	1.07	1.4	peripheral	unclassified Fungi	90.5	Fungi
OTU32280	0.01	26	b16	1.71	0.43	peripheral	unclassified Alveolata	93.6	unclassified Alveolata
OTU32294	0.02	43	b1	8.26	4.28	connector	Syndiniales	91.8	Dinophyceae
OTU32427	0	14	b14	4.13	4.17	connector	Perkinsea 1	95.8	Perkinsea
OTU32680	0.07	66	b14	0.12	0.41	peripheral	<i>Peridinium sp.</i>	98.1	Dinophyceae
OTU32918	0.01	29	b3	5.29	0.78	peripheral	unclassified Stramenopiles	95.1	Stramenopiles
OTU32995	0.03	89	b2	2.74	1.44	peripheral	<i>Pfiesteria sp.</i>	97.7	Dinophyceae
OTU33160	0.05	83	b1	3.99	0.83	peripheral	unclassified Fungi	99.2	Fungi
OTU33185	0.01	17	b12	3.02	3.12	connector	Cryptomycota CM1	93.9	Fungi
OTU33387	0.02	34	b20	0.92	0.35	peripheral	<i>Epispithidium sp.</i>	97.9	Ciliophora
OTU33538	0.01	26	b3	7.77	4.55	connector	<i>Blastodinium sp.</i>	96.6	Dinophyceae
OTU33787	0.01	29	b18	1.77	2.55	connector	<i>Geosiphon sp.</i>	87.6	Fungi
OTU33854	0.01	3	b19	1.05	0.83	peripheral	<i>Eimeria sp.</i>	82.6	Apicomplexa
OTU3406	0.04	74	b10	0.11	0.19	peripheral	Cryptomycota CM1	92.4	Fungi
OTU34105	0.22	89	b1	1.16	0.88	peripheral	Chaetophorales	94.7	Chlorophyta
OTU34164	0.03	51	b10	1.71	1.25	peripheral	unclassified Fungi	94.3	Fungi
OTU34287	0.03	83	b5	0.05	0.41	peripheral	unclassified Rhizaria	92.9	unclassified Rhizaria
OTU34333	0.01	11	b3	12.62	0.99	module_hub	Rhizophydiales	88.3	Fungi
OTU34633	0.54	49	b3	0.05	0.05	peripheral	unclassified Stramenopiles	99.6	Stramenopiles
OTU34674	0.25	94	b1	0.26	0.34	peripheral	<i>Spathidium sp.</i>	95.8	Ciliophora
OTU35041	2.26	100	b6	5.3	4.21	connector	Cryptomycota CM1	99.2	Fungi
OTU35093	0.01	29	b9	0.41	0.39	peripheral	unclassified Fungi	91.2	Fungi
OTU35462	0.02	20	b1	12.77	1.7	module_hub	unclassified Bicosoecida	83.3	Bicosoecida
OTU35588	0.04	40	b1	10.87	2.3	network_hub	<i>Schizangiella sp.</i>	89.9	Fungi
OTU35839	0.05	37	b0	0.86	0.82	peripheral	unclassified Stramenopiles	95.9	Stramenopiles
OTU36092	0.11	66	b3	6.66	2.02	connector	<i>Aulacoseira sp.</i>	99.6	Bacillariophyta
OTU36330	0.06	94	b1	14.89	2.66	network_hub	Agaricomycetes	99.2	Fungi
OTU36370	0	3	b5	0.77	0.48	peripheral	<i>Colpodidium sp.</i>	90.8	Ciliophora

OTU36502	0.02	60	b14	0.76	1.03	peripheral	Perkinsea 1	99.2	Perkinsea
OTU36614	0.78	97	b0	0.35	1.17	peripheral	<i>Chrysonebula</i> sp.	98.1	Chrysophyceae
OTU36769	7.46	100	b11	5.46	0.27	peripheral	Alveolate 1	97.7	Ciliophora
OTU36915	0.04	77	b5	1.73	0.5	peripheral	Dunaliellaceae	89.4	Chlorophyta
OTU37087	0.01	20	b5	0.1	0.15	peripheral	<i>Choricystis</i> sp.	97	Chlorophyta
OTU37710	0.01	14	b7	3.03	3.28	connector	Trachelidae	97.7	Ciliophora
OTU37719	0.03	31	b2	7.53	4	connector	unclassified Ciliophora	93.1	Ciliophora
OTU37915	0.02	43	b17	0.92	0.06	peripheral	Chytridiomycota	94.3	Fungi
OTU37989	0.01	37	b12	0.9	0.25	peripheral	Chytridiomycota	91	Fungi
OTU38254	0.01	20	b7	3.47	5.1	connector	<i>Gigaspora</i> sp.	93.3	Fungi
OTU38284	0.01	37	b2	1.75	1.13	peripheral	unclassified Alveolata	96.2	unclassified Alveolata
OTU38404	0.02	54	b10	1.56	0.35	peripheral	unclassified Dinophyceae	94	Dinophyceae
OTU38437	0.02	63	b20	0.64	1.49	peripheral	Chlamydomonadaceae	91.4	Chlorophyta
OTU3844	0.01	11	b3	2.99	1.26	peripheral	<i>Pyramimonas</i> sp.	91.7	Chlorophyta
OTU38638	0	11	b7	7.58	0.63	peripheral	unclassified Stramenopiles	95.1	Stramenopiles
OTU38647	0.01	9	b3	4.81	1.34	peripheral	<i>Chrysosaccus</i> sp.	93.6	Chrysophyceae
OTU38930	0.09	60	b2	4.8	4.19	connector	Cryptomycota LKM11	95.4	Fungi
OTU39067	0.04	51	b12	2.82	2.3	connector	<i>Paraglomus</i> sp.	89.1	Fungi
OTU3912	0.02	43	b3	0.35	0.27	peripheral	Chytridiomycota	92.9	Fungi
OTU3929	0.05	69	b6	6.18	6.81	connector	Cryptomycota CM1	98.5	Fungi
OTU3931	0.01	34	b9	0.13	0.19	peripheral	<i>Uronychia</i> sp.	92.1	Ciliophora
OTU39330	0.03	31	b10	2.9	2.24	connector	<i>Symbiodinium</i> sp.	90.3	Dinophyceae
OTU39477	0	6	b12	5.31	1.29	peripheral	unclassified Rhizaria	91.4	unclassified Rhizaria
OTU39617	0.02	57	b1	9.55	4.9	connector	Cryptomycota	97	Fungi
OTU39622	0.04	17	b1	15.81	2.71	network_hub	Cryptomycota	85.6	Fungi
OTU39639	0.02	34	b1	4.57	1.96	peripheral	unclassified Bicosoecida	83	Bicosoecida
OTU39683	0.01	17	b2	1.75	2.04	connector	Haptorida 1	97.7	Ciliophora
OTU39693	1.74	100	b4	0.63	0.29	peripheral	Thalassiosiraceae	99.6	Bacillariophyta
OTU39880	0	6	b7	8.44	0.81	peripheral	<i>Cyclonexis</i> sp.	91.4	Chrysophyceae
OTU40151	0.02	9	b14	2.05	0.42	peripheral	<i>Ambomucor</i> sp.	74.8	Fungi
OTU40294	0.01	9	b7	10.72	1.23	module_hub	unclassified Dinophyceae	96.6	Dinophyceae
OTU40328	0.01	17	b11	1.08	0.12	peripheral	Alveolate 1	95	Ciliophora
OTU40557	0.04	63	b6	2.58	3.98	connector	unclassified Chlorophyta	98.1	Chlorophyta
OTU40574	0.01	31	b4	0.35	0.29	peripheral	<i>Semispithidium</i> sp.	95.8	Ciliophora
OTU40576	0.01	37	b8	1.83	2.27	connector	<i>Peridinium</i> sp.	90.6	Dinophyceae
OTU41110	0.3	100	b1	9.83	6.05	connector	Cercozoa 2	96.3	Cercozoa
OTU41190	0.06	83	b6	2.5	3.25	connector	<i>Gigaspora</i> sp.	91.8	Fungi
OTU41196	0.74	100	b4	1.98	0.23	peripheral	<i>Arcuospathidium</i> sp.	98.1	Ciliophora
OTU41271	0.05	69	b1	7.3	4.04	connector	<i>Plasmodiophora</i> sp.	87.5	Cercozoa
OTU41324	0.06	80	b5	1.02	0.79	peripheral	Chytridiomycota	95.9	Fungi
OTU41597	0.02	34	b20	0.04	0.12	peripheral	<i>Pterosperma</i> sp.	88.5	Chlorophyta
OTU4194	0.01	14	b3	9.86	0.88	peripheral	unclassified Chrysophyceae	94.4	Chrysophyceae
OTU41946	0.02	31	b3	0.25	0.66	peripheral	unclassified Stramenopiles	96.2	Stramenopiles
OTU42187	0.14	100	b4	4.76	0.17	peripheral	<i>Semispithidium</i> sp.	98.1	Ciliophora
OTU42269	0.04	74	b20	1.16	0.28	peripheral	<i>Epispithidium</i> sp.	95.9	Ciliophora
OTU42452	0.02	43	b1	8.39	5.33	connector	<i>Microstroma</i> sp.	94.7	Fungi
OTU4261	0.01	43	b0	0.06	0.23	peripheral	<i>Plasmodiophora</i> sp.	89.7	Cercozoa
OTU42667	0.01	11	b1	5.04	3.29	connector	<i>Raabena</i> sp.	91.3	Ciliophora
OTU42697	0.46	80	b7	5.84	8.07	connector	<i>Cyclonexis</i> sp.	95.1	Chrysophyceae
OTU42788	0.02	37	b8	0.99	2.29	connector	<i>Paraglomus</i> sp.	88	Fungi
OTU43007	0.05	43	b11	4.07	0.13	peripheral	Alveolate 1	97.7	Ciliophora
OTU43015	0.01	14	b5	0.59	0.03	peripheral	unclassified Fungi	94.7	Fungi
OTU4310	0.09	37	b18	1.96	7.66	connector	<i>Bullera</i> sp.	87.7	Fungi
OTU4316	0.03	60	b3	0.37	0.42	peripheral	unclassified Stramenopiles	97	Stramenopiles
OTU43430	0.01	20	b3	7.75	2.69	connector	unclassified Alveolata	97.3	unclassified Alveolata
OTU43719	0.01	31	b8	2.4	2.87	connector	Cercozoa 2	94.3	Cercozoa
OTU43836	0	11	b10	2.87	0.26	peripheral	Chytridiomycota	96.6	Fungi
OTU44283	0.22	83	b2	7.36	3.42	connector	Cercozoa 1	97.7	Cercozoa
OTU44378	0.78	100	b17	1	0.79	peripheral	<i>Nannochloropsis</i> sp.	99.6	Eustigmatophyceae
OTU44745	0.02	77	b0	0.01	1.04	peripheral	Chytridiomycota	94.4	Fungi
OTU44833	0.01	34	b2	3.24	2.44	connector	Cercozoa 1	93.7	Cercozoa
OTU44875	0.04	66	b11	1.06	0.18	peripheral	Alveolate 1	92.7	Ciliophora
OTU44946	0.02	17	b8	9.41	3.93	connector	unclassified Rhizaria	95.5	unclassified Rhizaria
OTU44957	0.79	100	b12	3.44	5.14	connector	Alveolate 1	98.1	Ciliophora
OTU44985	0	6	b7	2.71	0.09	peripheral	unclassified Dinophyceae	97.4	Dinophyceae
OTU45044	0.03	20	b2	7.12	7.15	connector	<i>Chrysosphaerella</i> sp.	96.6	Synurophyceae
OTU45084	0.01	9	b3	5.06	0.22	peripheral	Cercozoa 1	93.3	Cercozoa
OTU45177	0.01	31	b3	3.3	1.97	peripheral	<i>Pseudoscourfieldia</i> sp.	89.7	Chlorophyta
OTU45351	0.03	71	b1	0.74	1.98	peripheral	<i>Enchelyodon</i> sp.	96.2	Ciliophora
OTU45378	0.01	11	b7	10.86	1.07	module_hub	unclassified Dinophyceae	95.9	Dinophyceae
OTU45478	0	3	b8	9.81	0.47	peripheral	Cercozoa 1	89.3	Cercozoa
OTU45514	0.56	54	b7	9.9	2.34	connector	unclassified Dinophyceae	98.9	Dinophyceae
OTU45600	0.13	89	b2	1.52	1.84	peripheral	<i>Didinium</i> sp.	100	Ciliophora
OTU45652	0.03	66	b11	0.04	0.27	peripheral	Cercozoa 1	99.2	Cercozoa
OTU45711	0.77	100	b5	1.76	0.81	peripheral	unclassified Fungi	99.2	Fungi
OTU4581	0.58	94	b10	4.19	1.87	peripheral	Cercozoa 1	88.9	Cercozoa
OTU46037	0.01	23	b9	2.78	0.11	peripheral	Basidiomycota	89.4	Fungi
OTU46247	0.02	29	b1	9.3	3.51	connector	Sarcosomataceae	89.4	Fungi
OTU46336	0.01	9	b13	9.02	1.76	peripheral	<i>Bandia</i> sp.	91.7	Ciliophora
OTU46617	0.01	9	b3	8.59	0.34	peripheral	Alveolate 1	92.3	Ciliophora
OTU46767	0.49	100	b4	1.58	0.31	peripheral	Haptorida	97.7	Ciliophora
OTU46895	1.79	100	b9	0.51	0.38	peripheral	unclassified Stramenopiles	98.1	Stramenopiles
OTU46993	1.83	100	b16	1.64	0.96	peripheral	unclassified Alveolata	97	unclassified Alveolata
OTU47170	0	9	b0	1.61	0.04	peripheral	Alveolate 1	90.1	Ciliophora
OTU47520	0.02	43	b6	1.3	1.8	peripheral	Cercozoa 1	83.5	Cercozoa
OTU47642	0.02	63	b2	0.12	0.1	peripheral	<i>Gigaspora</i> sp.	92.9	Fungi
OTU47703	0.01	29	b12	0.35	0.67	peripheral	<i>Hematodinium</i> sp.	87.3	Dinophyceae
OTU47759	0.01	6	b15	3.09	2.5	connector	<i>Vahlkampfia</i> sp.	81.2	Schizopyrenida
OTU47784	0.31	100	b4	2.45	0.21	peripheral	<i>Semispithidium</i> sp.	94.7	Ciliophora
OTU47889	0.05	43	b18	1.8	0.49	peripheral	Chytridiomycota	87.1	Fungi
OTU47973	0.03	77	b6	0.11	0.31	peripheral	unclassified Bicosoecida	98.5	Bicosoecida
OTU48080	0.05	66	b2	6.98	3.55	connector	unclassified Dinophyceae	97.4	Dinophyceae
OTU48386	0.04	51	b2	9	4.42	connector	unclassified Dinophyceae	96.6	Dinophyceae
OTU48579	0.02	40	b0	0.2	1.39	peripheral	unclassified Fungi	88.3	Fungi
OTU48842	0.02	54	b8	0.2	0.52	peripheral	Scenedesmeaceae	95.9	Chlorophyta
OTU49053	0.03	63	b14	0.79	2.47	connector	Cryptomycota CM1	96.2	Fungi
OTU49244	0.17	100	b1	6.68	3.04	connector	unclassified Alveolata	98.1	unclassified Alveolata
OTU4925	0.03	54	b2	2.83	3.95	connector	Saprolegniaceae	99.6	Oomycetes

OTU49267	0.01	14	b1	5.79	2.65	connector	unclassified Alveolata	92.8	unclassified Alveolata
OTU49428	0.02	89	b1	1.09	0.56	peripheral	Chlorellaceae	98.9	Chlorophyta
OTU49590	0.02	51	b7	0.45	1.3	peripheral	<i>Trachelophyllum sp.</i>	93.9	Ciliophora
OTU49624	0.02	34	b9	2.53	0.21	peripheral	<i>Hannaella sp.</i>	89.4	Fungi
OTU49839	0.02	40	b11	0.2	0.17	peripheral	Alveolate 1	94.3	Ciliophora
OTU49866	0.16	71	b11	4.29	0.15	peripheral	Alveolate 1	96.9	Ciliophora
OTU50185	0.15	97	b5	5.98	0.24	peripheral	<i>Chlamydomonas sp.</i>	95.1	Chlorophyta
OTU50309	0.04	51	b1	12.57	1.74	module_hub	unclassified Perkinsea	85.7	Perkinsea
OTU50326	0.04	40	b14	3.74	3.33	connector	Chytridiomycetes	98.1	Fungi
OTU50400	0.39	77	b4	0.2	0.83	peripheral	<i>Peridinium sp.</i>	98.5	Dinophyceae
OTU50665	0.05	40	b13	6.92	1.32	peripheral	<i>Kofoidinium sp.</i>	89.5	Dinophyceae
OTU50783	0.02	40	b7	3.45	0.81	peripheral	Haptorida	96.2	Ciliophora
OTU50831	0.01	14	b7	11.13	1.36	module_hub	unclassified Alveolata	98.1	unclassified Alveolata
OTU50998	0.05	54	b9	1.84	0.25	peripheral	<i>Cyllamyces sp.</i>	87.2	Fungi
OTU51006	0.02	20	b1	13.05	3.45	network_hub	Cryptomycota	92.4	Fungi
OTU51237	0.01	26	b16	1.94	0.77	peripheral	unclassified Alveolata	93.2	unclassified Alveolata
OTU51301	0.01	26	b14	0.28	0.7	peripheral	Cryptophyta 4	84.1	Cryptophyta 4
OTU51326	0.04	31	b1	13.3	3.56	network_hub	<i>Raabena sp.</i>	90.9	Ciliophora
OTU51389	0.18	80	b1	17.31	3.58	network_hub	unclassified Fungi	99.6	Fungi
OTU51536	0.05	77	b3	1.11	0.22	peripheral	unclassified Stramenopiles	97.7	Stramenopiles
OTU51758	0.02	14	b7	11.12	3.3	network_hub	unclassified Alveolata	98.1	unclassified Alveolata
OTU52355	0.02	6	b8	7.56	0.59	peripheral	<i>Chytriomycetes sp.</i>	88.6	Fungi
OTU52371	0.05	20	b12	1.86	2.95	connector	unclassified Dinophyceae	91.4	Dinophyceae
OTU52426	0.53	97	b9	1.6	1.2	peripheral	Chytridiomycota	88.3	Fungi
OTU52452	0.11	77	b2	0.04	0.19	peripheral	<i>Spirotaenia sp.</i>	95.5	Streptophyta
OTU52631	0.25	100	b9	0.13	0.36	peripheral	Lagenidiales	99.6	Oomycetes
OTU52649	0.06	66	b11	2.4	0.23	peripheral	<i>Blepharocorys sp.</i>	92.4	Ciliophora
OTU52687	0.67	54	b18	2.07	9.61	connector	Chytridiomycota	95.1	Fungi
OTU52711	0.01	43	b5	0.44	0.54	peripheral	unclassified Chlorophyta	98.9	Chlorophyta
OTU52779	0	6	b3	6.28	0.09	peripheral	<i>Uroglana sp.</i>	95.9	Chrysophyceae
OTU5279	0.01	9	b8	7.24	3.33	connector	Perkinsea 2	88.7	Perkinsea
OTU53119	0.01	29	b4	0.43	0.2	peripheral	<i>Rhizophydium sp.</i>	90.2	Fungi
OTU53141	0.34	46	b0	0	0.09	peripheral	Chlamydomonadales	98.9	Chlorophyta
OTU53445	0.04	69	b2	9.03	3.75	connector	unclassified Dinophyceae	97.9	Dinophyceae
OTU53489	0.53	100	b2	0.18	0.48	peripheral	Chytridiomycota	99.6	Fungi
OTU53512	0.02	31	b1	16.19	2.13	network_hub	unclassified Bicosoecida	81.1	Bicosoecida
OTU53639	0.01	11	b8	9.58	3.42	connector	unclassified Rhizaria	95.1	unclassified Rhizaria
OTU53757	0.01	6	b3	7.61	0.14	peripheral	unclassified Chrysophyceae	95.1	Chrysophyceae
OTU54068	0.04	89	b9	0.06	0.15	peripheral	<i>Sistotrema sp.</i>	93.6	Fungi
OTU54240	0.21	57	b1	13.55	2.76	network_hub	unclassified Alveolata	94.3	unclassified Alveolata
OTU54811	0.04	63	b4	0.45	0.3	peripheral	<i>Balanidion sp.</i>	96.6	Ciliophora
OTU55663	0.72	97	b1	0.56	0.61	peripheral	unclassified Alveolata	99.2	unclassified Alveolata
OTU5569	0.01	11	b3	4.41	2.54	connector	unclassified Stramenopiles	97.4	Stramenopiles
OTU55779	0.01	20	b1	9.94	3.69	connector	Cryptomycota	94.7	Fungi
OTU55910	0	11	b14	2.69	1.69	peripheral	unclassified Fungi	91.4	Fungi
OTU55923	0	3	b8	9.81	0.47	peripheral	Cercozoa 2	89.1	Cercozoa
OTU56078	0	17	b17	0.87	0.23	peripheral	Cercozoa 2	86.8	Cercozoa
OTU56181	0.03	34	b2	4.69	2.36	connector	unclassified Dinophyceae	89.1	Dinophyceae
OTU56337	0.01	14	b3	5.07	0.08	peripheral	unclassified Alveolata	91	unclassified Alveolata
OTU56391	0.04	51	b16	2.58	2.57	connector	unclassified Alveolata	95.5	unclassified Alveolata
OTU56415	0.03	60	b5	0.53	0.64	peripheral	Chytridiomycota	100	Fungi
OTU56483	0.01	20	b5	0.65	0.03	peripheral	unclassified Fungi	96.2	Fungi
OTU56528	0.03	31	b9	1.12	0.2	peripheral	Chytridiomycota	93.9	Fungi
OTU56729	0.02	37	b1	11.26	4.69	network_hub	<i>Plasmodiophora sp.</i>	90.1	Cercozoa
OTU56781	0.03	26	b9	2.64	0.09	peripheral	<i>Basidiobolus sp.</i>	83.2	Fungi
OTU56795	0.07	94	b3	0.1	0.28	peripheral	Cercozoa 2	93	Cercozoa
OTU56860	0.02	60	b4	1.28	0.46	peripheral	<i>Semispithidium sp.</i>	96.6	Ciliophora
OTU56915	0.01	17	b1	15.88	1.7	module_hub	unclassified Bicosoecida	80	Bicosoecida
OTU5696	0.01	3	b8	9.81	0.47	peripheral	<i>Plasmodiophora sp.</i>	87.9	Cercozoa
OTU57082	0.08	17	b1	1.43	1.03	peripheral	Cercozoa 1	93.3	Cercozoa
OTU57254	0.37	97	b7	2.39	2.77	connector	Chytridiomycota	93.5	Fungi
OTU57281	0.02	14	b1	4.94	4.8	connector	Chytridiomycota	91.7	Fungi
OTU57556	0.08	100	b3	0.05	0.31	peripheral	unclassified Fungi	99.6	Fungi
OTU57607	0.02	46	b1	9.33	1.25	peripheral	unclassified Fungi	98.1	Fungi
OTU57723	0	14	b8	3.41	0.32	peripheral	<i>Fragilaria sp.</i>	99.6	Bacillariophyta
OTU58118	0.03	20	b6	2.71	3.85	connector	unclassified Fungi	93.2	Fungi
OTU5827	0.05	60	b1	14.54	2.96	network_hub	unclassified Bicosoecida	83	Bicosoecida
OTU58429	0.01	23	b12	4.11	3.57	connector	<i>Lobulomyces sp.</i>	91.7	Fungi
OTU58447	0.02	20	b7	0.09	0.17	peripheral	<i>Ochromonas sp.</i>	97.4	Chrysophyceae
OTU58603	0.01	26	b2	6.19	3.71	connector	unclassified Dinophyceae	99.2	Dinophyceae
OTU58707	0.04	74	b1	0.08	0.14	peripheral	<i>Scherffelia sp.</i>	100	Chlorophyta
OTU58743	0.03	54	b9	0.13	1.13	peripheral	Chytridiomycota	87.5	Fungi
OTU58805	0.01	20	b18	1.01	0.55	peripheral	<i>Hygrocybe sp.</i>	62.5	Fungi
OTU58836	0	9	b19	4.9	5.62	connector	<i>Candida sp.</i>	91.3	Fungi
OTU59130	0.93	100	b12	0.7	1.18	peripheral	Cryptomycota LKM11	97.7	Fungi
OTU59140	0.05	43	b4	3.47	0.24	peripheral	<i>Semispithidium sp.</i>	95.1	Ciliophora
OTU59231	0.01	43	b0	0.02	0.26	peripheral	Cercozoa 1	92.1	Cercozoa
OTU59307	0.02	57	b0	0.28	0.64	peripheral	unclassified Dinophyceae	92.1	Dinophyceae
OTU59535	0.03	60	b1	6.17	1.97	peripheral	Cryptomycota	91.3	Fungi
OTU59597	0.01	14	b16	1.91	0.41	peripheral	unclassified Alveolata	95.1	unclassified Alveolata
OTU59816	0.01	9	b0	1.08	0.23	peripheral	unclassified Stramenopiles	95.9	Stramenopiles
OTU60288	0.02	46	b16	1.16	0.21	peripheral	Perkinsea 1	92.5	Perkinsea
OTU60340	0.02	26	b12	2.27	1.47	peripheral	Alveolate 1	97.7	Ciliophora
OTU60478	0.05	60	b11	0.36	0.41	peripheral	Rhizophydium clade	92.6	Fungi
OTU60578	0.02	51	b5	2.42	0.38	peripheral	Chytridiomycota	94.7	Fungi
OTU6087	0.01	31	b6	1.66	3.93	connector	unclassified Alveolata	89.4	unclassified Alveolata
OTU61006	0.04	69	b20	1.09	0.57	peripheral	<i>Epispithidium sp.</i>	95.8	Ciliophora
OTU61500	0.01	40	b10	1.79	0.26	peripheral	<i>Sabulodinium sp.</i>	93.2	Dinophyceae
OTU61517	0.03	26	b1	16.45	2.12	network_hub	unclassified Bicosoecida	81	Bicosoecida
OTU61585	0.04	74	b1	6.59	2.53	connector	Cercozoa 2	92.2	Cercozoa
OTU61890	0.01	9	b12	4.8	3.67	connector	<i>Ceriporia sp.</i>	77	Fungi
OTU62093	0.01	26	b12	2.52	2.62	connector	<i>Candida sp.</i>	94.7	Fungi
OTU62122	0.01	9	b8	7.51	3.78	connector	unclassified Rhizaria	93.3	unclassified Rhizaria
OTU62195	0.07	89	b6	4.09	7.98	connector	Amylovoracidae	87.8	Ciliophora
OTU6228	0.05	83	b5	4.45	0.35	peripheral	<i>Stigeoclonium sp.</i>	94.3	Chlorophyta
OTU62641	0.01	34	b8	2.15	0.99	peripheral	<i>Cosmarium sp.</i>	92.2	Streptophyta
OTU62964	0.02	49	b2	8.38	4	connector	unclassified Dinophyceae	97	Dinophyceae
OTU63063	0.02	29	b2	8.58	4.74	connector	unclassified Stramenopiles	91.7	Stramenopiles

OTU63251	0.1	89	b4	3.5	0.11	peripheral	<i>Semispithidium sp.</i>	96.3	Ciliophora
OTU63381	0.01	37	b1	4.01	2.99	connector	Cryptomycota	95.1	Fungi
OTU63382	0.01	9	b13	9.73	1	peripheral	<i>Triparticalcar sp.</i>	89.8	Fungi
OTU63696	0	17	b15	1.75	1.92	peripheral	unclassified Alveolata	95.1	unclassified Alveolata
OTU6403	0.01	11	b1	2.24	0.6	peripheral	Chytridiomycota	93.9	Fungi
OTU6405	0.03	80	b1	7.04	1.37	peripheral	Exobasidiomycetes	95.1	Fungi
OTU64074	0.19	97	b3	0.19	0.4	peripheral	unclassified Apicomplexa	99.6	Apicomplexa
OTU64186	0.05	83	b1	3.25	0.52	peripheral	unclassified Apicomplexa	98.1	Apicomplexa
OTU64284	0.04	63	b1	14.02	3.55	network_hub	Pleosporales	100	Fungi
OTU64423	0.02	77	b0	0.01	0.57	peripheral	<i>Woloszynskia sp.</i>	94.7	Dinophyceae
OTU64469	0.01	11	b1	3.62	0.32	peripheral	Perkinsea 1	90.6	Perkinsea
OTU64666	0	14	b0	0.36	0.17	peripheral	<i>Polydiniella sp.</i>	92.7	Ciliophora
OTU64686	0.03	37	b9	0.79	0.13	peripheral	unclassified Ciliophora	99.2	Ciliophora
OTU64740	0.31	94	b11	3.49	0.69	peripheral	Alveolate 1	92	Ciliophora
OTU64774	0.72	100	b4	0.24	0.6	peripheral	Chytridiomycota	89.8	Fungi
OTU64841	0.02	26	b1	14.05	2.92	network_hub	Cryptomycota	92.4	Fungi
OTU64848	0.02	46	b1	0.33	0.22	peripheral	Microbotryomycetes	99.6	Fungi
OTU64878	0.11	91	b1	2.02	1.89	peripheral	<i>Balantidium sp.</i>	97.3	Ciliophora
OTU64900	0.96	100	b5	0.94	0.78	peripheral	Fragilariaceae	99.3	Bacillariophyta
OTU6493	0.01	11	b14	5.41	4.18	connector	Perkinsea 1	94.7	Perkinsea
OTU65023	0.01	3	b4	0.86	0.08	peripheral	<i>Rhizoclostridium sp.</i>	84.2	Fungi
OTU65037	0.01	3	b3	2.84	0.46	peripheral	unclassified Chrysophyceae	97.7	Chrysophyceae
OTU65126	0.01	9	b7	6.73	1.29	peripheral	unclassified Chrysophyceae	92.2	Chrysophyceae
OTU6522	1.15	89	b1	16.6	2.5	network_hub	Chytridiomycota	81	Fungi
OTU65462	0.02	49	b4	0.26	0.09	peripheral	Thalassiosiraceae	95.1	Bacillariophyta
OTU65552	0.07	69	b1	4.05	2.25	connector	<i>Plasmodiophora sp.</i>	89.7	Cercozoa
OTU6563	0.13	49	b10	1.59	0.45	peripheral	Stichotrichia	98.9	Ciliophora
OTU65685	0.01	43	b6	0.63	0.55	peripheral	unclassified Fungi	85.7	Fungi
OTU65691	0.02	40	b12	1.09	0.63	peripheral	unclassified Fungi	88.4	Fungi
OTU6584	0.13	86	b11	2.05	0.95	peripheral	Spathidiidae	95.4	Ciliophora
OTU66161	0.02	46	b2	7.4	3.62	connector	unclassified Dinophyceae	97.4	Dinophyceae
OTU66244	0.02	57	b2	2.95	0.93	peripheral	unclassified Alveolata	97.4	unclassified Alveolata
OTU6638	0.02	37	b4	0.42	0.25	peripheral	unclassified Fungi	91.4	Fungi
OTU66458	0.18	89	b5	6.04	0.17	peripheral	<i>Chlamydomonas sp.</i>	96.6	Chlorophyta
OTU66524	0.01	34	b6	1.01	2.89	connector	Perkinsea 2	85.2	Perkinsea
OTU66608	0.01	11	b13	5.98	1.27	peripheral	unclassified Dinophyceae	87.6	Dinophyceae
OTU66983	0.03	9	b3	13.52	1.18	module_hub	unclassified Ciliophora	95.8	Ciliophora
OTU67181	0.03	71	b5	0.46	0.39	peripheral	unclassified Chlorophyta	97.7	Chlorophyta
OTU67227	0.01	31	b4	1	0.07	peripheral	<i>Arcuospithidium sp.</i>	97	Ciliophora
OTU67365	0.7	100	b1	3.36	1.23	peripheral	unclassified Apicomplexa	99.2	Apicomplexa
OTU67373	0	11	b3	5.84	0.38	peripheral	<i>Aulacoseira sp.</i>	99.3	Bacillariophyta
OTU67646	0.01	11	b3	10.4	1.58	module_hub	unclassified Chrysophyceae	98.5	Chrysophyceae
OTU67878	0	9	b19	4.88	5.32	connector	Basidiomycota	94.7	Fungi
OTU67926	0.02	40	b9	2.3	0.11	peripheral	Hematodinium sp.	87.8	Dinophyceae
OTU68011	0.12	43	b10	0.92	0.13	peripheral	<i>Gymnophrys sp.</i>	93.6	Cercozoa
OTU68358	0.02	14	b8	1.44	1.39	peripheral	<i>Haliphthoros sp.</i>	93.3	Oomycetes
OTU68386	0.09	57	b9	1.07	0.08	peripheral	unclassified Fungi	95.1	Fungi
OTU68416	0.07	74	b1	3.26	3.79	connector	Sarcocystidae	92.5	Apicomplexa
OTU68569	0	3	b8	9.81	0.47	peripheral	Alveolate 1	90.8	Ciliophora
OTU68950	0.02	29	b5	0.36	0.84	peripheral	unclassified Chlorophyta	97.4	Chlorophyta
OTU68999	0.03	40	b11	2.11	0.06	peripheral	Alveolate 1	97.3	Ciliophora
OTU69116	0	17	b17	0.67	0.18	peripheral	Chytridiomycota	95.5	Fungi
OTU69179	0.29	89	b6	4.02	7.2	connector	Cryptomycota	93.2	Fungi
OTU69251	0.02	54	b6	3.02	4.4	connector	Cryptomycota CM1	95.8	Fungi
OTU69893	0.01	49	b0	0.18	0.18	peripheral	unclassified Dinophyceae	97	Dinophyceae
OTU70016	0.05	77	b4	3	0.15	peripheral	<i>Semispithidium sp.</i>	97.7	Ciliophora
OTU70868	0.01	17	b1	7.11	1.87	peripheral	Dothideomycetes	99.2	Fungi
OTU71258	0.01	9	b3	12.16	1.5	module_hub	unclassified Ciliophora	93.1	Ciliophora
OTU71310	0.01	6	b7	3.73	0.1	peripheral	unclassified Dinophyceae	95.1	Dinophyceae
OTU7163	1.33	100	b4	3.8	0.21	peripheral	Haptorida	98.5	Ciliophora
OTU71740	0.04	40	b1	12.17	3.82	network_hub	<i>Condylostoma sp.</i>	60.3	Ciliophora
OTU71930	0.1	97	b9	2.25	0.27	peripheral	<i>Amylavorax sp.</i>	89.8	Ciliophora
OTU71949	0.06	37	b10	2.98	4.32	connector	Colpodidiidae	90.5	Ciliophora
OTU7203	0.03	3	b1	2.11	0.13	peripheral	<i>Schizosaccharomyces sp.</i>	84	Fungi
OTU72074	0.01	23	b12	1.44	0.28	peripheral	Alveolate 1	97.3	Ciliophora
OTU7229	0.02	37	b1	1.52	0.81	peripheral	<i>Plasmodiophora sp.</i>	89.7	Cercozoa
OTU72406	0.23	23	b3	11.41	1.26	module_hub	unclassified Dinophyceae	84.6	Dinophyceae
OTU72482	0.09	57	b2	3.65	3.04	connector	unclassified Rhizaria	94.8	unclassified Rhizaria
OTU72590	0.02	54	b5	0.39	0.54	peripheral	<i>Cyllamyces sp.</i>	88.2	Fungi
OTU72656	0.04	54	b4	2.52	0.28	peripheral	<i>Balantidium sp.</i>	97.7	Ciliophora
OTU72937	0	14	b16	1.44	0.74	peripheral	unclassified Alveolata	91.7	unclassified Alveolata
OTU72947	0.03	49	b2	8.82	5.44	connector	<i>Pfiesteria sp.</i>	97.4	Dinophyceae
OTU72984	0.02	31	b19	1.16	2.73	connector	Cryptomycota	99.2	Fungi
OTU7328	0.27	29	b2	4.82	2.54	connector	Tracheliidae	99.6	Ciliophora
OTU73320	0.02	26	b1	13.23	2.32	network_hub	<i>Condylostoma sp.</i>	60.4	Ciliophora
OTU73478	0.09	91	b4	1.11	0.34	peripheral	Cercozoa 1	91	Cercozoa
OTU73582	0.01	14	b15	3	3.99	connector	Cercozoa 2	90.7	Cercozoa
OTU73882	0.01	11	b3	2.35	0.83	peripheral	Zygomycota 1	90.8	Fungi
OTU73927	0.04	9	b10	3.64	0.18	peripheral	unclassified Rhizaria	89.9	unclassified Rhizaria
OTU74356	0.08	57	b6	0.76	1.83	peripheral	<i>Gymnophrys sp.</i>	80.6	Cercozoa
OTU74388	0.01	37	b9	1.95	0.2	peripheral	Chytridiomycota	91.3	Fungi
OTU74521	0.03	17	b7	10.16	2.97	network_hub	unclassified Dinophyceae	97.8	Dinophyceae
OTU74572	0.04	60	b1	3.73	0.6	peripheral	unclassified Apicomplexa	97.4	Apicomplexa
OTU74646	0.02	37	b14	0.22	1.11	peripheral	unclassified Fungi	98.9	Fungi
OTU74648	0.66	100	b12	0.38	1.74	peripheral	<i>Desmodesmus sp.</i>	99.6	Chlorophyta
OTU74955	0.02	9	b3	13.1	1.19	module_hub	unclassified Dinophyceae	90.2	Dinophyceae
OTU75007	0.01	34	b17	0.41	0.12	peripheral	unclassified Alveolata	88.3	Ciliophora
OTU75159	0.01	11	b15	3	3.79	connector	<i>Amylavorax sp.</i>	90.1	Fungi
OTU75565	0.01	23	b4	0.56	0.05	peripheral	Cryptomycota	90.1	Fungi
OTU75649	0.02	51	b1	1.25	0.59	peripheral	<i>Semispithidium sp.</i>	96.6	Ciliophora
OTU75839	0.02	46	b5	0.79	0.12	peripheral	Chytridiomycota	95.5	Fungi
OTU76015	0.04	57	b1	10.75	2.77	network_hub	<i>Stigeoclonium sp.</i>	96.6	Chlorophyta
OTU76139	1.63	100	b17	0.76	0.91	peripheral	Perkinsea 1	97.4	Perkinsea
OTU76283	0.03	31	b9	2	0.08	peripheral	Chytridiomycota	98.1	Fungi
OTU76285	0.03	40	b3	8.01	2.99	connector	unclassified Chlorophyta	91.8	Chlorophyta
OTU76411	0.12	86	b17	0.69	1.63	peripheral	unclassified Rhizaria	97.4	unclassified Rhizaria
OTU76483	0.04	51	b5	5.34	0.12	peripheral	Chytridiomycota	93	Fungi
OTU76545	0.01	51	b12	1.57	3.18	connector	<i>Chlamydomonas sp.</i>	98.9	Chlorophyta
							<i>Schizosaccharomyces sp.</i>	86.5	Fungi

OTU76581	0.01	14	b19	3.11	9.25	connector	Chytridiomycota	95.8	Fungi
OTU76624	0.02	11	b0	0.17	1.07	peripheral	unclassified Oomycetes	93	Oomycetes
OTU76774	0.05	54	b16	3.15	1.61	peripheral	unclassified Alveolata	93.2	unclassified Alveolata
OTU76795	0.01	34	b2	2.72	1.81	peripheral	<i>Enchelys sp.</i>	98.9	Ciliophora
OTU76947	0.23	100	b2	4.42	1.35	peripheral	<i>Woloszynskia sp.</i>	98.9	Dinophyceae
OTU77130	0	6	b13	10.18	0.99	module_hub	<i>Gigaspora sp.</i>	89.5	Fungi
OTU77154	0.05	14	b1	0.05	0.06	peripheral	Chlamydomonadaceae	97.3	Chlorophyta
OTU77166	0.01	40	b5	0.36	0.12	peripheral	unclassified Fungi	95.5	Fungi
OTU77575	0.06	100	b10	0.18	0.66	peripheral	Chytridiomycota	95.9	Fungi
OTU77620	0.08	11	b19	3.49	3.34	connector	unclassified Fungi	81.4	Fungi
OTU77654	0.01	11	b8	7.2	3.35	connector	Chytridiomycota	92.9	Fungi
OTU77667	0.06	91	b2	2.71	0.85	peripheral	unclassified Alveolata	96.3	unclassified Alveolata
OTU77701	0.01	40	b10	0.31	1.12	peripheral	Cercozoa 1	96.2	Cercozoa
OTU77782	0.11	83	b15	0.72	5.02	connector	Cercozoa 1	91.7	Cercozoa
OTU77847	0.02	20	b8	5.87	3.95	connector	Alveolata 1	95.8	Ciliophora
OTU77913	0.01	26	b2	5.96	3.94	connector	unclassified Dinophyceae	94.4	Dinophyceae
OTU77935	0.01	11	b3	9.21	2.31	connector	<i>Cyclonexis sp.</i>	95.1	Chrysophyceae
OTU78036	0.01	49	b2	4	3.36	connector	Cercozoa 1	93.6	Cercozoa
OTU7819	0.25	100	b16	3.96	1.16	peripheral	Perkinsea 1	91.7	Perkinsea
OTU78212	0.04	51	b1	6.35	4.42	connector	Cryptomycota	92.4	Fungi
OTU78235	0.01	37	b0	0.09	0.54	peripheral	unclassified Alveolata	90.6	unclassified Alveolata
OTU78302	0.12	49	b3	11.18	4.97	network_hub	unclassified Dinophyceae	99.2	Dinophyceae
OTU78659	0	6	b15	4.16	1.22	peripheral	Chytridiomycota	72.4	Fungi
OTU78742	0.02	31	b1	10.96	2.31	network_hub	unclassified Alveolata	94.7	unclassified Alveolata
OTU78816	0.01	11	b7	6.85	3.87	connector	<i>Chrysophaerella sp.</i>	95.5	Synurophyceae
OTU79123	0.04	60	b6	1.72	4.11	connector	<i>Woloszynskia sp.</i>	91	Dinophyceae
OTU79241	0.01	14	b1	7.44	0.93	peripheral	unclassified Alveolata	94.4	unclassified Alveolata
OTU79410	0.06	66	b1	17.87	4.72	network_hub	Ascomycota	99.6	Fungi
OTU79609	0.01	6	b4	0.41	0.46	peripheral	<i>Choricystis sp.</i>	96.6	Chlorophyta
OTU7964	0.05	69	b10	0.38	1.15	peripheral	<i>Epispathidium sp.</i>	97.3	Ciliophora
OTU79708	0.01	46	b3	1.35	1.67	peripheral	unclassified Stramenopiles	97.4	Stramenopiles
OTU79789	0.01	23	b7	5.1	5.49	connector	Cryptomycota LKM11	95.1	Fungi
OTU79981	0.06	97	b15	0.32	0.86	peripheral	unclassified Fungi	97.3	Fungi
OTU8005	0.09	46	b6	6.36	8.97	connector	Cryptomycota	90.9	Fungi
OTU80106	0.01	11	b8	8.22	0.35	peripheral	<i>Uronychia sp.</i>	89.7	Ciliophora
OTU80182	0.83	83	b9	2.37	0.08	peripheral	MAST-7A	91	MAST-7
OTU80462	1.44	91	b1	0.91	1.91	peripheral	unclassified Alveolata	99.2	unclassified Alveolata
OTU80493	0.05	83	b6	4.95	7.55	connector	Cercozoa 1	96.3	Cercozoa
OTU80720	0.01	46	b0	0.08	0.32	peripheral	unclassified Dinophyceae	89.8	Dinophyceae
OTU80731	0.15	100	b8	6.66	5.57	connector	unclassified Rhizaria	93.6	unclassified Rhizaria
OTU80895	0.1	86	b8	0.17	0.84	peripheral	unclassified Ciliophora	92.7	Ciliophora
OTU80942	0.01	14	b2	2.08	1.08	peripheral	unclassified Dinophyceae	83.2	Dinophyceae
OTU80967	0.91	100	b6	6.64	7.95	connector	Cryptomycota	97.3	Fungi
OTU80989	0.04	46	b18	2.01	5.94	connector	<i>Woloszynskia sp.</i>	88.1	Dinophyceae
OTU80994	0	3	b17	0.94	0.03	peripheral	<i>Pyramimonas sp.</i>	88.8	Chlorophyta
OTU81115	0.02	37	b6	2.82	7.1	connector	<i>Glomus sp.</i>	90.3	Fungi
OTU81418	0	3	b14	3.23	0.79	peripheral	<i>Kofofidinium sp.</i>	89.5	Dinophyceae
OTU8144	0.01	9	b13	8.32	3.07	connector	<i>Novistrombidium sp.</i>	91	Ciliophora
OTU81816	0.12	97	b16	2.31	3.26	connector	unclassified Alveolata	93.2	unclassified Alveolata
OTU82355	0.03	11	b3	6.21	0.08	peripheral	Ochromonadaceae	97	Synurophyceae
OTU82387	0.03	43	b9	1.02	0.34	peripheral	unclassified Fungi	90.5	Fungi
OTU82469	0.02	69	b2	1.32	0.48	peripheral	unclassified Alveolata	98.1	unclassified Alveolata
OTU82555	0.06	77	b11	1.87	0.33	peripheral	<i>Blepharocorys sp.</i>	92.4	Ciliophora
OTU82595	0.1	60	b8	3.42	2.31	connector	<i>Blepharocorys sp.</i>	92.4	Ciliophora
OTU82660	0.08	20	b8	9.12	3.42	connector	<i>Plasmodiophora sp.</i>	87.9	Cercozoa
OTU82680	0.05	57	b9	0.52	0.57	peripheral	Chytridiomycota	92.1	Fungi
OTU8276	0.13	46	b1	12.29	2.86	network_hub	unclassified Alveolata	94.9	unclassified Alveolata
OTU83264	0	6	b12	3.85	0.4	peripheral	Chytridiomycota	86.7	Fungi
OTU83408	0.01	3	b13	8.76	0.55	peripheral	Cercozoa 1	92.6	Cercozoa
OTU83430	0.02	23	b2	5.52	4.46	connector	<i>Chrysophaerella sp.</i>	96.2	Synurophyceae
OTU83483	0.04	34	b2	2.55	1.21	peripheral	Cryptomycota LKM11	85.7	Fungi
OTU83700	0.5	100	b4	1.27	0.34	peripheral	<i>Paraglomus sp.</i>	89.8	Fungi
OTU83736	0.01	14	b7	4.99	1.21	peripheral	<i>Chrysophaerella sp.</i>	94.4	Synurophyceae
OTU84226	0.01	11	b7	2.37	0.08	peripheral	unclassified Chrysophyceae	95.5	Chrysophyceae
OTU84691	0.02	60	b1	0.43	0.29	peripheral	unclassified Ciliophora	97.7	Ciliophora
OTU84899	0.02	57	b11	1.14	0.09	peripheral	Alveolata 1	92.3	Ciliophora
OTU85107	0.06	40	b1	14.19	2.03	network_hub	Glomeromycota	89.1	Fungi
OTU85122	0.01	17	b1	11.04	1.29	module_hub	unclassified Alveolata	93.6	unclassified Alveolata
OTU8542	0.01	46	b4	0.1	0.24	peripheral	Cryptomycota CM1	98.5	Fungi
OTU8547	0.02	23	b8	4.16	4.95	connector	unclassified Dinophyceae	91.7	Dinophyceae
OTU85607	16.75	100	b20	1.24	0.89	peripheral	<i>Epispathidium sp.</i>	98.9	Ciliophora
OTU85615	0.01	6	b18	3.17	2.38	connector	Cryptomycota CM1	81.7	Fungi
OTU85848	0.03	51	b1	0.21	0.26	peripheral	unclassified Haptophyceae	99.2	Haptophyceae
OTU85932	0.02	37	b2	0.32	0.31	peripheral	Haptorida	97.3	Ciliophora
OTU86081	0.5	100	b4	0.68	0.27	peripheral	Trebouxiophyceae	99.6	Chlorophyta
OTU86089	0.01	20	b7	6.09	1.21	peripheral	<i>Protospathidium sp.</i>	97.3	Ciliophora
OTU86217	0.01	14	b5	0.87	0.16	peripheral	Chytridiomycota	88	Fungi
OTU86364	0.01	9	b14	5.25	2.02	connector	Cryptomycota CM1	92.8	Fungi
OTU86440	0.01	6	b15	4.28	0.78	peripheral	unclassified Alveolata	92.4	unclassified Alveolata
OTU86753	0.01	20	b4	1.56	0.12	peripheral	<i>Trichosporon sp.</i>	91.5	Fungi
OTU86763	0.01	37	b6	0.62	0.95	peripheral	Chytridiomycota	98.1	Fungi
OTU86969	0.08	49	b9	0.15	0.11	peripheral	unclassified Stramenopiles	93	Stramenopiles
OTU87028	0.02	60	b4	0.83	0.15	peripheral	<i>Arcuospithidium sp.</i>	95.2	Ciliophora
OTU87225	0.34	100	b4	1.84	0.59	peripheral	unclassified Alveolata	98.5	unclassified Alveolata
OTU87443	0.01	9	b19	4.54	5.36	connector	Chytridiomycota	94.3	Fungi
OTU88053	0.3	97	b16	3.45	3.77	connector	<i>Diophrys sp.</i>	89.4	Ciliophora
OTU8823	0.08	91	b1	3.69	3.33	connector	Cryptomycota	95.4	Fungi
OTU88514	0.05	66	b11	1.17	0.18	peripheral	<i>Blepharocorys sp.</i>	91.6	Ciliophora
OTU88547	0.68	89	b10	3.14	3.43	connector	unclassified Fungi	91.3	Fungi
OTU88635	0.16	60	b6	3.72	6.58	connector	<i>Pyramimonas sp.</i>	90.8	Chlorophyta
OTU88701	0.02	34	b6	1.27	1.48	peripheral	Trichostomatia	87.5	Ciliophora
OTU89005	0	9	b12	3.04	0.99	peripheral	<i>Blepharocorys sp.</i>	92.1	Ciliophora
OTU89377	0.02	31	b6	5.95	5.3	connector	Cryptomycota CM1	98.5	Fungi
OTU89382	0.01	3	b0	1.74	0.02	peripheral	Alveolata 1	93.1	Ciliophora
OTU89415	0.02	26	b14	4.24	3.33	connector	Perkinsea 1	96.6	Perkinsea
OTU89472	0.5	100	b5	1.26	0.26	peripheral	<i>Mychonastes sp.</i>	99.6	Chlorophyta
OTU89625	0.01	43	b5	2.98	0.35	peripheral	<i>Chlamydomonas sp.</i>	94.3	Chlorophyta
OTU89954	0.03	31	b1	12.25	1.78	module_hub	unclassified Dinophyceae	92.5	Dinophyceae

OTU90077	0.02	57	b3	0.46	1.54	peripheral	<i>Pythium sp.</i>	99.6	Oomycetes
OTU90079	0.04	17	b15	1.81	9.05	connector	Cryptomycota CM1	86.7	Fungi
OTU90383	0.03	51	b11	2.23	0.14	peripheral	Alveolata 1	95.4	Ciliophora
OTU90426	0.02	46	b2	5.38	3.97	connector	unclassified Dinophyceae	100	Dinophyceae
OTU90829	0.02	14	b8	7	3.27	connector	Polycostidae	91.3	Ciliophora
OTU90984	0	9	b18	2.19	1.1	peripheral	<i>Eimeria sp.</i>	88.3	Apicomplexa
OTU91124	0.01	9	b13	4.51	4.11	connector	<i>Diophrys sp.</i>	62.8	Ciliophora
OTU91130	0.01	9	b7	8.95	1.14	peripheral	unclassified Dinophyceae	98.5	Dinophyceae
OTU91310	0.56	37	b3	13.83	2.05	network_hub	unclassified Ciliophora	98.9	Ciliophora
OTU91679	0.02	54	b2	0.06	0.81	peripheral	unclassified Alveolata	91.7	unclassified Alveolata
OTU9262	0.01	3	b15	4.5	0.88	peripheral	Pythiaceae	92.6	Oomycetes
OTU9388	0.03	60	b3	3.03	1.04	peripheral	<i>Mallomonas sp.</i>	98.9	Synurophyceae
OTU948	0.02	23	b17	1.05	0.2	peripheral	<i>Neocallimastix sp.</i>	84.6	Fungi
OTU9551	0.01	11	b7	10.45	2.6	network_hub	unclassified Dinophyceae	96.2	Dinophyceae
OTU9652	0.02	31	b2	8.61	3.71	connector	Chytridiomycota	93.9	Fungi
OTU9663	0.01	31	b2	0.44	0.63	peripheral	<i>Pythium sp.</i>	93.3	Oomycetes
OTU9748	0.01	43	b8	1.35	2.1	connector	Cercozoa 2	96.6	Cercozoa
OTU984	0	3	b13	8.76	0.55	peripheral	Desmidiiales	85.2	Streptophyta
OTU9991	0.16	80	b7	5.3	4.88	connector	Cryptomycota LKM11	95.8	Fungi





## V.9. Analyses complémentaires

### Quantification de l'abondance des chlorophytes par qPCR

Les résultats présentés dans cette partie sont issus des analyses réalisées par Léandre Vasseur au cours de son stage de BTS de 1<sup>ère</sup> année (juin 2016) co-encadré par Eric Capo, Cécile Chardon et Isabelle Domaizon

#### V.9.1. Contexte et stratégie d'étude

Les données obtenues dans les travaux basés sur le séquençage massif de l'ADN sédimentaire ont mis en évidence l'augmentation de l'abondance relative des chlorophytes dans le lac du Bourget en lien avec l'accroissement des teneurs en phosphore ( $> 80 \mu\text{g.L}^{-1}$ , Figures **V.1** et **V.S2**). Afin de confirmer (ou non) cette tendance, une quantification de l'ADN de chlorophytes a été réalisée par qPCR à partir des mêmes échantillons sédimentaires en utilisant des amorces nucléotidiques spécifiques aux chlorophytes (Turon *et al.* 2015).

Dans un premier temps, la spécificité des amorces nucléotidiques a été vérifiée par amplification PCR de l'ADN extrait des cultures pures d'algues (chlorophytes ou non) et en séquençant des banques de clones réalisées à partir d'ADN sédimentaire (lac d'Annecy et du Bourget). Après cette vérification, le protocole pour l'utilisation des amorces nucléotidiques en PCR quantitative a été mis au point (réalisation d'une gamme étalon, test sur des témoins positifs) puis les analyses qPCR ont été réalisées sur les extraits d'ADN sédimentaire des lacs d'Annecy et du Bourget.

#### V.9.2. Matériels et méthodes

##### *Extraction d'ADN à partir de cultures d'algues et échantillons environnementaux*

Afin de tester la spécificité des amorces spécifiques aux chlorophytes utilisées pour les analyses moléculaires, l'ADN a été extrait de 9 algues mises en culture au sein de la TCC (Thonon culture collection) (voir Tableau **V.4**). L'extraction a été réalisée à l'aide de la méthode GenElute<sup>TM</sup>-LPA (Sigma Aldrich, Saint-Louis, Missouri, USA). L'ADN des échantillons environnementaux (couches sédimentaires provenant des deux carottes sédimentaires présentées dans la partie principale de ce chapitre) a été extrait à l'aide d'un kit NucleoSpin<sup>®</sup> Soil en suivant les recommandations du fournisseur Macherey-Nagel (Düren, Germany). Le dosage de l'ADN extrait à partir des cultures d'algues et des échantillons environnementaux a été réalisé à l'aide d'un appareil Nanodrop 1000 Spectrophotometer V3.7 (Thermo Fisher Scientific, Ecublens, Suisse) (voir Tableau **V.4**).

Tableau V.4. Caractéristiques des 9 souches d'algues utilisées pour la vérification de la spécificité des amorces nucléotidiques

Nom de la souche	Affiliation Taxonomique	Groupe taxonomique	Concentration d'ADN (en ng.µL <sup>-1</sup> )
TCC3 A	<i>Chlamydomonas intermedia</i>	Chlorophyte	92.82
TCC3 B	<i>Chlamydomonas intermedia</i>	Chlorophyte	160.12
TCC343	<i>Chlamydomonas sp.</i>	Chlorophyte	41.91
TCC120	<i>Pediastrum duplex</i>	Chlorophyte	31.79
TCC144-1	<i>Scenedesmus quadricauda</i>	Chlorophyte	34.65
TCC150-1	<i>Chlorella vulgaris</i>	Chlorophyte	15.55
TCC167	<i>Synura petersenii</i>	Chrysophycée	33.11
TCC301	<i>Fragilaria crotonensis</i>	Diatomée	25.18
TCC139-1	<i>Nitzschia palea</i>	Diatomée	32.78

#### *Amplification PCR de fragments nucléotidiques spécifiques aux Chlorophytes*

Des amplifications PCR ont été réalisées à partir de l'ADN extrait des cultures d'algues et d'échantillons environnementaux des lacs d'Annecy et du Bourget. Une portion du gène codant l'ARNr 18S (70 pb) et spécifique aux chlorophytes a été amplifiée en utilisant les amorces INT-4F (5' TGGTGAAGTGTTCGGATTGG 3') et INT-5R (5' ARGTGGGAGGGTTTAATGAA 3'). Chaque réaction PCR a été réalisée dans un volume total de 30 µL contenant 1 µL ADN (25 ng.L<sup>-1</sup>), 3 µL de tampon de réaction 10X NH<sub>4</sub>, 1.2 µL de MgCl<sub>2</sub> (50 mM), 0.25 µL d'ADN polymérase BioTaq (Bioline), 0.24 µL de dNTP (10 mM), 0.36 µL de BSA (0.5 µg.µL<sup>-1</sup>) et 1 µL de chaque amorce. Les conditions d'amplification utilisées ont été constituées d'une étape dénaturation initiale de 3 min à 95 °C suivi par 40 cycles de 10 sec à 95 °C, de 30 sec à 56 °C, de 30 sec à 72 °C et en terminant par une étape de 5 min à 72 °C. Les produits PCR obtenus à partir des échantillons environnementaux ont été purifiés à l'aide du kit Illustra GFX PCR DNA and Gel Band purification (Fisher Scientific, Ecublens, Suisse) selon les recommandations du fournisseur.

#### *Construction de banques de clones*

Des banques de clones ont été réalisées à partir de produits PCR spécifiques aux chlorophytes obtenus à partir d'échantillons environnementaux. Le kit TOPO TA Cloning (Invitrogen) a été utilisé. Une ligation a été réalisée à partir de produits PCR purifiés obtenus à partir de 2 échantillons environnementaux du lac d'Annecy et 2 du lac du Bourget. Pour chaque réaction, 1 µL d'insert (produit PCR) purifié, 1 µL de solution saline, 4 µL d'eau stérile et 1 µL de plasmide Topo vector 2.1. Chaque réaction a été incubée 20 minutes à température ambiante. Ensuite, 2 µL de la réaction a été mélangé doucement dans un tube contenant des cellules compétentes d'*Escherichia coli* préalablement dégelées. Le mélange a été placé dans la glace pendant 10 minutes puis à 42 °C pendant 40 secondes puis dans la glace pendant 3 minutes. Enfin, 250 µL de milieu SOC a été ajouté et le mélange a été incubé pendant 1 heure 30 minutes à 37 °C. L'étalement de 50, 80 et 100 µL du mélange de transformation a été effectué sur milieu solide LB + Ampicilline (100 mg/L), préalablement recouvert avec du Xgal à 80 µg.mL<sup>-1</sup> et de l'IPTG à 50 µM. Les boîtes incubées pendant 18 h à 37 °C. Les colonies blanches obtenues ont été repiquées dans 20 µL d'eau ultra-pure. Les clones ont été placés au thermocycleur à 95 °C pendant 10 minutes.

Afin de vérifier la présence du plasmide dans chacun des clones repiqués, des amplifications PCR M13 ont été réalisées en utilisant des amorces spécifiques M13F (5' TTGTAACGACGGCCAGTG 3') et M13R (5' CACACAGGAAACAGCTATGACCATG 3'). Chaque réaction PCR a été réalisée dans un volume total de 30 µL contenant 1 µL ADN (25 ng.L<sup>-1</sup>), 3 µL de tampon de réaction 10X NH<sub>4</sub>, 1.2 µL de MgCl<sub>2</sub> (50 mM), 0.12 µL d'ADN polymérase BioTaq

(Bioline), 0.6 µl de dNTP (10 mM), et 1 µl de chaque amorce. Les conditions d'amplification utilisées ont été constituées d'une étape dénaturation initiale de 10 min à 94 °C suivi par 30 cycles de 1 min à 94 °C, de 1 min à 55 °C, de 1 min à 72 °C suivi d'une étape de 10 min à 72 °C. Enfin, les clones possédant le plasmide ligué ont été envoyés au séquençage Sanger réalisé par Genewiz (Cambridge, MA, USA). Les séquences nucléotidiques obtenues ont été nettoyées (retrait des séquences correspondant au plasmide) puis comparées aux séquences disponibles au sein de la base de données NCBI à l'aide du logiciel BLASTn (NCBI).

### *Réalisation d'une gamme étalon pour la PCR quantitative*

Afin de quantifier l'abondance des chlorophytes dans les échantillons environnementaux, la méthode de PCR quantitative (qPCR) a été utilisée. Une gamme étalon a été réalisée à partir d'un clone d'une des banques réalisées (voir partie V.9.2.4.). Le séquençage de l'insert transformé dans ce clone a permis d'obtenir une séquence nucléotidique affiliée à une chlorophyte proche de *Radiococcus* ou *Chlorella sp.* Après vérification de la spécificité de l'insert de ce clone, les colonies de ce clone ont ensuite été repiquées en milieu liquide LB + Ampicilline pendant 24 h à 37 °C. Afin de récupérer seulement le plasmide transformé, une purification a été réalisée à l'aide du kit QIAGEN Miniprep (Venlo, Pays-Bas) en suivant les recommandations du fournisseur. Une restriction enzymatique a été réalisée en utilisant une enzyme BamHI coupant un site de restriction sur la séquence plasmidique et ne possédant pas de sites de restriction sur l'insert. Le mélange réactionnel a été composé de 2 µL de BamHI, 4 µL de Tampon 10X, 10 µL d'ADN et 24 µL d'eau ultrapure. Le mélange a été incubé 4 h à 37 °C puis 20 min à 65 °C. Un dosage du plasmide + insert linéarisé a été réalisé à l'aide de la méthode Qubit (Invitrogen, Carlsbad, Californie, USA) en suivant les recommandations du fournisseur.

Le nombre de copies de la séquence d'intérêt par µL a été calculé à l'aide de la formule suivante :

$$\text{Nombre de copies} = ((\text{concentration en ADN}) / (\text{taille du plasmide} * 660)) * 6.02 * 10^{23}$$

Nombre de copies de gène cible par µL de solution de plasmide purifié et linéarisé

Concentration en ADN : g/µL

Taille du plasmide : pb

660 = masse molaire d'une paire de base : g/mol

$6.02 * 10^{23}$  = constante d'Avogadro : mol<sup>-1</sup>

A partir du nombre de copies/µL obtenu pour le clone d'intérêt ( $1.8 * 10^9$ ), des dilutions en cascade de raison 1/10<sup>ème</sup> ont été obtenus de  $1.8 * 10^9$  à  $1.8 * 10^0$ .

### *Quantification du nombre de copies de la séquence d'intérêt par qPCR*

Un volume réactionnel de 25 µL a été utilisé pour chaque réaction et composé de 1 µL d'ADN, 12.5 µL de sybrGreen Master Mix, 1 µL de chaque amorce nucléotidique (INT4F INT5R voir partie Amplification PCR) et 0.3 µL d'albumine sérique bovine (BSA). Les conditions d'amplification ont été constituées d'une étape dénaturation initiale de l'ADN 15 min à 95 °C suivi par 40 cycles de dénaturation de 10 sec à 95 °C, d'hybridation de 30 sec à 56 °C, d'élongation de 30 sec à 72 °C. Une acquisition continue du signal de fluorescence (FAM) permis de quantifier le nombre de copies de la séquence et de vérifier la spécificité des amorces.

### V.9.3. Résultats

#### *Vérification de la spécificité des amorces nucléotidiques*

Pour chaque extrait d'ADN obtenu à partir des cultures d'algues, des amplifications PCR ciblant des séquences nucléotidiques spécifiques aux chlorophytes ont été réalisées. Pour toutes les algues non chlorophytes, aucune amplification n'a été détectée alors que pour la plupart des extraits d'ADN de chlorophytes, des amplifications PCR positives ont été observées. Cependant aucune amplification n'a été détectée pour les souches de *Chlorella vulgaris* (TCC 150-1) et *Chamydomonas sp.* (TCC 3A). La spécificité des amorces nucléotidiques a également été testée sur 4 échantillons environnementaux (ADN extrait de strates sédimentaires): deux du lac d'Annecy, et deux du lac du Bourget. Pour chaque échantillon, 48 clones (positifs aux amplifications PCR avec les amorces M13) ont été sélectionnés pour le séquençage Sanger. Les résultats du séquençage Sanger ont permis d'obtenir 41 séquences nucléotidiques sur les 192 expédiées. Les assignations taxonomiques effectuées ont permis de confirmer la spécificité des amorces aux chlorophytes. En effet, seule une des 41 séquences obtenues possédait une affiliation ambiguë (aussi bien affiliée à une chlorophyte qu'à un apicomplexé (voir Tableau V.5).

#### *Quantification de l'abondance des chlorophytes dans les séries temporelles des lacs d'Annecy et du Bourget par qPCR*

Les données de qPCR obtenues ont permis de mettre en évidence une augmentation de l'abondance des chlorophytes dans les années 1980s dans le lac du Bourget de façon cohérente avec les données d'abondance relative observées dans les analyses basées sur le séquençage massif de l'ADN (Figure V.8). Pour le lac d'Annecy, les quantités de chlorophytes sont faibles et sont également cohérentes avec les résultats obtenus précédemment concernant l'abondance relative de ces taxons dans ce lac. De plus l'abondance des chlorophytes quantifiée par qPCR a été corrélée significativement ( $p$ -value < 0.01) avec les concentrations en phosphore dans les deux lacs notamment lors de l'eutrophisation forte dans les années 80s dans le lac du Bourget.

Tableau V.5. Affiliation taxonomique associée aux séquences obtenues pour les échantillons environnementaux de chaque lac

<b>Lac d'Annecy</b>	
<b>Affiliation taxonomique</b>	<b>Nombre de séquences d'ADN</b>
<i>Choricystis/Nannochloris sp.</i>	5
<i>Scenedesmus/Desmodesmus sp.</i>	4
<i>Chlorella sp.</i>	4
<i>Radiococcus sp.</i>	4
<i>Hafniomonas sp.</i>	3
<i>Schizomeris sp.</i>	2
<i>Pseudopediastrum/Stauridium sp.</i>	1
<i>Mychonastes sp.</i>	1

<b>Lac du Bourget</b>	
<b>Affiliation taxonomique</b>	<b>Nombre de séquences d'ADN</b>
<i>Mychonastes sp.</i>	10
<i>Chlorella sp.</i>	2
<i>Pseudopediastrum/Lacunastrum sp.</i>	1
<i>Tetraedron sp.</i>	1
<i>Desmodesmus sp.</i>	1
<i>Monoraphidium sp.</i>	1
<i>Uncultured Chlorophyta</i>	1

#### **V.9.4. Conclusion**

Les résultats obtenus apportent des connaissances originales sur la dynamique temporelle des chlorophytes dans les lacs ce qui n'avait jamais été réalisé précédemment notamment en utilisant de l'ADN sédimentaire. Les données qPCR confirment une tendance observée pour la dynamique temporelle des chlorophytes via l'approche de séquençage massif de la communauté microbienne eucaryote des deux lacs étudiés. Elles confirment l'impact positif des fortes augmentations de phosphore dans le lac du Bourget sur l'abondance des chlorophytes. Même si ces deux approches (séquençage massif de l'ADN et qPCR) ne permettent pas de conclure sur l'abondance réelle des chlorophytes dans la masse d'eau au vu des connaissances concernant la variabilité du nombre de copies de gènes en fonction des organismes, elle semble apporter néanmoins des informations concernant la réponse de ce groupe fonctionnel à l'eutrophisation de la masse d'eau. Les données acquises concernant la diversité des chlorophytes dans les deux lacs apportent une vision complémentaire des résultats précédemment acquis notamment concernant l'affiliation taxonomique de ces taxons. En perspectives de ces travaux, il serait intéressant d'étudier la dynamique temporelle d'autres groupes d'algues notamment les dinophycées qui semblent avoir été favorisé dans la période récente par le réchauffement climatique et les conditions oligotrophes des deux lacs (depuis le milieu des années 1980s).

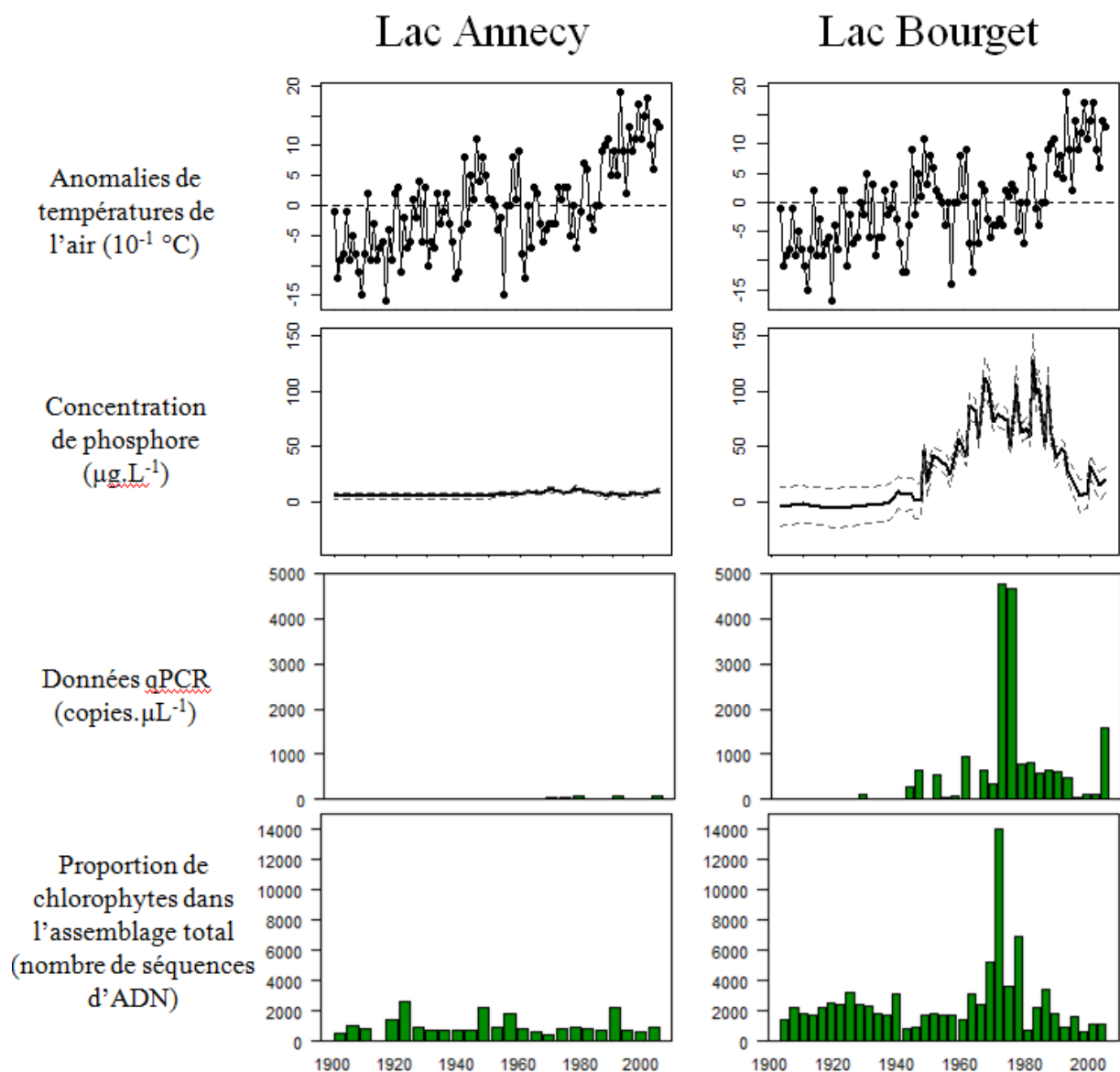


Figure V.8. Changements temporels au cours du siècle dernier dans les lacs Annecy et Bourget de la quantité de copies de séquences d'ADN de chlorophytes et la proportion relative des chlorophytes au sein des assemblages de micro-eucaryotes (résultats issus de la Figure V.1). Les forçages environnementaux (anomalies de température de l'air et estimation des teneurs en phosphore sont également présentés dans ce graphique ; voir Figure V.1 pour précision).

## VI. Discussion et Perspectives

L'approche de paléocéologie moléculaire appliquée dans ces travaux est basée sur l'utilisation du séquençage massif d'ADN sédimentaire pour tenter de reconstruire la dynamique temporelle des communautés microbiennes eucaryotes passées. Ces travaux ont permis de répondre à des questions d'écologie microbienne portant sur la dynamique temporelle de la communauté de micro-eucaryotes lacustres et ses réponses aux forçages environnementaux sur des pas de temps longs (2200 ans dans les travaux du chapitre IV et au cours du siècle dernier dans le chapitre V). Les données obtenues sont les premières à caractériser, via une approche de paléocéologie moléculaire, les modifications temporelles de la communauté totale des micro-eucaryotes en milieu lacustre (bien que cette approche ait déjà été appliquée par Coolen et al. (2013) à partir de sédiments de la Mer noire). Ainsi, contrairement aux études précédentes ciblant des groupes fonctionnels spécifiques (voir Table I.1), ces travaux permettent d'étudier la dynamique temporelle dans une communauté biologique complexe comprenant de nombreux taxons pouvant interagir au sein du réseau trophique aquatique (notamment relations proie-prédation, parasitisme, saprophytisme).

Ce domaine d'étude récent est toutefois encore en cours de calibration, aussi ce travail a inclus 2 volets visant à calibrer le signal ADN sédimentaire pour les communautés de micro-eucaryotes. L'étude de la préservation de l'ADN sédimentaire a permis de mettre en évidence l'efficacité d'archivage de l'ADN de la plupart des taxons dans les sédiments lacustres (chapitre II) et d'estimer les effets potentiels des processus diagénétiques au cours des premières années d'enfouissement sur la 'lecture' ADN de la composition et la structure des communautés microbiennes eucaryotes lacustres (chapitre III).

### VI.1. Autour de la validation du proxy ADN sédimentaire en milieu lacustre

Les sédiments sont connus pour être des archives pertinentes pour retracer les conditions physico-chimiques et biologiques passées des milieux lacustres (Meyers & Lallier-vergés 1999, Smol & Cumming 2000, Battarbee *et al.* 2012). Alors que la plupart des études se basent historiquement sur des biomarqueurs classiques (restes d'organismes, concentrations en lipides et pigments), l'ADN préservé dans les sédiments aquatiques est utilisé seulement



depuis une dizaine d'années pour reconstruire la dynamique temporelle des communautés biologiques sur des pas de temps longs (centenaires à millénaires ; voir Table I.1 pour liste des études). Cependant, peu d'études ont été réalisées pour évaluer si l'ADN provenant de la communauté biologique de la masse d'eau est efficacement archivé dans les sédiments lacustres. En effet, des modifications du signal ADN peuvent se produire au cours de la sédimentation dans la colonne d'eau (dégradation ou consommation dans la colonne d'eau), au cours du dépôt de l'ADN à l'interface eau-sédiment et pendant l'enfouissement dans les sédiments sous l'effet de processus diagénétiques (Dell'Anno *et al.* 2002, Dell'Anno & Corinadelsi 2004, Corinadelsi *et al.* 2008, 2011, Torti *et al.* 2015). Les travaux de cette thèse apportent des connaissances nouvelles sur certains aspects de calibration de l'archivage de l'ADN dans les sédiments. Les principaux résultats de ces travaux sont synthétisés dans la Figure VI.1.

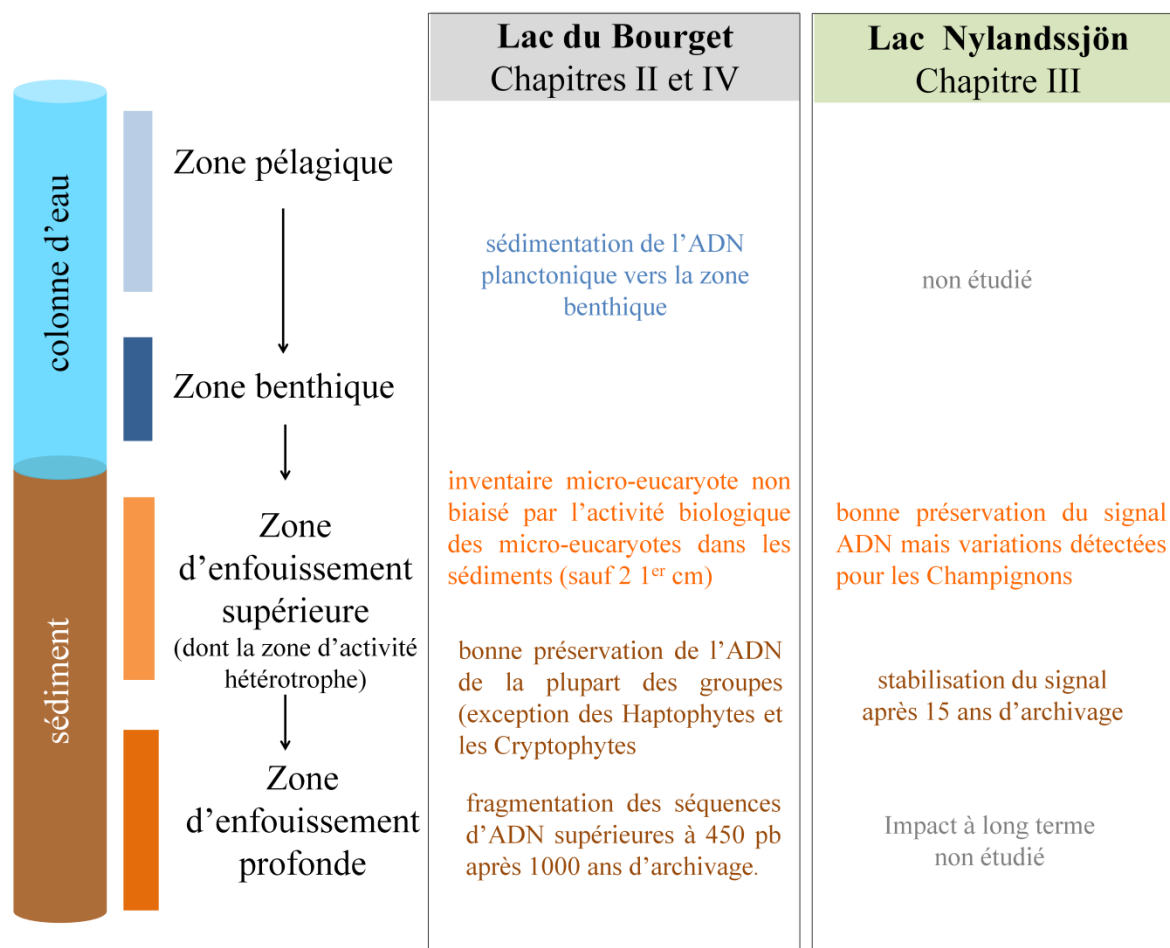


Figure VI.1. Synthèse des principaux résultats portant sur la préservation de l'ADN des micro-eucaryotes au cours de son archivage dans les sédiments lacustres.

### **VI.1.1. De la masse d'eau au sédiment : une bonne préservation de l'ADN des communautés de micro-eucaryotes de la masse d'eau (lac profond tempéré)**

Dans le chapitre II, des informations nouvelles ont été apportées par la comparaison des inventaires moléculaires de la masse d'eau (profondeur 2 m et 130 m) et des archives sédimentaires du lac du Bourget (sédiments récents : de 2 à 60 ans). Parmi les résultats principaux de cette étude, il a été montré que plus de 70 % des unités phylogénétiques présentes dans la masse d'eau sont également retrouvées dans les sédiments lacustres notamment au sein des groupes de chlorophytes, dinophycées, chrysophycées, ciliés et champignons. Cependant, l'ADN de certains groupes comme les cryptophytes et les haptophytes semblent moins bien archivé en termes de richesse dans les sédiments.

Un des processus pouvant être impliqué dans la perte du signal ADN de ces taxons dans les archives sédimentaires est leur consommation par d'autres organismes au cours de la sédimentation dans la colonne d'eau. Les données produites dans l'étude du chapitre II montrent que l'ADN des cryptophytes et haptophytes présent à 2 m de profondeur dans la colonne d'eau est également présent à 130 m, et suggèrent donc que les processus biologiques ayant lieu au cours de la sédimentation des cellules mortes dans la colonne d'eau ne sont pas responsables de la perte du signal ADN. Toutefois, alors que les prélèvements mensuels de la masse d'eau ont été réalisés à la même date pour les échantillons à 2 m et à 130 m, on peut supposer qu'un échantillonnage décalé dans le temps (~1 semaine) permettrait d'optimiser le suivi de la sédimentation. Une stratégie d'étude complémentaire pourrait donc être le suivi de la modification du signal ADN au cours de la sédimentation des cellules dans la colonne d'eau via l'utilisation des trappes sédimentaires comme récemment décrit pour l'étude de la sédimentation de *Phaeocystis* en milieu marin (Wolf *et al.* 2016). Alors que cette algue était dominante dans la zone pélagique (bloom), seulement quelques séquences d'ADN de cette algue ont été retrouvées à l'aide de trappes sédimentaires suggérant que cette algue pourrait avoir été recyclée au cours de sa sédimentation dans la colonne d'eau. Nous avons fait le choix de ne pas mettre en place ce type de procédure expérimentale lors de cette thèse car cela nécessitait une forte implication en termes de suivi de terrain à long terme (au moins deux ans) pour obtenir des données solides permettant de répondre à cette question.

Dans ces travaux, nous avons utilisé des données moléculaires provenant de deux projets différents donc qui n'était pas originellement mis en place pour l'étude de l'archivage de l'ADN ; il s'agissait (i) pour la masse d'eau, d'un projet au sein de l'observatoire des lacs alpins visant à la caractérisation de la diversité génétique des communautés microbiennes par

un suivi mensuel de la masse d'eau du lac du Bourget (ii) pour les archives sédimentaires, le projet ANR IPER-RETRO visant à l'exploration de plusieurs biomarqueurs sédimentologiques et biologiques pour reconstituer les trajectoires évolutives des 3 lacs péri-alpins (à partir des restes d'organismes, de signatures isotopiques, ou d'ADN ; pour synthèse des résultats voir Perga *et al.* 2015). Les inventaires moléculaires, issus du séquençage 454 de la région V4 du gène codant l'ARNr 18S, étaient donc disponibles au commencement de cette thèse, ces données ont donc été utilisées pour comparer le signal ADN pour la communauté actuelle de la masse d'eau (2011-2012) et celui enregistré dans les sédiments récents (27 premiers cm de la carotte sédimentaire). Cependant, une stratégie d'étude plus efficace, mais nécessitant une durée d'étude incompatible avec la durée de thèse, aurait été de comparer des inventaires moléculaires issus d'échantillons de la masse d'eau et d'archives sédimentaires de la même année (par exemple, l'échantillonnage mensuel de la masse d'eau en 2007 comparé à la strate sédimentaire correspondante). Ce type de données permettrait de conforter les données acquises concernant le bon archivage de l'ADN de la majorité des taxons. Cette procédure expérimentale pourrait potentiellement être planifiée puisque les collections ADN de la masse d'eau ont été poursuivies de 2011 à 2015, l'échantillonnage d'une carotte sédimentaire en 2017 ou au-delà permettrait de réaliser cette comparaison.

#### **VI.1.2. Les effets de la diagénèse précoce sur l'ADN des micro-eucaryotes lacustres : effets très modérés pour les sédiments varvés du Lac Nylandssjön**

Au cours de l'enfouissement dans les sédiments aquatiques, l'ADN est connu pour être impacté par des processus diagénétiques entraînant la fragmentation et des mutations des séquences d'ADN (dépurination, déamination, création de sites abasiques) (Dell'Anno & Corinadelsi 2004, Corinadelsi *et al.* 2008, Torti *et al.* 2015). Ces processus ont été détectés sur des pas de temps relativement longs c'est-à-dire plusieurs millénaires ou dizaines de millénaires (Boere *et al.* 2011a, Kirkpatrick *et al.* 2016). Il y a peu d'études concernant la diagénèse précoce (ici définie comme se déroulant au cours des premières années d'archivage) de l'ADN dans les sédiments lacustres.

Grâce aux travaux du chapitre III, nous avons étudié l'effet des premières années d'enfouissement de l'ADN sur les inventaires moléculaires des communautés microbiennes eucaryotes. La stratégie d'étude consistait à comparer les inventaires moléculaires de deux carottes sédimentaires prélevées dans le lac Nylandssjön à 6 ans d'intervalles pour révéler les effets de la diagénèse précoce. Ce travail a été réalisé à partir d'une collection de carottes

sédimentaires unique au monde, prélevées par les chercheurs de l'équipe de paléolimnologie de l'Université d'Umeå (Suède). Ces sédiments annuellement varvés permettent de sélectionner précisément chaque dépôt annuel et de le suivre au cours du temps grâce aux carottes échantillonnées (et congelées) régulièrement dans le lac ; c'est pour cela que les sédiments de ce lac ont été largement utilisés pour étudier les effets de la diagénèse précoce sur les teneurs de plusieurs éléments (carbone, azote, fer, soufre, mercure...) et la compaction des varves sédimentaires (Gälman *et al.* 2008, 2009a, 2009b, Rydberg *et al.* 2008, Maier *et al.* 2013, Rydberg & Martinez-Cortizas 2014).

Les analyses moléculaires ont montré que (i) la quantité d'ADN extrait des varves sédimentaires, la richesse et la diversité de la communauté microbienne eucaryote ne semblent pas avoir été modifiées au cours de l'enfouissement (ii) les dynamiques temporelles des communautés de micro-eucaryotes (structure des communautés révélée par les analyses nMDS) obtenues pour les deux carottes sont congruentes (iii) dans les strates les plus récentes (< 15 ans), des différences dans la composition et la structure des communautés sont observées alors que le signal ADN des strates plus anciennes est stabilisé (> 15 ans jusqu'à 40 ans) (iv) les différences dans les inventaires moléculaires des deux carottes sont majoritairement liées aux variations dans l'abondance de quelques OTUs dominantes (notamment des chlorophytes et des Stramenopiles non affiliés) et au changement plus net de la proportion et de la richesse des champignons (comme discuté dans le chapitre III). Bien que la diminution des teneurs de carbone et d'azote ait été mise en évidence dans les sédiments du lac Nylandssjön au cours des 5 premières années d'archivage (Gälman *et al.* 2008, Rydberg & Martinez-Cortizas 2014), les effets de la diagénèse précoce sur l'ADN sédimentaire se produisent sur une période de temps plus longue (jusqu'à 15 ans). Cela illustre le fait que des processus peuvent impacter la teneur des éléments chimiques et de molécules complexes (notamment leur (bio)disponibilité pour le recyclage ou la dégradation) sur des pas de temps différents. Ces résultats pourraient mettre en évidence que la zone d'activité microbienne (surtout bactérienne) des sédiments du lac Nylandssjön où la matière organique est transformée (dont l'ADN, Fischer *et al.* 2002, Haglund *et al.* 2003, Dell'Anno & Danovaro 2005) correspond aux 15 premières années d'archivage. Ici, seulement 12 strates ont été sélectionnées pour chaque carotte (datés entre 1973 et 2006) afin de limiter le coût de l'analyse. Cependant, une approche plus résolutive consisterait à obtenir les inventaires moléculaires de chaque strate (environ 35 strates communes entre les deux carottes). Cela permettrait notamment de préciser la « zone de rupture » ou le signal ADN des strates des

deux carottes n'est plus modifié. En complément, la comparaison avec des carottes sédimentaires prélevées à des intervalles de temps plus courts (1 an d'intervalle par exemple en utilisant une carotte sédimentaire prélevée en 2012) permettrait d'étudier comment le signal change à un rythme annuel.

Il est toutefois important de mentionner que les inventaires moléculaires des deux carottes utilisées permettent d'obtenir une visualisation très similaire de la dynamique temporelle des communautés (cf analyse procrustéenne, Figure **III.8**) ; les variations mises en évidence lors de la diagenèse précoce demeurent donc relativement subtiles et a priori assez modérément impactantes pour le signal global obtenu pour les communautés de micro-eucaryotes. Seuls les résultats obtenus pour le groupe des champignons soulèvent des questions qu'il sera nécessaire d'investiguer, notamment concernant leur état actif dans les sédiments. En effet l'augmentation de la richesse et de la proportion des champignons entre les deux carottes suggère un changement dans les conditions physico-chimiques et biologiques des sédiments pouvant conduire à une modification dans l'activité de certains microorganismes en fonction des carottes (variabilité spatiale), notamment les champignons connus pour être actifs dans les sédiments aquatiques superficiels (Edgcomb *et al.* 2011, Orsi *et al.* 2013).

### **VI.1.3. Effet du vieillissement des sédiments sur l'ADN de micro-eucaryotes au cours de longues périodes de temps**

Comme mentionné précédemment, l'ADN est impacté au cours de son vieillissement dans les sédiments, a priori en particulier pour des périodes de temps relativement longues c'est-à-dire plusieurs millénaires ou dizaines de millénaires (Boere *et al.* 2011a, Torti *et al.* 2015, Kirkpatrick *et al.* 2016). La fragmentation des séquences d'ADN a ainsi été étudiée dans certains travaux pour montrer les effets du vieillissement sur la longueur des séquences d'ADN (Coolen & Overmann 1998, Boere *et al.* 2011b).

Dans cette thèse, bien que les travaux ne portent pas sur ces aspects méthodologiques, les essais menés pour valider le fragment ADN utilisé comme barcode pour ces analyses paléoécologiques ont permis de mettre en évidence une perte du signal de l'amplification du fragment V4 (350 pb, du gène codant l'ARN 18S) dans des sédiments du lac du Bourget âgés de plus de 1000 ans, alors que le fragment V7 plus court (260 pb, du gène codant l'ARN 18S) été amplifié sur toute la période étudiée (2200 ans).

Alors que le fragment long (région V4, 350 pb) a été utilisé pour les analyses du chapitre II, nous avons privilégié le choix du fragment plus court (région V7, 260 pb) pour les autres études (chapitres III, IV et V) permettant de répondre aux exigences habituelles des approches sur l'ADN dit ancien (c'est à dire analyse de fragments ADN courts ; Coolen & Gibson 2009, Torti *et al.* 2015). Il aurait été intéressant de comparer les inventaires moléculaires obtenus à partir de ces deux fragments mais nous avons choisi de privilégier l'étude des changements temporels des communautés microbiennes de plusieurs lacs plutôt que cette question méthodologique pour ces travaux de thèse. Certaines études portant sur des échantillons environnementaux actuels ont par ailleurs montré qu'il existe des différences dans la vision de la communauté obtenue à partir de différents fragments d'ADN (Taïb *et al.* 2013, Bradley *et al.* 2016) et suggèrent l'utilisation combinée de plusieurs barcodes pour la caractérisation de la communauté étudiée ; ce type de stratégie très couteuse en termes de temps et d'argent n'a pas été appliqué ici.

#### **VI.1.4. Une préservation différentielle de l'ADN des communautés biologiques du lac au cours de l'enfouissement dans les sédiments ?**

Plusieurs études suggèrent que l'ADN des micro-eucaryotes pourrait présenter une préservation différentielle au cours de leur archivage dans les sédiments, donc variable en fonction des groupes taxonomiques considérés (principalement des études sur les sédiments marins, Boere *et al.* 2011a, Kirkpatrick *et al.* 2016). Pour les sédiments lacustres, les résultats des chapitres II et III mettent en évidence que (i) l'ADN de la majorité des taxons planctoniques du lac du Bourget est plutôt bien archivé dans les sédiments que ces taxons laissent ou non des restes morphologiques dans les sédiments (kystes, spores,...) (ii) dans les sédiments du lac Nylandssjön, l'abondance relative de certains taxons a été modifiée au cours des premières années d'archivage mais il n'est pas apparu de lien visible avec la capacité de ces taxons à former des cellules de résistances.

La préservation différentielle de l'ADN dans les sédiments pourrait être liée aux effets (combinés) de plusieurs facteurs :

- **L'architecture cellulaire** : L'architecture cellulaire (ex les parois) pourrait avoir un rôle important dans la préservation de l'ADN dans les sédiments. Kirkpatrick *et al.* (2016) suggèrent que l'ADN des diatomées semble bien mieux archivé dans les sédiments que celui de la plupart des taxons photosynthétiques en se basant sur la comparaison de la proportion de l'ADN chloroplastique et de l'ADN associé aux diatomées (par rapport au pool total d'ADN

16S chloroplastique) dans des sédiments très anciens (plus de 100-200 000 années). Les travaux de Boere et ses collègues (2011a) montrent également que la quantité de l'ADN des diatomées est plus stable que celle des dinophycées qui déclinent graduellement au cours du temps (notamment pour des strates sédimentaires âgées de plus de 1500 ans). En revanche, dans la même étude, la quantité d'ADN des bactéries vertes sulfureuses ne diminue pas avec l'âge des sédiments alors qu'elles ne possèdent pas de structures extracellulaires suggérant que l'architecture cellulaire n'est pas le seul facteur pouvant expliquer une préservation différentielle. De plus, bien qu'une architecture cellulaire robuste semble être un avantage pour la préservation de l'ADN (notamment structure des membranes, présence de paroi), elle pourrait également être un facteur biaisant négativement la représentativité de certains groupes de protistes au sein des assemblages microbiens. Ainsi, Koid et al. (2012) ont montré que les ciliés marins sont mieux représentés dans les assemblages eucaryotes quand une lyse mécanique est utilisée pour l'extraction de l'ADN alors que la lyse chimique semble bénéficier à d'autres taxons comme les diatomées. Les diatomées, qui sont classiquement utilisées comme biomarqueurs en paléoécologie grâce à leurs frustules bien préservées dans les sédiments au cours du temps (Battarbee 2000, Battarbee *et al.* 2002, Davidson & Jeppesen 2013) sont paradoxalement peu diverses et abondantes dans les inventaires moléculaires obtenus à partir des sédiments (voir chapitres III, IV et V). Cela pourrait résulter de la difficulté à extraire l'ADN de leurs cellules à cause de frustules difficile à casser voire de la forte adsorption de l'ADN aux particules siliceuses (de la frustule). Ainsi, la correspondance entre les données moléculaires et les restes fossilisés n'est aujourd'hui pas encore validée et des travaux sont actuellement menés sur ce sujet (notamment au sein de l'UMR CARRETEL : thèse de Valentin Vasselon). La difficulté à extraire l'ADN intracellulaire pourrait également expliquer le faible pourcentage d'Haptophytes de la masse d'eau retrouvés dans les sédiments du lac du Bourget (chapitre II, Capo *et al.* 2015). En effet, les Haptophytes ne possèdent pas de paroi cellulaire mais les prymnésiophytes sont recouverts d'écailles généralement calcaires (les coccolithes) qui peuvent conférer une certaine protection à l'ADN intracellulaire. Ainsi, globalement, il n'y a pas à l'heure actuelle de réel consensus (et relativement peu d'études) concernant les effets de l'architecture cellulaire sur l'efficacité d'archivage des taxons.

La forme sous laquelle l'ADN atteint le sédiment est un facteur à considérer car la préservation différentielle de l'ADN pourrait être liée à cette caractéristique ; l'ADN peut arriver sous forme extracellulaire, sous forme intracellulaire dans des cellules simples ou des

cellules de résistances (kystes par exemple) donc physiquement protégé des nucléases extracellulaires. L'ADN extracellulaire est ainsi considéré comme plus exposé à la dégradation par les nucléases microbiennes et les processus diagénétiques tels que l'hydrolyse et l'oxydation (Dell'Anno *et al.* 2002, Corinaldesi *et al.* 2005, 2008, Willerslev & Cooper 2005) en comparaison avec l'ADN intracellulaire. Afin d'avoir une vision globale de la communauté de micro-eucaryotes, nous avons choisis d'utiliser, dans ces travaux de thèse, le kit Macherey Nagel NucleoSpin Soil qui permet en théorie d'extraire de l'ADN intra et extracellulaire en même temps (sauf les fragments < 50 bp qui peuvent être perdus au cours de l'extraction selon le fournisseur). Des analyses complémentaires basées sur la comparaison de différentes méthodes d'extraction de l'ADN, permettant d'extraire l'ADN extracellulaire, intracellulaire ou les deux, apporteraient des éléments de réponse intéressants sur la part des taxons représentés par ces différents types d'ADN (extra ou intracellulaire) dans les sédiments. Des protocoles existent actuellement pour extraire préférentiellement l'ADN extracellulaire et peuvent être adaptés aux sédiments, ils utilisent par exemple des tampons saturés en phosphate pour désorber l'ADN extracellulaire (Taberlet *et al.* 2012, Lever *et al.* 2015). De plus, des études ciblant indépendamment des groupes spécifiques, comme dans Boere *et al.* (2011a), et combinant des données de paléocéologie classique (comptage de restes, teneurs en lipides) et moléculaire (qPCR et séquençage à haut débit) permettraient d'apporter de nouvelles connaissances quant à la préservation différentielle de l'ADN des certains groupes microbiens dans les sédiments.

- les **conditions physico-chimiques dans les sédiments** : L'adsorption de l'ADN extracellulaire aux particules minérales et organiques dépend de plusieurs caractéristiques qui peuvent varier en fonction du type de dépôt sédimentaire et également des modifications des conditions physico-chimiques au cours du temps. Ces facteurs comprennent notamment le type de minéraux présents, le pH du milieu, la concentration et la valence des cations présents ( $\text{Na}^+$ ,  $\text{K}^+$ ,  $\text{NH}_4$ , Torti *et al.* 2015). La protection de l'ADN extracellulaire contre les processus biologiques (nucléases microbiennes) ou chimiques (hydrolyse, désoxygénation) (Dell'Anno & Corinadelsi 2004, Corinadelsi *et al.* 2008, Torti *et al.* 2015) via l'adsorption aux particules dépend de ces conditions physico-chimiques. Ces travaux de thèse n'apportent pas de connaissances concernant cet aspect de calibration. Il serait pourtant intéressant d'envisager une étude de suivi temporel des conditions physico-chimiques de l'interface eau/sédiment et des strates sédimentaires superficielles (profils sédimentaires), pour cerner la variabilité de ces conditions ; en parallèle le couplage avec des expérimentations visant à tester les effets de



ces changements de conditions physico-chimiques sur l'ADN de la communauté pourrait être mené via par exemple des tests en mésocosmes.

- la **biosphère active des sédiments** : le sédiment est l'archive naturelle des conditions passées mais il s'agit également d'un écosystème actif (importante zone de re-minéralisation et recyclage de la matière organique). Des microorganismes, principalement des bactéries et archées mais potentiellement aussi des Champignons, sont actifs dans les sédiments en particulier dans la partie supérieure des sédiments où se trouve une zone d'activité hétérotrophe dans laquelle la matière organique (dont l'ADN archivé) peut être dégradée (Fischer *et al.* 2002, Haglund *et al.* 2003, Wurzbacher *et al.* 2016). Par exemple, dans les sédiments du lac du Bourget, la présence de bactéries et archées actives a été détectée dans les 30 premiers cm de la carotte sédimentaire correspondant à une zone d'activité hétérotrophe importante illustrant le fait que cette biosphère active pourrait avoir une influence sur les molécules d'ADN archivées.

De plus, la présence de micro-eucaryotes vivant dans les sédiments pourrait induire un biais dans la caractérisation de la communauté microbienne eucaryote à partir de l'ADN sédimentaire. En effet, la présence de micro-eucaryotes actifs (principalement des champignons, ciliés, perkinsées) a été mise en évidence dans des sédiments marins mais principalement dans les premiers centimètres des carottes sédimentaires (Coolen & Shtereva 2009, Edgcomb *et al.* 2011, Orsi *et al.* 2013, Chambouvet *et al.* 2014), cette activité étant évaluée à l'aide d'analyses basées sur les transcrits d'ARNr des sédiments (RT-qPCR, métatranscriptomique). Au sein du groupe des champignons, considérés comme aérophiles, plusieurs espèces fongiques pourraient être actives en conditions anoxiques (Scupham *et al.* 2006) ce qui a notamment été suggéré par plusieurs études menées sur les sédiments (synthétisées dans Wurzbacher *et al.* 2010). Bien que les limites de l'utilisation de l'ARN comme indicateur de l'activité microbienne sont encore discutées (Blazewicz *et al.* 2013), cette approche reste aujourd'hui une des plus fiables pour évaluer cette activité. Nous avons étudié l'activité des micro-eucaryotes dans les sédiments du lac du Bourget pour évaluer si les inventaires moléculaires obtenus à partir de l'ADN sédimentaire pourraient être biaisés par une communauté active de micro-eucaryotes (chapitre II). Des rétro-transcrits d'ARNr (ADNc) de micro-eucaryotes ont été amplifiés par PCR mais seulement dans les deux premiers centimètres de la carotte sédimentaire alors que pour les bactéries et les archées, des amplifications PCR ont été obtenus sur les 30 premiers centimètres de la même carotte (données non présentées). Le séquençage de banques de clones a permis de montrer que ceux

sont principalement des OTUs de Ciliés qui ont été amplifiés mais ces taxons n'ont pas été retrouvés comme abondants dans les inventaires moléculaires retraçant les communautés passées à partir de ces sédiments. Ainsi, pour ce lac, l'activité biologique des micro-eucaryotes ne semble pas avoir eu d'influence sur le signal ADN qui représente essentiellement l'ADN du passé. Pour le lac Nylandssjön, la présence de champignons actifs est suspectée et pourrait expliquer les différences dans la proportion et la diversité des champignons (dont dikarya, chytrides, gloméromycètes, voir Figure **III.S2**) détectées dans les 2 carottes, alors même que les autres groupes de micro-eucaryotes sont très stables.

#### **VI.1.5. Focus sur quelques choix méthodologiques et les potentielles limites concernant l'approche déployée dans ces travaux**

Parmi les potentielles limites méthodologiques pouvant impacter la description faite des communautés microbiennes via les inventaires moléculaires obtenues à partir d'ADN sédimentaire, nous avons effectué certains choix concernant :

(i) le choix de la séquence nucléotidique ciblée qui est un paramètre important pour les études d'écologie microbienne et peut sensiblement impacter la lecture faite de la diversité et la composition des communautés microbiennes eucaryotes (e.g. Taïb *et al.* 2013, Bradley *et al.* 2016). Concernant les barcodes utilisés dans ces travaux, comme expliqué précédemment (parties I.2.2 et VI.1.3.), le choix a été fait d'une part, après des tests comparatifs d'amplification PCR avec un fragment court (V7 260 b) et un fragment long (V4 350 pb) pour l'ADN extrait des sédiments, d'autre part après des tests menés au LMGE pour comparer le taux de couverture des différentes régions (V1 à V9) et les potentielles différences d'assignation taxonomique en fonction des régions ciblées et ceci pour les outils bioinformatiques utilisés (comme expliqué dans le chapitre IV). L'ensemble de ces essais nous ont amené à travailler soit avec le V4 pour les sédiments très récents (< 60 ans) soit avec le V7 pour les sédiments plus anciens. L'avantage de l'utilisation du barcode V4 a été, compte tenu de sa longueur, de permettre une classification phylogénétique assez robuste (Outils PANAM <https://github.com/panammeb>, Taïb *et al.* 2013). A partir des OTUs 95 %, un arbre phylogénétique a ainsi été construit et a permis de regrouper les OTUs au sein d'unités phylogénétiques (PU) qui semblent être un bon proxy de l'espèce écologique (voir explication matériels et méthodes du chapitre II). Ceci n'a pas été appliqué pour le barcode V7 plus court (réduisant la robustesse des analyses phylogénétiques), qui est utilisé dans les travaux des chapitres III, IV et V. Lorsque c'est possible (fragment ADN assez long amplifiables), l'utilisation des PU est intéressante en paléolimnologie, elle permet en effet de regrouper des

OTUs et de travailler à une résolution taxonomique un peu moins fine que celle de l'OTU, ceci peut être utile dans un cas de polymorphisme lié à un vieillissement de l'ADN (possible augmentation du nombre d'OTUs pour un taxon donné).

(ii) les modalités d'extraction de l'ADN

Outre les précautions pour prévenir la contamination par de l'ADN extérieur (travail en conditions stériles et en utilisant des hôtes à flux laminaires localisées dans des salles différentes pour chaque étape de la procédure) et la vérification des blancs aux différentes étapes du processus (contrôles négatifs aux étapes de sous-échantillonnage des sédiments, extraction de l'ADN et amplification PCR), nous n'avons pas dans ces travaux testé de manière spécifique l'impact que peut avoir la méthode d'extraction ADN utilisé. Nous avons utilisé un kit standard pour l'extraction de l'ADN (Macherey Nagel NucleoSpin Soil). Comme mentionné précédemment, la distinction d'ADN intracellulaire et extracellulaire pourrait être envisagée afin de cerner les répercussions sur les inventaires moléculaires qui découlent de ces différentes fractions d'ADN (voir notamment les études de Nguyen *et al.* 2011, Eland *et al.* 2012, Willner *et al.* 2012, Wesolowska-Andersen *et al.* 2014, Rubin *et al.* 2014, Wagner Mackenzie *et al.* 2015). Toutefois un autre paramètre potentiellement important concernant l'extraction de l'ADN des sédiments est la quantité de sédiments utilisée. En effet, en fonction de la quantité de sédiments utilisés, la diversité biologique recélée et donc révélée dans l'inventaire moléculaire pourrait être modifiée. Nous avons utilisé, dans ces travaux de thèse, 0.5 g de sédiments pour réaliser toutes les extractions d'ADN en duplicats (car c'est la quantité maximale pour l'utilisation du kit) mais il existe d'autres procédures permettant de travailler avec une quantité de sédiment plus importante notamment les kits PowerMax soil isolation (Mobio Laboratories Inc., Carlsbad, California, USA). Klouch *et al.* (2016) ont par exemple extrait l'ADN à partir de 5-7 g de sédiments en utilisant ce kit pour étudier la dynamique temporelle d'un Dinoflagellé (*Alexandrium minutum*). Dans nos travaux, un autre choix a été fait pour essayer de pallier à une potentielle hétérogénéité de l'ADN dans une strate sédimentaire (via le séquençage de duplicats pour chaque strate), mais l'utilisation d'une quantité supérieure de sédiments pour l'extraction de l'ADN pourrait avoir eu une influence sur la richesse totale obtenue.

(iii) la stringence des traitements bioinformatiques et la prise en compte de répliques (plusieurs inventaires moléculaires par strate sédimentaire)

Le nettoyage des données issues du séquençage massif nous a permis d'obtenir des séquences nucléotidiques de qualité. Ainsi pour les barcodes V7, les séquences avec des bases indéfinies, d'une taille trop petite (en fonction des barcodes choisis) et pour lesquelles des erreurs de séquençage ont été retrouvées dans les amorces nucléotidiques, ont été retirées des analyses. En complément, les séquences chimériques ont été retirées des jeux de données. Pour les deux barcodes (régions V4 et V7), nous avons choisi de regrouper les séquences nucléotidiques au sein d'OTUs dont les séquences possèdent une similarité de 95 %. Comme expliqué dans la partie Matériels et Méthodes du chapitre IV, ce choix de seuil pour des OTUs de micro-eucaryotes a été fait après différents tests et vise à réduire la présence d'OTUS artificielles. Par ailleurs afin de consolider les inventaires produits nous avons également appliqué la stratégie de conserver uniquement les OTUs présentes dans les 2 réplicats de séquençage pour chaque strate (2 extraits d'ADN par strate). Cette étape permet de retirer non seulement les singletons mais également des OTUs qui peuvent être relativement abondantes dans 1 des 2 inventaires mais absentes dans l'autre. Seulement quelques OTUs avec ce type de patron ont pu être détectées lors de ces traitements mais la grande majorité des OTUs retirées étaient peu abondantes en nombre de 'reads'.

Dans la majorité des cas, les deux réplicats partagent une richesse commune relativement faible (inférieur à 50 % des OTUs inventoriées) comme cela est illustré ci-dessous pour la richesse (nombre d'OTUs) spécifique ou commune aux deux réplicats de séquençage (2 extraits ADN distincts) d'une même strate sédimentaire (échantillons sédimentaires du lac du Bourget, Figure VI.2). En conservant uniquement les OTUs communes aux deux réplicats, nous avons privilégié la 'robustesse' des OTUs conservées pour l'analyse des dynamiques temporelles de la communauté.

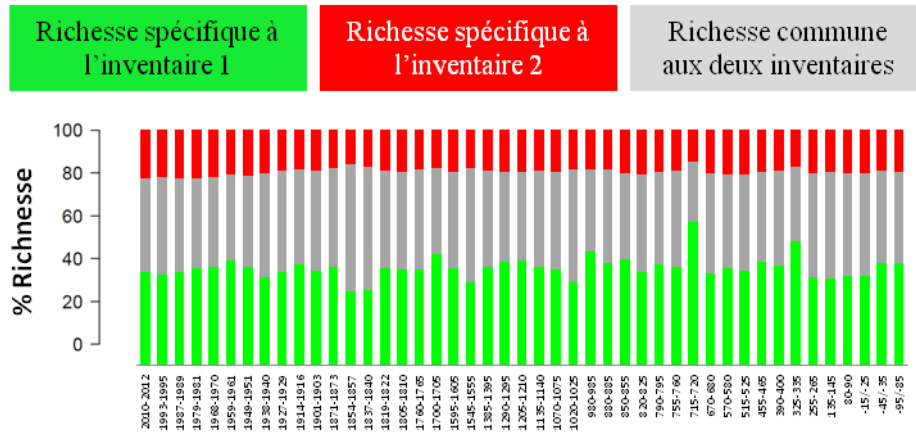


Figure VI.2. Pourcentage de richesse commune et spécifique pour chaque répliat de séquençage obtenu à partir des 2 extraits d'ADN issus d'une même strate sédimentaire (échantillons sédimentaires du lac du Bourget).

## **VI.2. Vision à long-terme de la dynamique temporelle des communautés lacustres de micro-eucaryotes**

### **VI.2.1. Vers une étude de la dynamique temporelle des communautés multi-échelles : les réarrangements de la communauté vus à court et à long terme**

L'étude de la dynamique temporelle des communautés microbiennes est importante pour comprendre leur rôle dans le fonctionnement dans les écosystèmes et leur réponse aux forçages environnementaux (Allison & Martiny 2008, Shade *et al.* 2012, Faust *et al.* 2015). Il peut être intéressant de s'interroger sur la pertinence du pas de temps pris en compte pour étudier les réarrangements au sein des communautés microbiennes en lien avec les variations des conditions environnementales ou les perturbations. Alors que quelques études s'intéressent aux réarrangements des communautés planctoniques eucaryotes à très court terme (par exemple Mangot *et al.* 2013), la plupart des suivis environnementaux s'effectuent actuellement à partir d'échantillons prélevés tous les mois (ou deux fois par mois) et les plus longues séries temporelles disponibles et adéquates pour les analyses moléculaires représentent au mieux 10 à 15 ans (principalement en milieu marin, Giovannoni & Vergin 2012, Gilbert *et al.* 2012, Chow *et al.* 2013, Cram *et al.* 2015). Ainsi, certaines études illustrent la forte dynamique saisonnière dans la composition et la structure des assemblages de micro-eucaryotes lacustres (Nolte *et al.* 2010, Charvet *et al.* 2012, 2014, Simon *et al.* 2015, Debroas *et al.* 2015) ; mais il y a à l'heure actuelle peu de travaux sur la dynamique à long terme de ces communautés en particulier en milieu lacustre.

Les données apportées par ces travaux de thèse sont donc assez originales ; elles révèlent de manière inédite la concomitance des changements observés dans la dynamique temporelle des communautés de micro-eucaryotes avec ceux observés dans les conditions environnementales (pressions anthropiques locales et climatiques) au cours des 2200 dernières années ou de pas de temps plus restreints (100 dernières années). Les modifications observées dans ces assemblages biologiques concernent à la fois leur richesse, leur composition et leur structure. Parmi les résultats majeurs de ces travaux, de nouvelles informations sont apportées concernant l'existence et le 'timing' des réarrangements dans les assemblages microbiens eucaryotes, la présence de taxons persistants au cours du temps, la détection de groupes d'OTUs co-occurrentes au cours du temps, la contribution de la rare biosphère dans les changements temporels à long terme et la réponse de ces assemblages eucaryotes aux forçages environnementaux. Les connaissances relatives à la dynamique temporelle des communautés de micro-eucaryotes lacustres bénéficieraient du couplage entre approches limnologiques et

paléolimnologiques afin de comparer l'amplitude des changements s'opérant sur des pas de temps contrastés (journalier, mensuel, annuel, pluriannuel) ainsi que de connaissances supplémentaires concernant l'identité des taxons impliqués dans ces réarrangements. A titre d'exemple, la biosphère rare semblait peu impliquée dans la dynamique à court terme (quelques jours) comme rapporté par Mangot *et al.* (2013) alors que dans nos travaux ces taxons rares semblent pouvoir être mobilisés en réponse à des forçages environnementaux à long terme (réponse à l'eutrophisation par exemple). Des études emboitant des échelles de temps contrastées sur un ou plusieurs lacs permettraient notamment de caractériser des taxons importants dans la communauté pour la réponse aux variations des conditions environnementales.

### **VI.2.2. Dynamique temporelle des taxons abondants**

Un apport de ces travaux de thèse aux connaissances en écologie microbienne concerne le patron de distribution de l'abondance des espèces (SAD, Shade & Gilbert 2015) et vient en particulier alimenter le **concept d'hyperdominance** (comme décrit dans de Vargas *et al.* 2015). Ce concept, suggérant qu'une communauté microbienne est composée principalement par quelques OTUs dominantes et une majorité de taxons rares, a été mise en évidence par les premières études utilisant le séquençage massif de l'ADN environnemental pour l'étude de la diversité microbienne (e.g. Sogin *et al.* 2006, Pedrós-Alió 2012) et depuis dans de nombreux écosystèmes (Bates *et al.* 2013, de Vargas *et al.* 2015, Filker *et al.* 2016). Tous les inventaires moléculaires obtenus dans ces travaux de thèse ont été caractérisés par la présence de quelques OTUs dominantes et une majorité d'OTUs rares. Par exemple, dans les inventaires moléculaires des lacs Annecy et Bourget (couverture temporelle : 100 ans, chapitre V), moins de 0.5 % des OTUs représentent plus de 50 % du nombre total de reads. Toutefois, de larges oscillations ont été observées dans l'abondance relative de ces OTUs au cours du temps (comme illustré pour quelques OTUs abondantes de la série temporelle du lac Annecy, voir Figure VI.3). Ces résultats soulignent que le caractère dit 'hyperdominant' d'un taxon doit être défini au regard de sa dynamique temporelle et pourrait varier selon le pas de temps considéré. Globalement, ces travaux de thèse (chapitres IV et V) révèlent des grandes oscillations dans l'abondance de ces OTUs hyper-dominantes sur des pas de temps très longs (plusieurs décennies voir plusieurs millénaires) illustrant le fait qu'une vision temporelle devrait être intégrée pour la description des communautés microbiennes comme suggéré par Shade & Gilbert (2015).

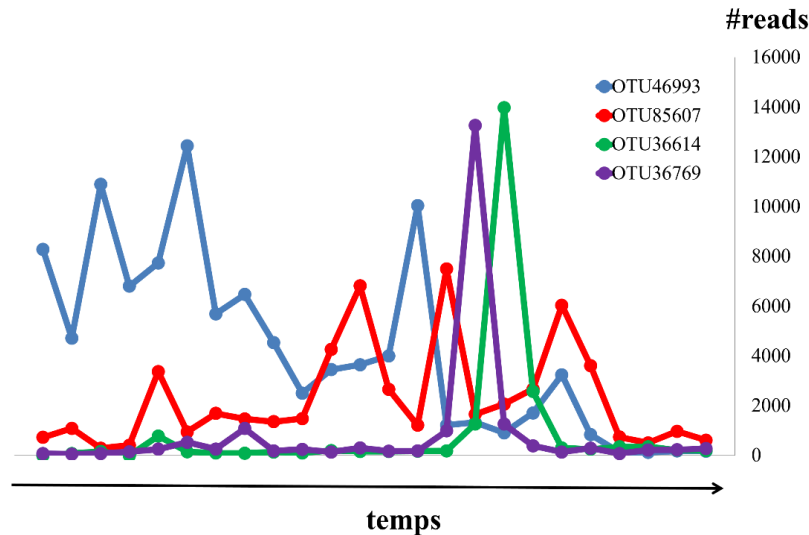


Figure VI.3. Abondance relative des OTUs les plus abondantes dans la série temporelle de la communauté microbienne eucaryote du lac Annecy au cours des 100 dernières années.

Dans ces travaux nous nous sommes appuyés sur l'analyse des abondances relatives des OTUs (proportion en termes de nombre de séquences ADN). L'utilisation des données d'abondances relatives des OTUs au sein d'un assemblage est encore aujourd'hui débattue. En effet, il a été montré que le nombre de copies d'un même gène au sein du génome des micro-eucaryotes peut varier de 1 à plusieurs dizaines de milliers (Zhu *et al.* 2005, Godhe *et al.* 2008). Ainsi, un taxon hyper-dominant au sein d'un assemblage pourrait correspondre à un nombre modéré de cellules possédant un nombre de copies très élevé et apporter une distorsion entre l'image fournie par l'abondance des OTUs et leur abondance cellulaire réelle comme rapporté par l'étude de Props *et al.* (2016) pour les bactéries de réservoir d'eau douce. Cependant, certaines études suggèrent que l'abondance relative des OTUs pourrait être une vision pertinente pour refléter l'abondance cellulaire notamment pour les micro-eucaryotes (de Vargas *et al.* 2015, Giner *et al.* 2016). Ainsi, Giner *et al.* (2016) ont mis en évidence une corrélation entre les inventaires microscopique (abondance cellulaire) et moléculaire (nombre de séquences nucléotidiques) dans des échantillons d'eau. Certainement, des calibrations doivent encore être réalisées afin de cerner les biais liés à l'utilisation des proportions relatives issues d'inventaires moléculaires comme proxy de l'abondance des divers groupes micro-eucaryotes lacustres ; toutefois, dans le cadre de nos travaux nous avons choisi de considérer les changements de proportion relative des OTUs au cours du temps comme une information a priori pertinente. Ce choix est conforté d'une part, par une cohérence entre les données ADN sédimentaire et des données de suivi de la masse d'eau (en termes de



proportion des ‘reads’), il s’agit notamment des données présentées dans le chapitre II, mais également de travaux précédemment effectués en ciblant l’ADN cyanobactérien (Savichtcheva *et al.* 2015). Par ailleurs, dans la majorité des analyses effectuées nous avons vérifié la cohérence des résultats acquis en considérant les proportions de ‘reads’ et en considérant seulement la présence/absence des OTUs (utilisation des indices de Bray-Curtis et Jaccard, respectivement).

### **VI.2.3. Contribution de la rare biosphère aux réarrangements**

L’importance relative de la rare biosphère au sein des inventaires microbiens a été découvert grâce au séquençage à haut débit de l’ADN environnemental (Sogin *et al.* 2006) et a entraîné plusieurs questions pour les études d’écologie microbienne concernant (i) la réalité de ces rares taxons, représentés par quelques séquences d’ADN, (ii) leur rôle potentiel dans le fonctionnement des écosystèmes.

La **réalité des taxons rares** est aujourd’hui bien admise dans la communauté scientifique au vu du nombre d’études et de synthèses portant sur ce compartiment (Jones & Lennon 2010, Pedrós-Alió 2012, Logares *et al.* 2015, Lynch & Neufeld 2015, Shade & Gilbert 2015). Cependant, il est important de noter que leur présence et richesse dans les inventaires est fortement dépendant des traitements bioinformatiques. Ainsi, le traitement des données issues du séquençage à haut débit de l’ADN peut entraîner plusieurs biais dans l’évaluation de la part de la biosphère rare (en lien avec la profondeur de séquençage, le traitement des séquences rares et la stringence des processus bioinformatiques). La profondeur de séquençage, variant de quelques milliers à 1.7 million de reads par échantillon en fonction des études (de Vargas *et al.* 2015 pour le maximum), impacte fortement le nombre d’OTUs rares retrouvés dans les assemblages. Dans les études écologiques présentées ici (chapitre IV et V), le nombre de reads, environ 30-35 000 par échantillon, peut être considéré comme faible pour décrire la diversité d’une communauté mais reste suffisant pour étudier les réarrangements dans une communauté au cours du temps. De plus, les courbes de raréfaction réalisées en complément de ces travaux montrent qu’avec une profondeur de séquençage de 35 000 reads, une grande partie de la richesse est retrouvée dans les inventaires moléculaires (Figure **VI.4**).

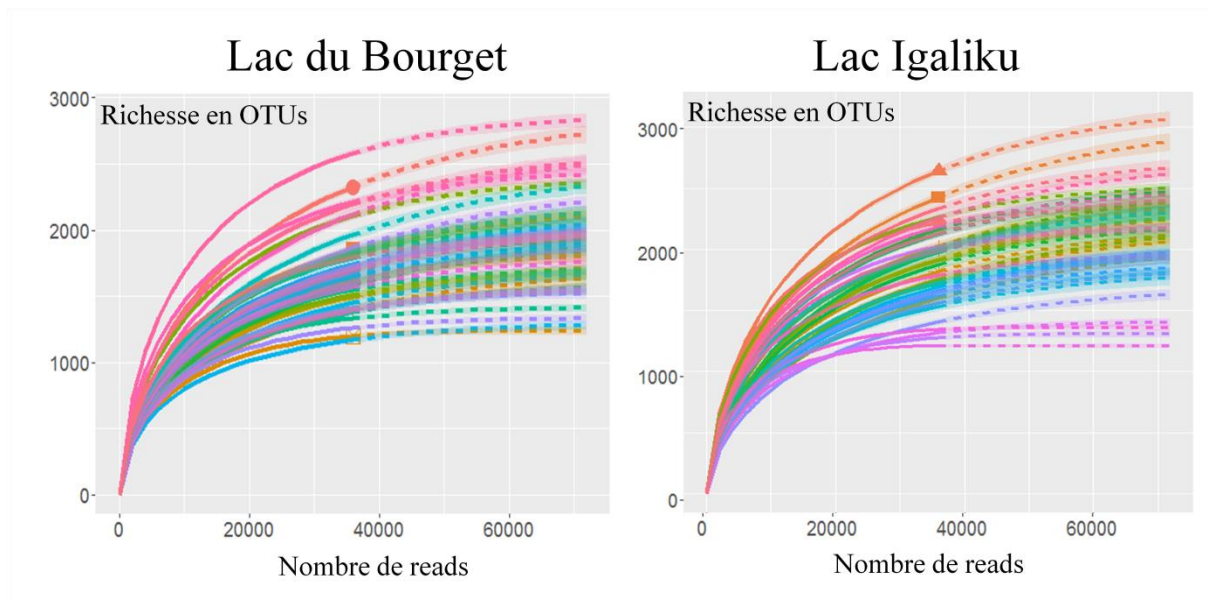


Figure VI.4. Courbes de Raréfaction/Extrapolation obtenues pour la richesse spécifique de chaque inventaire moléculaire des analyses du chapitre IV.

Cette profondeur de séquençage relativement modérée est due au traitement utilisé pour les analyses des inventaires moléculaires. En effet, afin de s'assurer que les OTUs présentes dans les inventaires ne sont pas des artefacts dus à des erreurs de séquençage, nous avons choisi de ne garder que les OTUs présentes dans les deux répliquats d'inventaires moléculaires obtenus pour chaque échantillon sédimentaire. Ainsi, une grande partie des OTUs est retirée à cette étape réduisant la table d'OTUs (de 80 081 à 16 396 OTUs dans les données du chapitre IV). Une autre caractéristique importante dans l'étude de la biosphère rare est le seuil pris en compte pour caractériser les taxons rares. Une OTU peut être donc définie comme rare si elle possède moins de 0.01 ou 0.1 % des reads de la série temporelle (e.g. Logares *et al.* 2015). Dans cette thèse, nous avons choisis de définir plusieurs niveaux de rareté (Figure VI.5), également lié à un seuil arbitraire mais défini en lien avec les changements temporels dans l'abondance relative des taxons (i) les taxons dit *permanently rares* sont définis comme les OTUs qui ne possèdent pas plus de 0.1 % des séquences pour chaque échantillon (ii) les *oscillating OTUs* sont caractérisées par de larges oscillations dans leur abondance relative au cours du temps transitant d'état abondant à rare (< 0.01 %), ou vice versa. Dans ces travaux de thèse, nous avons commencé à explorer la dynamique des CRT (*conditionally rare taxa*, tels que défini par Shade *et al.* 2014) qui sont des OTUs devenant abondantes très ponctuellement au cours du temps (alors que les *oscillating OTUs* restent à l'état rare ou dominant sur des périodes longues). La dynamique des CRT peut être intéressante à

considérée comme le suggère les analyses de réseaux dans lesquels ils apparaissent et sont impliqués au sein de modules dont la dynamique temporelle est liée aux changements environnementaux. C'est par exemple le cas pour des OTUs Alveolate\_1 liées aux changements de phosphore dans les lacs Annecy et Bourget (chapitre V).

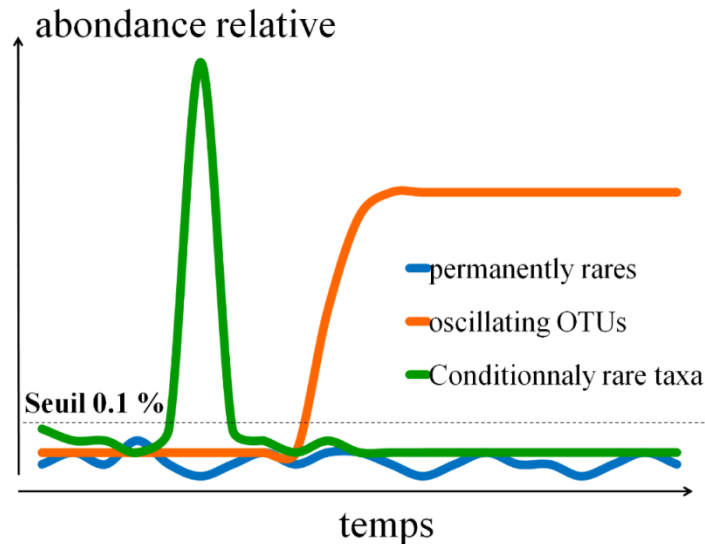


Figure VI.5. Schéma illustrant la définition des niveaux de rareté considérés dans les inventaires moléculaires. Bien que les *conditionally rare taxa* n'est pas été étudié en tant que tel, les traitements des seuils de rareté ont permis de les intégrer dans les réseaux de cooccurrence.

**Le rôle des taxons rares** dans le fonctionnement des écosystèmes est encore peu compris pour les micro-eucaryotes même si des études récentes ont mis en évidence l'activité de certains micro-eucaryotes rares dans plusieurs milieux aquatiques (Whilhem *et al.* 2014, Logares *et al.* 2014, Debroas *et al.* 2015) suggérant que la rare biosphère pourrait avoir un rôle important dans le fonctionnement des écosystèmes. Plusieurs hypothèses (non exclusives) sont attribuées au rôle de la rare biosphère : elle pourrait agir comme un réservoir de diversité via par exemple des taxons en dormance pouvant être mobilisés en cas de perturbation de l'écosystème, contribuer à la redondance fonctionnelle au sein de la communauté (Jones & Lennon 2010, Logares *et al.* 2015), ou représenter des taxons actifs qui restent à l'état de rares mais contribuent au fonctionnement de l'écosystème (e.g. Debroas *et al.* 2015).

Des travaux ont déjà rapportés le rôle important des 'rares' dans la réponse de la communauté aux forçages environnementaux (Sjöstedt *et al.* 2012, Shade *et al.* 2014, Baltar *et al.* 2015, Lynch & Neufeld 2015) notamment au sein des communautés de micro-eucaryotes (Logares *et al.* 2014, 2015, Shade & Gilbert 2015); nos analyses (chapitres IV et V) tendent à

confirmer ceci sur la base d'une vision à long terme des micro-eucaryotes lacustres. Les inventaires moléculaires issus de la série temporelle longue (2200 ans) du lac du Bourget ont ainsi permis de mettre en évidence l'émergence (en terme de nombre de séquences d'ADN) de plusieurs taxons rares en lien avec l'eutrophisation du lac du Bourget. En effet, 16 taxons affiliés à des chlorophytes sont très abondants à partir des années 1960s et contribuent à plus de 55 % des réarrangements temporels dans la structure de la communauté dans les années 1980s (Figure IV.6). Cette augmentation de chlorophytes a également été observée dans l'étude portant sur les 100 dernières années où un module principalement composé de chlorophytes a été trouvé fortement corrélé aux augmentations des teneurs en phosphore (Figure V.4). Au cours de ces 100 dernières années, d'autres changements ont été observés notamment concernant l'abondance relative de plusieurs micro-eucaryotes rares. En effet, l'augmentation des teneurs en phosphore dans la masse d'eau du lac du Bourget au début des années 1940s semble avoir eu un impact plus fort sur la rare biosphère que les augmentations des teneurs en phosphore dans le lac d'Annecy car un nombre plus élevé d'OTUs (76 OTUs) sont devenus rares lors de ce point de changement dans le lac du Bourget (Figure V.7).

#### **VI.2.4. Les données temporelles au service de la description d'un *core microbiome* eucaryote**

Les données à long terme obtenues dans cette thèse permettent d'apporter des nouvelles connaissances pour caractériser un *core microbiome* temporel pour les communautés de micro-eucaryotes de plusieurs lacs. Le *core microbiome* tel que définie par Shade & Handelsman (2012) est caractérisé par plusieurs paramètres de la communauté dont la persistance d'OTUs dans tous les échantillons d'une série temporelle et la co-occurrence temporelle de certaines OTUs (synchronie au cours du temps).

La persistance de taxons dans des séries temporelles longues pourrait être une caractéristique importante pour la stabilité d'une communauté et sa réponse aux forçages environnementaux (notamment l'eutrophisation et les changements climatiques). Cette notion de persistance a déjà été mise en évidence pour différentes communautés par exemple dans les écosystèmes humains ou marins (Gilbert *et al.* 2009, Caporaso *et al.* 2011, Shade & Handelsman 2012, Gibbons *et al.* 2013) mais jamais en intégrant des données temporelles aussi longues. Dans ces travaux de thèse, la persistance d'OTUs de micro-eucaryotes a été révélée notamment au cours de 2 millénaires dans les communautés des lacs du Bourget et Igaliku (chapitre IV). Ainsi, moins de 1 % des OTUs ont été retrouvées comme persistantes au cours du temps et

représentent plus de 50 % de l'abondance en nombre de reads dans l'assemblage. Il est possible qu'une grande partie des micro-eucaryotes rares soient également persistants au cours du temps mais que la profondeur de séquençage dans chaque échantillon et/ou la stringence des analyses bioinformatiques empêchent leur détection comme cela a été suggéré pour les bactéries rares persistantes en milieu marin (Gibbons *et al.* 2013). Une grande diversité est observée dans la composition des *core microbiomes* caractérisés par la présence de ciliés, de chlorophytes, de cercozoaires ou encore de champignons illustrant le fait que ces persistants possèdent des fonctions variées et importantes pour le fonctionnement des écosystèmes (dont la photosynthèse, le recyclage et la production de matière organique). En perspective des travaux réalisés sur quelques lacs, il serait intéressant de dresser un portrait du *core microbiome* (taxons persistants) de lacs provenant de différentes écorégions en se basant sur des données temporelles contrastées (limnologiques et paléolimnologiques). Cette étude de comparaison permettrait de détecter des taxons communs à plusieurs types d'écosystème et ainsi peut être définir un *core microbiome* global des milieux lacustres (s'il existe).

L'utilisation de données moléculaires issues de séries temporelles permet de créer des réseaux de co-occurrence basés sur les changements dans l'abondance relative des OTUs au cours du temps (Faust *et al.* 2015). Chaque paire d'OTUs est ainsi caractérisée par une valeur de corrélation en fonction de la synchronie des OTUs dans le temps. La plupart des études se basent sur des données spatiales pour définir ces interconnexions (c'est-à-dire des corrélations basées sur l'abondance relative des OTUs dans les écosystèmes étudiés, Chaffron *et al.* 2010, Eiler *et al.* 2012, Faust & Raes 2012, Lima-Mendez *et al.* 2015) quand certaines, incluant des protistes, révèlent la synchronie des OTUs au cours du temps (dans les milieux lacustres : Vick-Majors *et al.* 2014, Kammerlander *et al.* 2015). Les réseaux de co-occurrence permettent d'obtenir de nouveaux indices pour caractériser la communauté microbienne : le degré de connectivité des taxons, le diamètre du réseau, sa densité ou encore sa modularité. Une de ces caractéristiques les plus intéressantes est le regroupement des taxons fortement connectés entre eux au sein de modules. Dans le chapitre IV et V, les réseaux de cooccurrence construits à partir des séries temporelles (2200 ans ou 100 ans) des lacs d'Annecy, du Bourget et Igaliku ont permis de détecter d'une part des modules plutôt diversifiés (comme ceux regroupant des OTUs de cercozoaires, dinophycées, ciliés et champignons) et d'autre part des modules regroupant des OTUs à la taxonomie très similaire comme ceux composés de ciliés Alveolate\_1. Ces derniers (modules mono-groupe taxonomique) pourraient résulter de (i) la présence de taxons très proches génétiquement possédant une 'niche écologique' semblable (ii) d'un biais méthodologique engendrant l'occurrence de plusieurs OTUs alors que les

séquences ne proviennent que d'un seul taxon (erreur de séquençage par exemple) (iii) d'un biais lié plus particulièrement au seuil de clustérisation de 95 % qui ne serait pas forcément toujours adapté pour définir des espèces de micro-eucaryotes en particulier pour l'ADN sédimentaire (en lien avec un potentiel polymorphisme génétique engendré par des processus diagénétiques dans les sédiments). Les données de topologie des réseaux présentées dans ces analyses permettent de lier les changements temporels de la structure des modules avec les forçages environnementaux (discuté dans la partie VI.3). Cette vision de la cooccurrence des OTUs de micro-eucaryotes au cours du siècle dernier dans les lacs Annecy et Bourget suggère que ces communautés ne se sont pas structurées de façon aléatoire avec le temps. Contrairement aux réseaux de cooccurrence au sein des communautés bactériennes, le coefficient de clustérisation de ces 2 réseaux à des valeurs supérieures à 0.6 montrant des liens plus forts entre les OTUs des communautés microbiennes eucaryotes que bactériennes (Chaffron *et al.* 2010, Steele *et al.* 2011, Lima-Mendez *et al.* 2015). Ceci semble logique si l'on considère les multiples interactions se déroulant au sein même de la communauté microbienne eucaryote (proie/prédateur, hôte/symbionte, hôte/parasite, coopération) en comparaison avec les interactions trophiques connues au sein des communautés bactériennes. La présence de taxons connecteurs (qui connectent fortement les modules entre eux), pouvant représenter des taxons clés dans la structure du réseau, a également été détectée dans les réseaux de cooccurrence. Cependant, la plupart de ces OTUs ont une classification grossière (au niveau taxonomique Alveolata, Chlorophyta ou encore Dinophyceae). Il n'est donc pas possible d'inférer leurs rôles fonctionnels dans le milieu sauf pour certains taxons appartenant à des groupes exclusivement photosynthétiques. Un des points importants dans les études d'écologie microbienne est de compléter des bases de données de séquences nucléotidiques eucaryotes (par isolement, mise en culture et séquençage du génome de chaque organisme) pour mieux caractériser la diversité taxonomique au sein des communautés de micro-eucaryotes (comme par exemple dans le projet SINGEK <http://www.singek.eu/>).

### **VI.3. La réponse des communautés de micro-eucaryotes aux pressions anthropiques et climatiques**

L'eutrophisation et le changement climatique sont deux perturbations environnementales connues pour entraîner des modifications dans la communauté biologique des écosystèmes lacustres (Adrian *et al.* 2009, Moss 2012, Randsalu-Wendrup *et al.* 2016). L'étude des forçages environnementaux dans ces travaux de thèse a porté sur ces deux facteurs. Les périodes au cours desquelles ils semblent avoir impacté les communautés des 3 lacs étudiés sont présentés sur la Figure **VI.6**. Les analyses menées dans les travaux des chapitres IV et V ont permis de mettre en évidence (i) la réponse des communautés eucaryotes à des niveaux d'eutrophisation contrastés (ii) les impacts des changements de teneurs en phosphore sur ces communautés (en terme de richesse, composition, structure) et sur les potentielles interactions biologiques au sein des réseaux trophiques (iii) l'existence de potentiels indicateurs biologiques jusque-là non considérés dans les proxies paléolimnologiques, comme par exemple les chlorophytes répondant positivement aux fortes augmentations des teneurs en phosphore (iv) des résultats inédits concernant l'impact des fluctuations climatiques au cours des 2200 dernières années sur les assemblages de micro-eucaryotes et (v) les effets combinés du réchauffement climatique et de la ré-oligotrophisation sur les communautés de micro-eucaryotes des lacs expliquant probablement la trajectoire de non-retour à un état de pré-eutrophisation pour ces communautés.

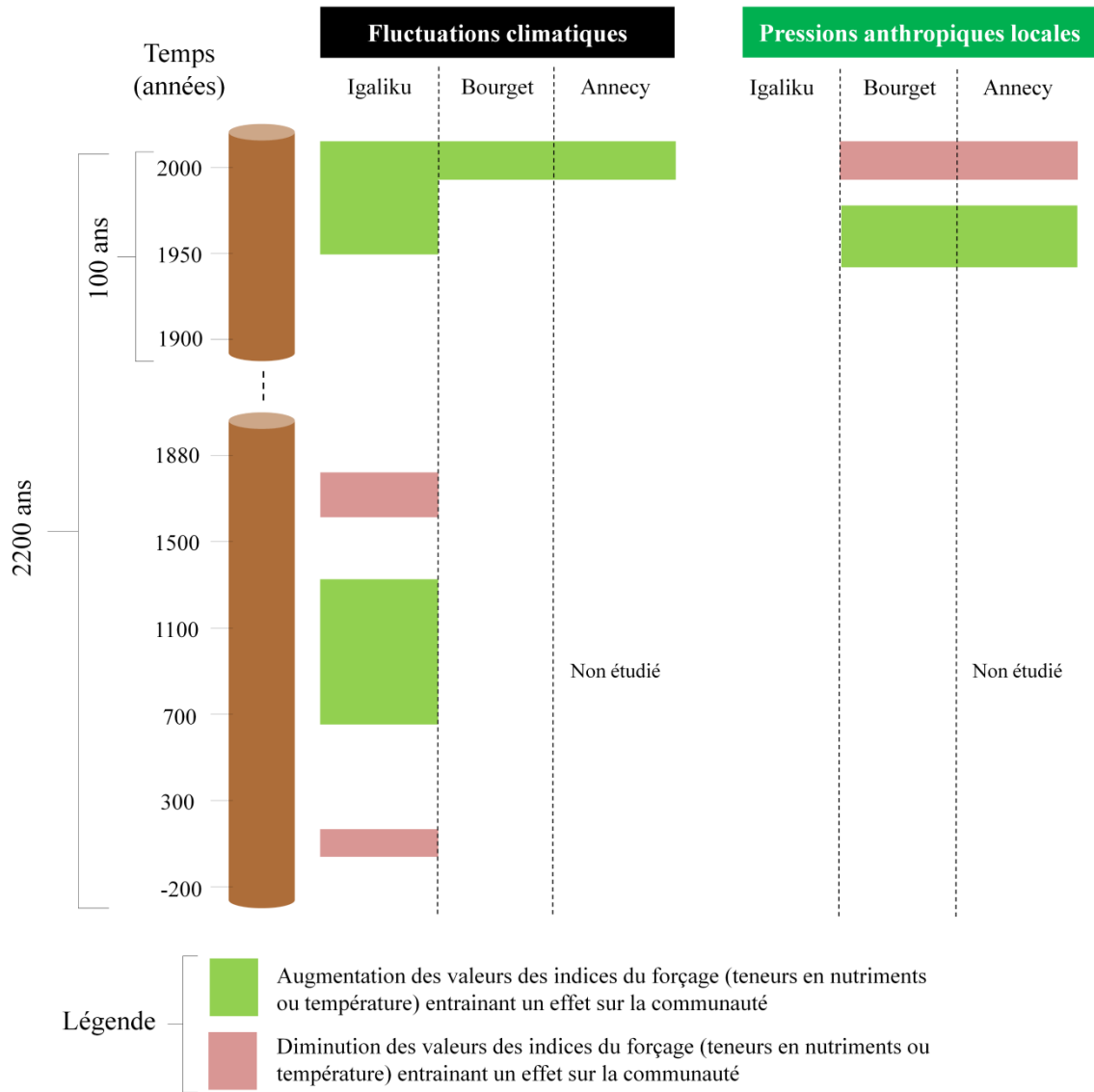


Figure VI.6: Schéma simplifié illustrant les périodes temporelles au cours desquelles les variations des deux forçages pris en compte semblent avoir eu un effet sur la communauté de micro-eucaryotes des lacs étudiés.



### **VI.3.1. Réponse des communautés de micro-eucaryotes lacustres à différents niveaux d'eutrophisation**

Les données obtenues via l'ADN archivé dans les sédiments ont permis de mettre en évidence que l'effet des forçages anthropiques, en particulier l'augmentation des teneurs en phosphore, est détectable sur l'ensemble de la communauté de micro-eucaryotes lacustres, incluant non seulement des groupes phototrophes pour lesquels une réponse à l'eutrophisation était connue, mais aussi des groupes purement hétérotrophes. Alors que la communauté microbienne eucaryote du lac du Bourget semble avoir été relativement stable au cours des dernières 2200 années, un point de basculement a été détecté dans la richesse, la composition et la structure de cette communauté eucaryote à partir des années 1940s en lien avec les forçages environnementaux récents et notamment les fortes augmentations des teneurs en phosphore (Figure IV.I). La détection de ce point de basculement a été révélée par de 2 méthodes analytiques différentes. D'une part, à l'aide d'un outil analytique appelée MRT (*Multivariate Regression Tree*) se basant sur les données d'abondance relative des OTUs pour déterminer à quel moment il y a une rupture dans la communauté, et qui est donc une analyse contrainte dans le temps (*constrained analysis*). A l'opposée, la détermination de groupes d'échantillons en se basant sur le calcul de la dissimilarité de Bray-Curtis (et l'analyse nMDS qui en découle) n'est pas contrainte par le temps (*unconstrained analysis*) et a, pour le lac du Bourget, également permis de mettre en évidence un point de basculement majeur dans les années 1940s. L'utilisation de ces deux types d'analyse a donc conforté l'hypothèse que l'eutrophisation a impacté fortement la communauté de micro-eucaryote de ce lac ; les travaux du chapitre V, ciblant les 100 dernières années, ont fourni une résolution temporelle relative fine et ininterrompue permettant de relier assez clairement l'augmentation des teneurs en phosphore à partir des années 1940-50s à des modifications de richesse et structure des communautés de micro-eucaryotes dans le lac du Bourget, mais également dans le lac d'Annecy pourtant affecté par un enrichissement des eaux bien inférieur à celui du Bourget (teneurs maximales  $\sim 15 \mu\text{gP.L}^{-1}$  sur le lac d'Annecy). Ces données suggèrent que dans une situation de forte limitation par les nutriments comme c'est le cas sur ce lac oligotrophe, la communauté eucaryote peut répondre avec une forte amplitude même à des augmentations modérés de nutriments, dès lors qu'elles atteignent un certain seuil (il est utile de rappeler que dans les catégories trophiques des lacs, le passage de l'état oligotrophe à mésotrophe se situe aux environs de  $20 \mu\text{gP.L}^{-1}$ ).

Parmi les effets de ce forçage sur les communautés des deux lacs, il faut mentionner la relation unimodale qui a été détectée dans la richesse totale de la communauté eucaryote (incluant tous les groupes taxonomiques) en réponse aux modifications des concentrations en phosphore (avec des valeurs maximales de richesse pour des teneurs proches de  $15 \mu\text{gP.L}^{-1}$ ) ce qui est en adéquation avec les travaux portant sur l'impact de l'eutrophisation sur la richesse spécifique phytoplanctonique via des observations microscopiques (Reynolds *et al.* 1993, Sommer *et al.* 1993, Jeppesen *et al.* 2000, Soares *et al.* 2013).

### **VI.3.2. Effets de l'eutrophisation au sein d'un réseau trophique lacustre : quelques hypothèses issues des données paléolimnologiques concernant les interactions biotiques**

Les données obtenues pour les analyses sur les séries temporelles longues des lacs Bourget et Igaliku (2200 ans, chapitre IV) permettent d'avancer des hypothèses relatives aux interactions biotiques qui s'opèrent au sein des réseaux trophiques. L'un des exemples qui nous semble le plus convaincant est celui des ciliés du lac du Bourget qui semblent avoir été favorisés, en termes de richesse et d'abondance relative, par l'augmentation des teneurs en phosphore à partir des années 1940s, ceci alors même que l'on note une augmentation des abondances relatives des picocyanobactéries, connues pour être une source de nourriture des ciliés, dans ce lac comme le montre l'étude de Domaizon *et al.* (2013). Ces données suggèrent que ces deux groupes, probablement reliés par une relation trophique proie-prédateur présentent une dynamique comparable en réponse à l'eutrophisation du lac, sûrement en lien avec des pressions *bottom-up* sur le réseau trophique lacustre.

L'accroissement des teneurs en nutriments n'a cependant pas que des effets positifs sur les communautés notamment pour certains groupes d'hétérotrophes, tels que les apicomplexés, les perkinsés et les champignons. En ce qui concerne les groupes fongiques saprotrophes, leur déclin à partir des années 40s (lac du Bourget en particulier) pourrait être dû à une modification de la qualité de la matière organique dans la masse d'eau (apport de matière allochtone comme décrit dans Giguet-Covex *et al.* 2010), ces organismes étant connus pour dégrader préférentiellement des matières organiques complexes (Stelzer *et al.* 2003, Pascoal *et al.* 2005, Tank *et al.* 2010). Pour ce qui concerne les autres parasites (perkinsés notamment), il est difficile d'expliquer les raisons de leur déclin avec le changement d'état trophique au vu des connaissances actuelles encore très incomplètes concernant leurs interactions au sein des réseaux trophiques lacustres. Par exemple, les perkinsés sont connus pour parasiter des espèces phytoplanctoniques (notamment des chlorophytes et des

chrysophycées) mais il n'y a pas de connaissances précises concernant le spectre des hôtes de ces parasites et le niveau de spécificités de ces relations parasitaires (Mangot *et al.* 2011, 2013).

Les hypothèses qui peuvent être soulevées à partir de ces données historiques pourront être investiguées plus précisément, notamment par l'analyse plus spécifique des taxons concernés (ciblage spécifique de groupes parasites et de leurs hôtes potentiels par exemple).

### **VI.3.3. Les chlorophytes, de potentiels indicateurs de l'état trophique du lac ?**

Dans les assemblages eucaryotes, nos données ont montré que le groupe des chlorophytes est celui qui semble être le plus nettement favorisé par l'accroissement des teneurs en phosphore. L'augmentation de la richesse et l'abondance relative dans la communauté du lac du Bourget avec l'eutrophisation a été assez clairement détectée dans l'étude couvrant 2200 ans (chapitre IV) et a été confirmé avec une résolution d'échantillonnage plus fine pour les 100 dernières années (chapitre V), les chlorophytes devenant particulièrement abondantes (abondance relative) quand les teneurs en phosphore atteignent ou dépassent les  $80 \mu\text{g.L}^{-1}$  (d'après nos données). Ces résultats ont également été confirmés par des analyses de PCR quantitative réalisées sur les séries temporelles des lacs d'Annecy et du Bourget en utilisant des amorces nucléotidiques spécifiques aux chlorophytes (Turon *et al.* 2015, Figure V.8). Même si les quantifications doivent être considérées avec certaines précautions (en lien les biais inhérents au nombre de copies du gène 18S variable selon les espèces), la convergence des données issues des deux types d'analyses (abondance relative dans l'assemblage + abondance quantifiée par qPCR) tendent à confirmer la prédominance des chlorophytes dans l'assemblage en lien avec un fort niveau d'eutrophisation. De plus, ces résultats sont cohérents avec les connaissances concernant l'abondance de ces algues en conditions eutrophes révélée par des études limnologiques (Jensen *et al.* 1994, Naselli-Flores & Barone, 2000, Moss *et al.* 2003, Soares *et al.* 2013) et paléolimnologiques (Fee 1979, Engstrom *et al.* 2006). Les chlorophytes qui semblent devenir plus particulièrement dominantes pendant la période récente appartiennent (ou sont proches) au genre *Chlamydomonas*, un groupe connu pour être ubiquiste et abondant dans de nombreux lacs et plus particulièrement dans les lacs en conditions eutrophes (Naselli-Flores & Barone 2000, Lepère *et al.* 2008, Chen *et al.* 2008). Ce patron de distribution devrait être vérifié à partir de séries temporelles d'autres lacs afin de visualiser des gradients d'eutrophisation contrastées (incluant des systèmes ayant subis de fortes eutrophisations).

#### **VI.3.4. L'effet des fluctuations climatiques perceptibles sur les communautés eucaryotes**

De façon inédite, nous avons mis en évidence, pour le lac Igaliku, une concomitance entre les changements de structure de la communauté eucaryote et les grandes périodes climatiques définies au cours des dernières 2200 années (Mann 2003, Kaufman *et al.* 2009, McKay & Kaufman 2014). La communauté microbienne eucaryote du lac Igaliku semble en effet avoir été principalement impactée par les modifications des conditions climatiques (Groenland) au cours de cette période alors que la communauté du lac du Bourget a été plus fortement impactée par les forçages environnementaux récents (prioritairement l'eutrophisation et également le réchauffement climatique).

Nous avons ainsi mis en évidence d'une part (sur la base d'analyses non contraintes par le temps), une structuration des communautés cohérentes avec grandes périodes climatiques et d'autre part (sur la base d'analyses contraintes) que le point de basculement majeur de la communauté microbienne eucaryote a été détecté au cours du 17<sup>ème</sup> siècle dans le lac Igaliku ce qui correspond à la période la plus froide des 2200 dernières années (entre 1640 et 1780 ans ap JC). Les autres études réalisées sur les communautés biologiques du lac Igaliku n'ont mis en évidence que de faibles changements dans la dynamique temporelle des compartiments biologiques (dont chironomes et groupes phytoplanctoniques) avant les années 1980s (date de début d'activités agricoles modernes) ce qui contraste avec nos résultats (Perren *et al.* 2012, Millet *et al.* 2014).

Une des modifications importantes au sein de la communauté microbienne eucaryote du lac Igaliku au cours des 2200 dernières années est l'augmentation de la richesse et l'abondance des OTUs de cercozoaires pendant la période médiévale. Les cercozoaires sont connus pour être prédateurs de bactéries et de micro-algues dans les écosystèmes pélagiques et benthiques et ils peuvent également être parasites de taxons phytoplanctoniques (Tilman *et al.* 1999). Mais le lien entre la dynamique temporelle des cercozoaires et le réchauffement climatique est largement inconnue. Une réponse positive des cercozoaires aux augmentations de température a été récemment mise en évidence dans une tourbe arctique (Tveit *et al.* 2015) ce qui pourrait être un indice, avec nos résultats (chapitre IV) du rôle potentiel de ces organismes dans la réponse de la communauté au réchauffement climatique. En comparaison, les séries temporelles des communautés du lac du Bourget et Annecy obtenues dans le chapitre V montrent que les cercozoaires sont principalement impactés négativement par les modifications des teneurs en phosphore mais qu'aucune relation avec le réchauffement climatique récent au cours des 30 dernières années ne peut être observée. Au vu de l'ensemble

de ces données, la dominance des cercozoaires pendant la période médiévale reste donc assez largement inexplicée.

Les données acquises sur 2200 ans n'ont pas permis de quantifier la contribution relative des forçages anthropiques locaux et climatiques sur les communautés, notamment car la recherche de corrélations entre les différents indices issus des inventaires moléculaires (indices de diversité et structure, abondance relative de certains groupes) et les changements des conditions environnementales (état trophique du lac et fluctuations climatiques) n'a pas pu être effectuée. En effet nous ne disposons des estimations de teneurs en phosphore (indicateur de l'état trophique du lac) pour les archives sédimentaires du lac du Bourget que pour les 100 dernières années d'archivage et non au cours des 2100 années précédentes. Par ailleurs, la recherche de corrélations avec les fluctuations climatiques définies via la reconstruction des températures était périlleuse, les strates sédimentaires étudiées couvrant des pas de temps variables (allant jusqu'à 10 ans) d'où une difficulté à associer les températures à chaque strate sédimentaire étudiée. Ce problème ne s'est pas posé pour l'étude couvrant le siècle dernier, puisque nous disposons d'une résolution temporelle fine et ininterrompue à la fois pour les inventaires moléculaires et pour les 2 facteurs de forçage étudiés.

Sur la base de la modélisation statistique (GAM), cette étude a permis de confirmer le rôle accru du réchauffement récent dans la dynamique de la communauté eucaryote (communauté globale ou certains groupes comme les dinoflagellés et chrysophycées) à partir des années 80s, avec des effets détectés en particulier sur la structure (proportion relative des OTUs) de ces assemblages. Sur la base de ces données, les chrysophycées et dinoflagellés semblent particulièrement positivement impactés par les augmentations de température récentes avec un accroissement de leur richesse et de leur abondance relative. Toutefois l'effet du réchauffement récent est enregistré de manière concomitante avec l'effet de la ré-oligotrophication des lacs et l'estimation de leur part respective demeure délicate.

### **VI.3.5. Effets combinés des pressions anthropiques et climatiques sur les micro-eucaryotes: le cas de lacs en cours de ré-oligotrophisation exposés au réchauffement climatique**

La plupart des études concernant les effets combinés des modifications des teneurs en nutriments (en particulier le phosphore) et du réchauffement climatique ont été réalisées dans

des lacs eutrophes ou hyper-eutrophes. Ces travaux ont révélé la présence d'effets synergétiques entre les deux facteurs même si l'eutrophisation reste le facteur le plus fortement structurant pour la communauté phytoplanctonique (Jeppesen *et al.* 2005, Elliott & May 2008, Elliott 2012, Rigosi *et al.* 2014). Dans les lacs oligotrophes, l'importance relative des effets des deux facteurs sur les communautés est encore floue car certaines études montrent que le phosphore est le facteur structurant quand d'autres suggèrent que le réchauffement climatique a un effet plus fort sur les communautés (Blenckner *et al.* 2007, Shimoda *et al.* 2011, Rigosi *et al.* 2014). Dans les travaux du chapitre V, des effets combinés des deux forçages étudiés (diminution des teneurs en phosphore et augmentation des anomalies de température de l'air) ont été détectés pour des taxons photosynthétiques et hétérotrophes notamment au sein des chrysophycées, des ciliés, des dinoflagellés, des perkinsés et des champignons. La plupart des effets synergétiques ont été détectés au sein de la communauté biologique du lac ayant été fortement eutrophisé (lac du Bourget), de façon cohérente avec les observations de Rigosi *et al.* (2014) concernant les cyanobactéries. Nos analyses suggèrent, à l'instar de ces travaux sur les cyanobactéries, que l'histoire trophique d'un lac a un impact sur la réponse des communautés aux fluctuations climatiques.

Comme déjà mis en évidence à partir des archives sédimentaires pour les lacs d'Annecy et du Bourget par l'étude de différents bioindicateurs notamment les cladocères et les chironomes (Alric *et al.* 2013, Frossard *et al.* 2013, Perga *et al.* 2015), la ré-oligotrophisation de la masse d'eau n'a pas entraîné un retour de la communauté microbienne eucaryote à son état initial (en terme de diversité et structure). Cette tendance a également été observée pour la communauté biologique (phyto et zooplanctonique) de plusieurs autres lacs Européens (Jeppesen *et al.* 2007, Sondergaard *et al.* 2007). Pour la première fois, les travaux présentés dans le chapitre V ont permis de mettre en évidence cette tendance pour la communauté globale des micro-eucaryotes incluant donc à la fois des compartiments phytoplanctoniques et hétérotrophes. Ces trajectoires de « non-retours » sont vraisemblablement causées par les effets combinés de la ré-oligotrophisation et du réchauffement climatique sur les communautés depuis les années 1980s dans les deux lacs. La réponse des communautés est notamment marquée par l'émergence de plusieurs taxons (en termes de richesse et abondance relative) constituant des groupes de taxons important dans la structure des réseaux de co-occurrence (Figures **IV.7** et **V.4**). C'est notamment le cas pour les dinoflagellés, un groupe connu en milieu aquatique pour être sensible aux fluctuations climatiques (Vernal *et al.* 2001, Londeix *et al.* 2009, Chen *et al.* 2011) et aux modifications des teneurs en phosphore (Klouch *et al.* 2016). Ici non

seulement les deux facteurs semblent avoir des effets synergétiques sur la richesse et l'abondance de ces deux groupes, mais de plus, un module (groupes de taxons interconnectés) composé principalement d'OTUs de dinoflagellés a été détecté dans les deux lacs illustrant une réponse commune des communautés des deux lacs à ces forçages externes. Même si quelques OTUs de dinoflagellés ont pu être affiliées à des taxons connus pour être photosynthétiques (*Symbiodinium* sp.), la plupart d'entre elles ne sont pas affiliées au niveau du genre et donc les données produites ne permettent pas de conclure sur leur rôle fonctionnel, et donc de savoir si la réponse est spécifique à des taxons photosynthétiques ou non. Il serait intéressant de compléter les données existantes avec par exemple le séquençage massif de l'ADN sédimentaire des carottes sédimentaires des lacs d'Annecy et du Bourget en utilisant des amorces nucléotidiques ciblant les dinoflagellés (gènes codant l'ARNr 18S, *rbcl*...). Ces données permettraient peut être de révéler/caractériser plus précisément la diversité taxonomique et fonctionnelle au sein de ce groupe et les caractéristiques (morphologiques, physiologiques, métaboliques), expliquant sa réponse aux forçages.

Ces travaux de thèse se sont attachés à identifier les effets de deux facteurs, l'eutrophisation et le réchauffement climatique. Bien que ces deux facteurs soient connus pour avoir un fort impact sur les communautés biologiques lacustres, il existe cependant de nombreux autres facteurs pouvant avoir un effet sur les réseaux trophiques notamment la gestion des espèces piscicoles (induisant des effets en cascade sur les maillons trophiques inférieurs), d'autres pollutions d'origine anthropique et notamment des polluants nouvellement pris en compte tels que les micropolluants médicamenteux et les micro-plastiques (Chèvre 2014, Faure *et al.* 2015, Perga *et al.* 2015) ; et il ne peut pas être complètement exclu que ces facteurs n'ont pas affecté la réponse des micro-eucaryotes observée dans ces études. Toutefois, en l'état actuel des données environnementales disponibles, il n'était pas possible d'intégrer ces facteurs dans nos études, même si l'exploration des effets combinés de ce genre de facteurs pourrait fournir des informations très attendues concernant leur part respective dans la réponse des communautés lacustres.

## VII. Conclusion

Ces travaux de thèse ont permis d'apporter de nouvelles connaissances concernant la préservation de l'ADN dans les archives sédimentaires lacustres et à propos de questions d'écologie microbienne centrées sur les communautés eucaryotes. Les données obtenues ont mis en évidence :

- l'efficacité d'archivage de l'ADN planctonique dans les sédiments du lac du Bourget pour la plupart des groupes microbiens eucaryotes dont les chrysophycées, les chlorophytes, les cercozoaires, les ciliés et les dinoflagellés. L'ADN de certains groupes semble cependant moins bien archivé dans les sédiments et soulève des questions quant à la préservation différentielle de l'ADN de la communauté planctonique.

- les effets de la diagenèse au cours de premières années d'enfouissement dans les sédiments montrant que la quantité d'ADN et la richesse des communautés microbiennes eucaryotes ne semblent pas être impactées alors que des variations modérées sont observées dans la composition et la structure de la communauté au cours des 15 premières années d'archivage avec une stabilisation du signal dans les strates sédimentaires plus anciennes.

- les réarrangements au sein des communautés microbiennes eucaryotes en lien avec des forçages environnementaux externes plus particulièrement les fluctuations climatiques au cours des 2200 dernières années et les forçages environnementaux récents (eutrophisation de la masse d'eau des lacs, réchauffement climatique au cours des 30 dernières années). L'eutrophisation a notamment été identifiée comme le facteur le plus structurant pour la communauté microbienne eucaryote.

- l'intérêt de cibler certains groupes microbiens dont les caractéristiques (richesse, abondance) sont corrélées aux forçages environnementaux et qui pourraient constituer des indicateurs biologiques intéressants pour les études rétrospectives, il s'agit notamment des ciliés et chlorophytes en lien avec leur capacité à refléter l'état d'eutrophisation et potentiellement des dinoflagellés qui semblent être favorisés par le réchauffement climatique.

Ces travaux ont ainsi permis de démontrer que, malgré le besoin de calibration concernant la préservation de l'ADN au cours de l'enfouissement dans les sédiments, l'approche du séquençage massif de l'ADN sédimentaire a un fort potentiel pour les études d'écologie microbienne permettant d'avoir accès à des connaissances préalablement inaccessibles



concernant la dynamique temporelle de ces communautés et plus généralement la caractérisation de la diversité biologique de l'écosystème étudié.

Parmi les perspectives proposées dans la discussion générale, deux volets semblent prioritaires et pourraient apporter des réponses à court-moyen termes. Il s'agit d'une part de l'étude de la dynamique temporelle des communautés de micro-eucaryotes (lacustres ou autres écosystèmes) sur des pas de temps contrastés (journalier, mensuel, saisonnier) qui permettrait de comparer l'amplitude des remaniements s'opérant dans ces assemblages biologiques à divers pas de temps et de mieux hiérarchiser le poids des forçages journalier, saisonnier ou plus globaux (changement d'état trophique par exemple). D'autre part, le déploiement de l'approche de séquençage massif de l'ADN sédimentaire lacustre, dont nous avons démontré le potentiel dans ces travaux, sur plusieurs lacs ayant des histoires trophiques différentes, des typologies contrastées voire présents dans différentes écorégions, semble extrêmement prometteur. Ceci permettrait de tester la généralisation des observations faites ici pour quelques lacs (effet structurant majeur de l'état trophique, mobilisation des taxons rares lors de changements environnementaux drastiques, etc) et en particulier de valider l'intérêt de certains groupes/taxons en tant qu'indicateurs de perturbations anthropiques locales et climatiques (par exemple les chlorophytes et ciliés pour l'eutrophisation et les dinophycées pour les conditions de température).

# **VIII. Annexe : Application des méthodes moléculaires en paléolimnologie, de nouvelles opportunités pour étudier la dynamique à long terme de la biodiversité lacustre**

**Application of DNA-based methods in paleolimnology: new opportunities for investigating long-term dynamics of lacustrine biodiversity**

Papier en révision dans *Journal of Paleolimnology*

Isabelle Domaizon<sup>1</sup>, Amanda Winegardner<sup>2</sup>, Eric Capo<sup>1</sup>, Joanna Gauthier<sup>2</sup>, Irene Gregory-Eaves<sup>2</sup>

<sup>1</sup> CARRTEL, INRA, Université de Savoie Mont Blanc, 74200, Thonon-les-bains, France

<sup>2</sup> Department of Biology, McGill University, Montréal, Canada

## **VIII.1. Abstract**

The emergence of DNA analyses of lake sediments has opened up many new areas of inquiry, including the study of taxa that were traditionally not considered because they do not leave distinct morphological fossils. In this review, we discuss the potential and challenges associated with the study of DNA in paleolimnology to address critical research questions in lacustrine ecology. We examine some recent applications by highlighting studies that have quantified centennial to millennial-scale dynamics, and that considered a diversity of planktonic groups (including bacteria, diatoms, zooplankton). We also summarize the main methodological precautions to be taken into account for implementing these types of DNA analyses. Based on our read of the literature, the emerging topics we have identified include the application of sedimentary DNA to study: i) the spread, establishment and effect of invasive species, ii) the study of past fish population dynamics, iii) the interactions within lacustrine communities as identified through network analyses, iv) the potential application of metabarcoding for transfer functions. Overall, there are many new exciting questions that could be addressed using sedimentary DNA and this will no doubt be an area of continued expansion in the field of paleolimnology for many years to come.

## **VIII.2. Potential of DNA tools in paleolimnology**

Scientists unanimously recognize the widespread presence of multiple stressors on lake systems. An active area of aquatic science is quantifying the relative importance of the unique and synergistic impact that these stressors have on lake functioning and biodiversity. Long-term data are essential to address this issue (Smol 2010) and paleolimnology has a well-established reputation for providing valuable insights into the drivers of biological assemblages over long time scales (Gregory-Eaves & Beisner 2011). However, there is still room for considerable expansion of the field of paleolimnology, as many taxa that do not leave diagnostic morphological features in sediments, but could leave a DNA signature (Figure VIII.1). In addition, DNA based methods have the potential to drive clear advances in the study of taphonomy, intra-specific diversity and ecosystem functioning.

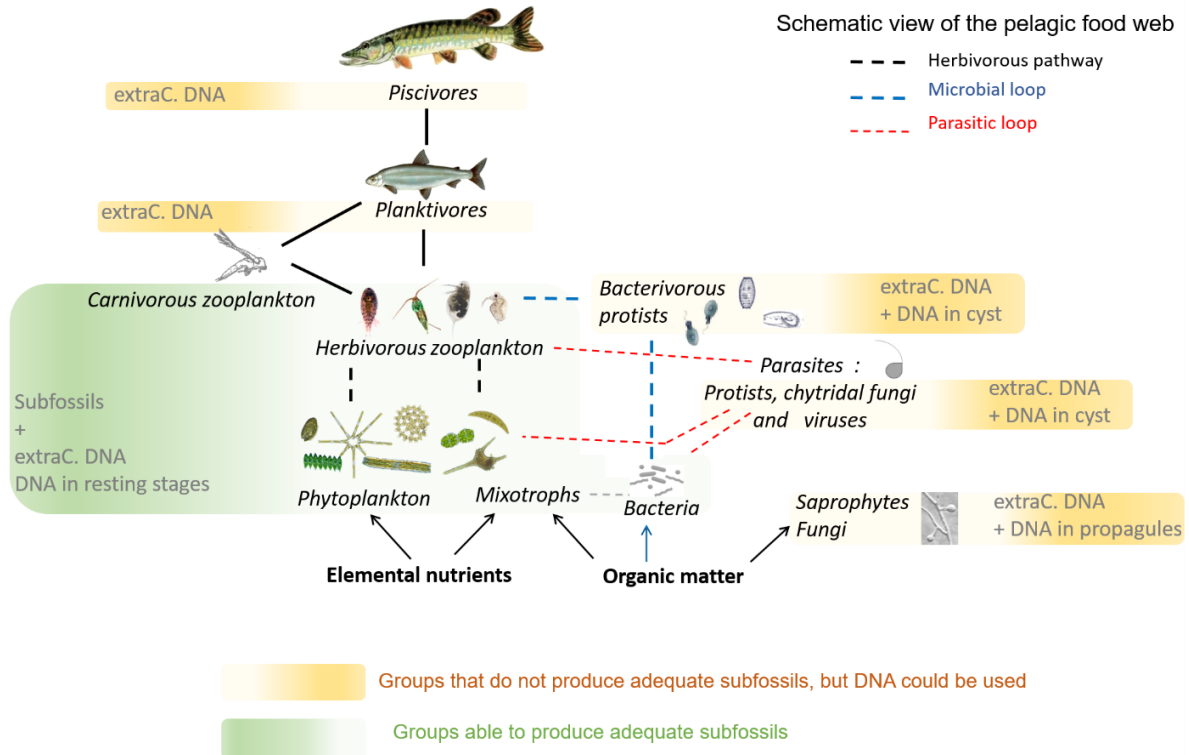


Figure VIII.1. Schematic view of the lacustrine pelagic food web highlighting how DNA analysis can expand the range of organisms in paleolimnology. The biological groups that typically don't leave a morphological subfossil, and consequently are not used as proxies in paleolimnology while their DNA can be preserved in sediment, are underlined by an orange frame. The type of DNA is specified as being intracellular DNA in cysts, propagules, resting stages, or extracellular DNA (extraC. DNA).

### VIII.2.1. DNA-based methods complement classical paleolimnology proxies and expand the range of organisms detectable in lake sediments.

An increasing number of publications show that aquatic sedimentary archives act as DNA repositories, preserving genetic information about organisms that live (or once lived) in the sediments or the water column above (Pawlowski *et al.* 2011, Coolen *et al.* 2013). The diversity of planktonic groups that are archived in the sediments over centuries to millennia and that can be detected from their DNA is non-trivial. For example, Capo *et al.* (2015) demonstrated that 71 % of the taxonomic units present (and detectable) in the water column over a ~2 year period were also identified by extracting and sequencing DNA from lake sediments cores that were ~100 years old. Numerous studies have also shown that aquatic organisms can be identified from DNA extracted from lake sediments that are up to 270 000 years (Randlett *et al.* 2014). The key advantages of DNA paleoecological approaches include: (i) an increase in the number of species that can potentially be studied at high taxonomic resolution, including those without diagnostic features retained in the sediment records (e.g. picocyanobacteria) (Domaizon *et al.* 2013, Hou *et*

*al.* 2014), (ii) the potential to explore evolutionary trajectories, by studying genetic variation among individuals over time (Epp *et al.* 2011) or by examining co-evolutionary patterns between taxa like host and parasite (Kyle *et al.* 2015a), (iii) the possibility to target specific genes that are indicative of key ecosystem functions like the DNA of methane oxidizing bacteria as a proxy for methanotrophy (Belle *et al.* 2014, 2015). With careful consideration of the methodologies needed to work with environmental DNA (aka eDNA; Poinar & Cooper 2000, Hofreiter *et al.* 2001, Rizzi *et al.* 2012), there are many exciting new insights that can be gained from the study of sedimentary DNA.

### **VIII.2.2. DNA-based methods are uniquely positioned to answer critical questions in paleolimnology, some of which are not possible with classic paleoecological indicators.**

In an attempt to define a forward vision for the field of paleolimnology, Seddon *et al.* (2014) identified 50 priority research questions in paleolimnology. Many of these questions would benefit from the integration of DNA-based methods and, in fact, DNA approaches may be the most realistic approach to drive real advancements in knowledge. For example, DNA approaches would be very useful for the question “How can the paleoecological record be applied to understand the interactions between native and invasive species?” (Seddon *et al.* 2014), because of their potential to track taxonomic groups at a finer resolution than morphological fossils. To date, a few paired comparisons of an organismal group identified using morphological remains and genetic markers have been conducted and, in some cases, considerably more taxonomic units have been identified based on DNA approaches (e.g. Stoof-Leichsenring *et al.* 2012).

On a more general level, the expansion of paleolimnology to include DNA-based approaches on a more routine basis has the potential to challenge existing limits within the field, as well as further promote the integration of paleolimnology with other domains. For example, Seddon *et al.* (2014) poses “What common questions can be addressed by ecologists and paleoecologists to bridge the contrasting spatial and temporal scales between the two disciplines effectively?” This is an important question as there is still sometimes disconnect between the fields of paleolimnology and other contemporary ecology fields (e.g. community ecology, population ecology etc.) (Rull 2014). This is not to say that the fields do not interact significantly, but rather that paleoecological data is not always used to its full potential in the other fields of aquatic ecology. The integration of DNA-based approaches into paleolimnology will continue to expand the field and bridge communication with molecular ecologists, enriching the questions addressed and approaches utilized by paleolimnologists. The opposite is also true – the paleolimnological

record offers microbial ecologists a unique temporal perspective to ask fundamental and applied questions.

Finally, an additional question highlighted by Seddon *et al.* (2014) was “How do paleoecologists encourage hypothesis testing rather than data-dredging approaches when exploring relationships between proxies and records?”. The introduction of molecular data into paleolimnology drastically increases the magnitude of data available. This likely means an initial increase in exploratory studies in paleolimnology because of the need to develop tools and machinery to handle this amount of data in the paleolimnology realm. However, the potential for hypothesis-driven research may also grow as new and different data are available for generating hypotheses over long temporal scales. We envision that DNA paleolimnological approaches will not replace the more classical physical, geochemical and morphological indicators that are widely used (see Smol 2008 for review), but rather the molecular methods will allow investigators to more fully identify mechanisms for the observed changes (e.g. methane-oxidizing bacteria supporting chironomid diets; Belle *et al.* 2014).

### **VIII.3. Some methodological aspects: precautions, limits and possible solutions**

Various issues should be addressed to ensure that the findings based on sedimentary DNA are robust. Some of these protocols are well documented, such as the need: i) to apply strict clean-lab procedures; ii) to include controls of non-contamination by recent DNA; iii) and to optimize protocols for ancient DNA that can be fragmented (Boere *et al.* 2011a, Rizzi *et al.* 2012, Torti *et al.* 2015). However, other issues are still not completely clarified. For instance, there is still a fair degree of uncertainty regarding the potential for differential preservation of DNA, which appears to depend on the physical and chemical conditions of the lake, as well as the target organisms.

### VIII.3.1. The way DNA is archived in sediments matters (Figure VIII.2)

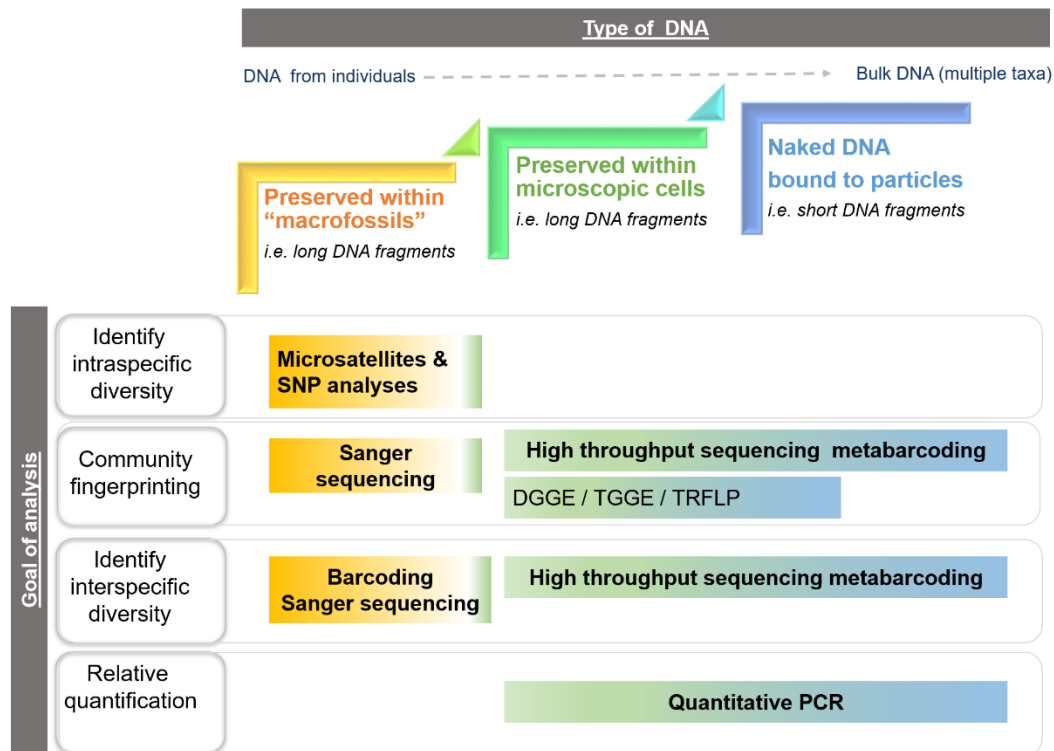


Figure VIII.2. Possible DNA analyses according to the goal of the study and the way DNA is archived. *SNP*: Single Nucleotide Polymorphism; *DGGE*: Denaturing Gradient Gel Electrophoresis; *T-RFLP*: Terminal Restriction Fragment Length Polymorphism; *Barcoding*: identification of species using a specific DNA region; *Metabarcoding*: identification of multiple taxa from a complex environmental sample using a specific DNA region.

DNA of lacustrine taxa may be archived in sediments in different forms depending on its biological origin. This variation, in turn, affects the level of protection and preservation of DNA, and consequently the analytical strategy needed to analyse the DNA (e.g. extraction methods, length of fragments targeted, DNA markers targeted, etc.). Many microscopic aquatic organisms produce some form of resting stage which affords them physical protection and allows them to withstand stressful conditions. Resting stages vary across organismal groups from the akinetes produced by some cyanobacteria, to the resting propagule formed by some protists (Harnstrom *et al.* 2011) and the ephippia produced by cladocerans (Piscia *et al.* 2015, Rogalski 2015, Morton *et al.* 2015). DNA may be extracted directly from these resting stages, or organisms can be resurrected and then the DNA can be extracted. In contrast, other aquatic organisms may have DNA that is less protected (found within easily lysed cell membrane) or may release their DNA directly into the environment (through the shedding tissues, releasing metabolic wastes, or via decomposition after death). Due to these contrasting origins, one can target primarily

intracellular DNA or the extracellular DNA fraction. Aquatic sediments are generally characterized by relatively high DNA concentrations in comparison to those found in the water column (Eichmiller *et al.* 2014, Turner *et al.* 2015), and most of this DNA is represented by extracellular DNA (Corinaldesi *et al.* 2005). Extracellular DNA is adsorbed onto mineral and organic particles and this limits its degradation by nucleases (Dell'Anno *et al.* 2002, Corinaldesi *et al.* 2005, 2008).

Despite the general dominance of extracellular DNA, it is generally hypothesized that taxa known to produce rigid cell walls (e.g. diatoms, dinophyceae, some chlorophyceae like Desmids) should be well represented in the sediment record due to the strong preservation of the DNA. In contrast, groups like Cryptophyta, whose DNA is protected solely by a simple membrane, have been identified as being underrepresented in recent sediments of a deep alpine lake compared to its water column (Capo *et al.* 2015). However, there is considerable variability in cell wall types even within taxa that might be considered to have “protected” DNA (i.e. organic-walled dinoflagellates tend to be sensitive to chemical conditions; Zonneveld *et al.* 1997), which in turn likely affect DNA preservation in the sediments. As such, we caution that categorizing organismal groups as having protected or unprotected DNA (i.e. cyst vs noncyst-forming species) is a rather simplistic vision.

DNA leaching across the temporal stratigraphy of sediments would significantly complicate the interpretation of results of paleoreconstruction based on the analysis of extracellular DNA. At present, DNA leaching is thought to be unlikely in lake sediments (Anderson-Carpenter *et al.* 2011), because of the significant temporal congruence that has been observed between paired DNA and morphological fossil records (e.g. Stoof- Leichsenring *et al.* 2012; Savichtcheva *et al.* 2014). However, we are unable to definitely rule out the risk of leaching based on our current state of knowledge and we suggest that this is a fruitful area for experimental testing (Haile *et al.* 2007).

### **VIII.3.2. How do lakes characteristics affect DNA preservation?**

There are now numerous studies where DNA has been successfully extracted from lake sediments and inference of past biological dynamics have been generated from these natural archives (Table VIII.1). Our list of study site examples vary from shallow polar lakes to deep temperate ecosystems. Furthermore, our review illustrates the variety of the biological groups targeted and the diversity of issues addressed. From this pool of literature, there has been some knowledge developed about the extent to which DNA is preserved and the conditions that



control the preservation of DNA in lacustrine sediments. Anoxic conditions, absence of bioturbation, and low temperature appear to be among the most crucial factors promoting DNA preservation (Boere *et al.* 2011a). Similarly, deep hardwater lakes were found to have very good preservation of cyanobacteria DNA, with long sequences archived and strong congruence observed between paleolimnological and historical records (i.e. Domaizon *et al.* 2013). In these hardwater lakes, the calcite precipitation process that occurs each year during spring and summer allows for the rapid sedimentation and burial of organic matter (such as planktonic taxa from the water column (Dittrich *et al.* 2004), which in turn, are processes that likely promote DNA preservation. Likewise, organic matter preservation is likely to be favorable in lakes with a concave morphology as these conditions tend to result in direct particle sedimentation and limit sediment remobilisation and resuspension. Moreover, colder, deeper habitats as well as a lack of bioturbating organisms that generate laminated sediments are likely to promote the preservation of DNA once buried in the sediments.

Table VIII.1. Our review of the literature focusing on the study of sedimentary DNA analyses in freshwater systems. Here we summarize meta-data for these studies: the sites, the biological targets, the main goals and the methods used.

CS: cloning sequencing qPCR: quantitative PCR

Targeted organisms	Climatic regions	Time Period (years)	Goal	Issue	Studied forcing factors	Molecular techniques	References		
Zooplankton	Copepods	65 000	Methodology	Sed-DNA amplification	-	CS (28S rRNA gene ; from eggs and sediments)	Bissett <i>et al.</i> 2005		
		surface sediments	Methodology	Sed-DNA amplification	-	CS (28S rRNA gene)	Xu <i>et al.</i> 2011		
		10 000	Ecology	Community dynamics	climate, salinity	454 pyrosequencing (28S rRNA gene; primers copepods)	Epp <i>et al.</i> 2015		
	temperate	10 000	Ecology	Community dynamics	climate, salinity	PCR-DGGE (partial copepod 18S rDNA 317 bp)	Coolen <i>et al.</i> 2013		
	Rotifers ( <i>Brachionus</i> )	arid	~ 200	Ecology	Population dynamics	volcanism	CS (COI gene)	Epp <i>et al.</i> 2010	
Daphnia	temperate	~ 50	Methodology	Sed-DNA amplification	-	12s rRNA; PCR	Ishida <i>et al.</i> 2012		
Microbial Eukaryotes	Microbial Eukaryotes	surface sediments	Methodology	Sed-DNA amplification	-	CS (18S rRNA gene)	Takishta <i>et al.</i> 2007		
		temperate	60	Methodology	Preservation of Sed-DNA	-	454 pyrosequencing (V4 18S; ~360 bp)	Capo <i>et al.</i> 2015	
		temperate, polar	~2 000	Ecology	Community dynamics	climate, nutrients	Illumina MiSeq sequencing (V7 18S ; ~280 bp)	Capo <i>et al.</i> in press	
	Microbial photosynthetic groups	phytoplankton	temperate	10 000	Ecology	Community dynamics	climate, salinity	454 sequencing (V1V3 region of 18S rRNA gene, 563 bp)	Coolen <i>et al.</i> 2013
			alpine	~ 3000	Ecology	Community & Population dynamics	climate, salinity, nutrients	PCR-DGGE (23S rDNA ~410 bp) + qPCR (~ 300 bp)	Hou <i>et al.</i> 2014
		Dinoflagellates	polar	~18 500	Ecology	Community dynamics	climate	PCR-DGGE (23S rDNA ~410 bp) + Sequencing	Li <i>et al.</i> 2016
			3 000	Ecology	Community dynamics	climate	PCR-DGGE (18S rRNA gene)	Boere <i>et al.</i> 2009	
		Diatoms	arid	surface sediment	Methodology	Sed-DNA amplification; DNA versus microscopic analysis	-	PCR-DHPLC (SSU gene)	Epp <i>et al.</i> 2011
			200	Ecology	Community & Population dynamics	climate	CS ( <i>rbcl</i> gene) ~650 - 750 bp	Stoof-Leichsenring <i>et al.</i> 2012	
		Diatoms ( <i>Staurisira</i> sp.)	polar	~11 500	Ecology	Community dynamics	climate, salinity	454 pyrosequencing (gene <i>rbcl</i> , ~100 bp)	Epp <i>et al.</i> 2015
			boreal	7 000	Ecology	Population dynamics	climate, geographic distance	CS ( <i>rbcl</i> gene; ~75 bp)	Stoof-Leichsenring <i>et al.</i> 2015
		Haptophytes, Diatoms	polar	~10 000	Ecology	Community & Population dynamics	salinity	PCR-DGGE (498 bp and 556 bp)	Coolen <i>et al.</i> 2004
		Haptophytes	temperate	270 000	Ecology	Community dynamics	climate	qPCR-DGGE (18S rRNA gene ; ~500 bp)	Randlett <i>et al.</i> 2014
		Chytrids (Fungi)	boreal	30	Ecology	Population dynamics; parasitic interaction	climate, nutrients	qPCR (ITS ; ~150 bp)	Kyle <i>et al.</i> 2015b
		Prokaryotes	Type I and II methanotrophic bacteria	polar	~ 10 000	Ecology	Population dynamics	salinity	qPCR-DGGE (specific primers on 16S rRNA gene)
Methanotrophic bacteria	temperate		~1 500	Ecology	Detection of anoxy, methan-associated processes	climate, nutrients	16S rRNA gene ; qPCR (~110 to 200 bp)	Belle <i>et al.</i> 2014 2015	
<i>Bifidobacterium</i> sp., <i>Bacteroides-Prevotella</i>	boreal		~ 3 000	Ecology	Detection of human activities	-	PCR (~300bp, 500 bp, 650bp) and nested PCR	Madeja <i>et al.</i> 2010	
Bacteria & Archaea	temperate		~ 11 000	Ecology	Community dynamics	climate	16S rRNA ; DGGE (V3 region ~250 bp)	Nelson <i>et al.</i> 2007	
All cyanobacteria (focus on <i>Synechococcus</i> sp.)	temperate		100	Ecology	Community dynamics; Population dynamics	climate, nutrients	16S rRNA ITS ; (up to ~1000 bp) & qPCR (up to ~300bp)	Domaizon <i>et al.</i> 2013	
All cyanobacteria			~200	Ecology	Community dynamics	climate, land use	16S rRNA; qPCR (will look up details of bp soon)	Pal <i>et al.</i> 2015	
Picocyanobacteria ( <i>Synechococcus</i> sp.)	alpine		~ 3 000	Ecology	Community dynamics; Population dynamics	climate, salinity, nutrients	CS (ITS ; 1100 bp) + qPCR (~288bp)	Hou <i>et al.</i> 2014	
All cyanobacteria (focus on <i>Planktothrix</i> )	temperate		100	Ecology	Community dynamics; Population dynamics	climate, nutrients	16S rRNA gene & ITS ; qPCR (up to 300 bp) & CS (up to ~1000 bp)	Savichtcheva <i>et al.</i> 2015	
<i>Planktothrix</i> sp. (cyanobacteria)			70	Methodology	Preservation of Sed-DNA	-	qPCR (16S rRNA ITS ; up to ~300bp)	Savichtcheva <i>et al.</i> 2011	
<i>Planktothrix</i> sp. (cyanobacteria)	boreal		300	Methodology	Sed-DNA amplification	-	qPCR ( <i>ociB</i> gene ; from 50 to ~380 bp)	Kyle <i>et al.</i> 2015a	
<i>Planktothrix</i> sp. (cyanobacteria)	boreal	30	Ecology	Population dynamics; parasitic relationships	-	qPCR ( <i>ociB</i> gene ; from 50 to ~380 bp)	Kyle <i>et al.</i> 2015b		
Others	Plants	temperate	5 000	Ecology	Detection of plant DNA	-	Spacer region between <i>atpB</i> & <i>rbcl</i> ; qPCR & CS (82 bp)	Anderson-Carpenter <i>et al.</i> 2011	
		10 000	Ecology	Community dynamics	climate, nutrients	Illumina MiSeq (chloroplast <i>trnL</i> gene ~ 65-155 bp)	Pansu <i>et al.</i> 2015		
	Animals, Plants	alpine	3 000	Ecology	Detection of human activities	-	Illumina MiSeq (mitochondrial 16S gene ~60-84 bp ; chloroplast <i>trnL</i> gene ~ 65-155 bp)	Giguet-Covex <i>et al.</i> 2014	
	Vascular plants, Bryophytes	polar	10 000	Ecology	Community dynamics	climate, salinity	454 pyrosequencing chloroplast <i>trnL</i> gene ( <i>bryo_P6E_1a</i> & <i>bryo_P6R</i> ; <i>trnL-g</i> & <i>trnL-h</i> ;	Epp <i>et al.</i> 2015	

Despite these advances in understanding, there has no systematic study that has quantified the degree to which DNA preservation is optimized in sediments (either in the field or in the lab), and thus this area is ripe for further development. A recent study of extracellular DNA from water does provide some interesting insight, showing evidence of slower decay rates at lower temperature and in waters with higher concentrations of dissolved organic carbon (Eichmiller *et al.* 2016). Additional insight could be gained from tapping into earlier studies that have examined how different environmental conditions affect the preservation of the more classically-studied morphological remains. We know, for example, that dinoflagellates cysts generally preserve less well under oxygen-rich conditions (Zonneveld *et al.* 1997), and so by extension one could predict that DNA from dinoflagellates would also be poorly archived under such conditions. Likewise, the silica frustules of diatoms are sensitive to high pH and thus under these conditions, there is strong potential alter the preservation of diatom DNA.

In addition, we recommend comparative studies of DNA from the water column, sediments traps and sediments cores to figure out taphonomic biases. Collections of sediments cores from the same site but sampled in different years (*sensu* Gälman *et al.* 2008) is another complementary method that should allow investigators to measure how the DNA signal is modified during burial.

### **VIII.3.3. What are the requisite precautions to ensure the authenticity of results?**

There are numerous precautions that should be followed at each stage of the analysis in a DNA paleolimnological project that optimize the authenticity of the results (Table VIII.2). A common risk at each analytical step is the potential contamination by ‘modern’ DNA (i.e. DNA coming from outside the sediments). Thomsen and Willerselev (2015) recently summarized the main pitfalls regarding eDNA studies, which are linked to contamination, inhibition, erroneous DNA sequences and reliability and completeness of reference DNA database; this manuscript, however, was not focused on paleolimnological studies and thus we provide some specific details here.

#### **Sampling and handling of samples**

The isolation of DNA from the core should be carried out in a specific DNA lab, physically separated from other molecular biology laboratory (especially due to the potential contamination by PCR products). Rigorous controls have to be employed against contamination, which includes (i) separate stations for sampling, DNA extractions, and PCR preparations, (ii) the use of sterile

disposable labware, (iii) the introduction of negative controls at all analytical steps including sampling, extraction and amplification, (iv) replicates (at least duplicates) for extractions and sequencing of each sediment sample. Several authors have previously reported and precisely described the recommendations for the analysis of ancient DNA (e.g. Gilbert *et al.* 2005, Rizzi *et al.* 2012).

#### DNA extraction and molecular analyses

The type of extraction (with or without cell lysis step) and the DNA region (called also barcode region) used to track the identity of past biological taxa might vary according to the targeted biological groups and the age of the sediments analysed (Table **VIII.2**). Very short barcodes (< 150 bp) are well adapted to degraded DNA (older sediments or extracellular DNA). However, it is not always possible to find very short barcodes that allow for a precise taxonomic assignment (i.e. identification to the species or intraspecific level). Often shorter barcodes are not robust enough to perform phylogenetic analyses. The choice of a barcode is therefore a compromise between the length (short for ancient/fragmented DNA) and the need for high taxonomic resolution and phylogenetic analyses. In a recent paper, Kyle *et al.* (2015b) elegantly demonstrated this trade-off with cyanobacteria primers as they showed that long barcodes (DNA fragments > 350 bp) could only be amplified in the younger sediments layers deposited (> 20 years) whereas short fragments (50 bp) were amplified in cores spanning 300 years.

The level of PCR inhibition due to co-extracted molecules (e.g. humic acid) also needs to be estimated in order to limit false negative results. As described by Savichtcheva *et al.* (2011), the level of potential inhibition can be verified using qPCR assays performed with a series of increasing amounts of genomic DNA of the biological target (isolated strain for instance), in the presence of increasing amounts of sediments DNAs (potentially containing inhibitors). A comparison of the efficiencies of qPCR performed with the resulting mixtures is used to estimate the level of inhibition; in this case, the approach applied is based on the assumption that inhibitors are diluted out when a log-linear relationship is achieved between C<sub>q</sub> and the dilution factor (Lloyd *et al.* 2010).

In certain situations, checking the physiological state (active versus inactive) of the biological group that is targeted may be important. For example, Capo *et al.* (2015) compared molecular inventories obtained from DNA and from RNA (i.e. transcripts of the gene, indicating that cells

are active) to differentiate the active and inactive biosphere. This step was done in order to verify that the presence of active micro-eukaryotes in the sediments did not induce any bias in the inventory of past eukaryotic diversity because active taxa may be over-represented. Such a verification step is even more important when studying bacteria of the past (particularly non-photosynthesizing bacteria when light is not present). Clearly, sediments represent an environment where many bacteria and Archaea can be active (even in relatively ancient sediments (Haglund *et al.* 2003); therefore, RNA must be targeted along with DNA to differentiate between past bacterial communities and those living in the sediments.

Once the targets are established, a decision must be made regarding the sequencing platform and bioinformatics approaches. High Throughput Sequencing (HTS) technologies applied for metabarcoding are rapidly evolving, where 454 pyrosequencing have largely been replaced by Illumina and Ion Torrent techniques that generate more DNA sequences at lower costs. Recent studies compared the main HTS platforms and showed that they perform equally well for HTS metabarcoding (Frey *et al.* 2014). It seems that the bioinformatics pre-processing, as well as the library preparation, have larger effects on the molecular diversity inventories generated (Esling *et al.* 2015, Pylro *et al.* 2014). A major issue is indeed eliminating / minimizing sequence errors (erroneous DNA sequences), which can have various origins. Numerous precautions should be applied all along the bioinformatics processing, which include applying stringent filtering and cleaning procedures to raw data and checking for the presence of chimera. In addition, one could run replicated analyses (from extractions to PCR), and then minimize the presence of falsely detected rare sequences/OTUs due to PCR bias (Lange *et al.* 2015) by only retaining DNA sequences/OTUs that are common to the replicates. Across the bioinformatics process, there are trade-offs between error trimming and keeping as much information as possible, and clearly it depends on a variety of factors including the gene target, molecular tools used, level of taxonomic resolution sought, and interest in phylogenetic analyses.

Finally, the completeness of the reference of DNA sequence databases is also crucial to provide relevant taxonomic list through metabarcoding. As has been illustrated in numerous publications, there is still a lack of references for numerous planktonic taxa found in environmental molecular inventories and a strong need to continue to develop barcode reference libraries, especially for microbes (del Campo *et al.* 2014). For larger organisms, research programs such as the International Barcode of Life (IBOL; <http://ibol.org/>) have facilitated the standardization and reusability of barcodes.

Table VIII.2. Main analytical steps, risks and possible solutions for implementation of DNA analyses in paleolimnological studies.

MAIN ANALYTIC STEPS		CHOICE AND PRECAUTIONS
<b>1 – Sampling and Handling of samples</b>	Field coring	<ul style="list-style-type: none"> <li>Choice of coring methods (gravity corer, piston corer, freeze corer)</li> <li>Adapt handling to avoid external contamination</li> </ul>
	Handling and preservation of sediment cores	<ul style="list-style-type: none"> <li>'Immediate' subsampling with specific room and lab disposable, and rapid storage of subsamples (from 4 °C to 80 °C or via lyophilisation according to the duration of storage)</li> <li>Delayed subsampling (preservation at 4 °C, darkness, limitation of water circulation within the core and oxydation)</li> </ul>
	Sub-sampling of sediment cores	<ul style="list-style-type: none"> <li>Select the center of the core to reduce contamination &amp; oxydation risks</li> <li>Adapt aliquots volume/size according to the amount of material required for extraction</li> <li>Include blank/controls for aerial contamination</li> <li>Specific work stations isolated from DNA labs</li> </ul>
<b>2 – DNA extraction</b>	Choice of DNA extraction methods	<ul style="list-style-type: none"> <li>DNA from cysts/protected cells : DNA extraction with specific lysis step</li> <li>Total DNA : Soil/sediment DNA commercial kits, « in house » protocols</li> <li>Extracellular DNA : desorption and precipitation of 'free' eDNA</li> </ul>
	Limit & estimate the contamination by modern DNA	<ul style="list-style-type: none"> <li>Include blanks/controls to track potential contamination</li> <li>Use specific work stations dedicated to rare/ancient DNA</li> </ul>
<b>3 Estimation of PCR inhibitors (optional)</b>	qPCR assays	<ul style="list-style-type: none"> <li>Estimation of inhibitors presence via qPCR assays</li> <li>Adaptation of protocols to remove potential PCR inhibitors (additional purification steps with silica column, ethanol precipitation...)</li> </ul>
<b>4 – PCR amplification from sedimentary DNA</b>	Choice of the molecular probes (primers)	<ul style="list-style-type: none"> <li>Verify specificity by sequencing amplicons</li> </ul>
	Trade-off between Taxonomic resolution Barcodes length	<ul style="list-style-type: none"> <li>Adapt genes/DNA regions targeted to the required taxonomic resolution</li> <li>Relevance and completeness of molecular reference database used for taxonomic assignment</li> <li>Select short barcodes (&lt; 150 bp) or medium size barcodes (150 to 600 bp) : choice according to DNA fragmentation and available barcodes</li> </ul>
<b>5 – Preparation of amplicons and HTS sequencing</b>	Choice of the HTS technology	<ul style="list-style-type: none"> <li>Compromise between sequencing depth and number of samples</li> <li>Optimization according to barcode length (eg. short barcodes well adapted to Illumina techniques, long barcodes treated by 454 pyrosequencing)</li> </ul>
	Choice of the number of replicates	<ul style="list-style-type: none"> <li>Include replicates in DNA extraction and PCR amplification steps depending on the goal of the study (ecology or methodology)</li> </ul>
	Choice of tags for each sample	<ul style="list-style-type: none"> <li>Estimating the saturation level</li> <li>Estimate mistagging, Include alien DNA in HTS samples</li> </ul>
	Optimisation of PCR protocols & quality controls	<ul style="list-style-type: none"> <li>Optimisation of PCR conditions, work in specific rooms dedicated to rare DNA, manipulation of a limited number of samples and introduction of blanks to detect potential cross contamination</li> </ul>
<b>6 – Post-sequencing treatments</b>	Bioinformatics processing	<ul style="list-style-type: none"> <li>Stringence of cleaning procedures e.g., <i>merged paired-end reads with large overlap and no mismatch allowed</i></li> <li>Limit false positive reads e.g., <i>Keep only reads/OTUs found in at least two replicates</i></li> <li>Take into account DNA polymorphism due to diagenetic processes e.g., <i>track preferential mutation sites</i></li> </ul>

## **VIII.4. Applications of sedimentary DNA analyses to advance lacustrine ecology**

### **VIII.4.1 Both intra specific diversity and complex biological group can be described from sedimentary DNA**

To date, the sequencing of sedimentary DNA and quantitative PCR have been applied to a large spectrum of organisms, ranging from microbial eukaryotes (including phytoplanktonic taxa) and cyanobacteria (Coolen *et al.* 2004, Stoof-Leichsenring *et al.* 2012, Domaizon *et al.* 2013, Savichtcheva *et al.* 2014, Pal *et al.* 2015) to zooplanktonic groups (Copepods (Bisset *et al.* 2005); rotifers (Epp *et al.* 2009); Cladocerans (Ishida *et al.* 2012) and fish (Stager *et al.* 2015, Turner *et al.* 2015). The methodological approaches applied across these different studies also vary as in some cases DNA is extracted from individual resting stages (e.g. studies of the DNA from cladoceran ephippia; Bisset *et al.* 2005) whereas in other cases DNA is extracted from bulk sediment samples to target a wide range of phylogenetic groups (e.g. studies of 100s of eukaryotic groups; Coolen *et al.* 2013). Likewise, studies vary in their focus from interspecific diversity to intraspecific diversity. There have been numerous intra-specific studies of cladocerans as this field has been active for some time (e.g. Hairston *et al.* 1999, Pollard *et al.* 2003), but other organismal groups are now also been targeted, such as diatoms (Stoof-Leichsenring *et al.* 2012) and cyanobacteria (Domaizon *et al.* 2013). Overall, the knowledge gained from this work is very diverse in scope, ranging from quantifying natural variability in population and community dynamics to understanding how these variables respond to anthropogenic disturbances locally (at the watershed scale) or more globally (e.g. climatic change).

### **VIII.4.2. DNA from zooplankton resting stages**

One of the most established applications of sedimentary DNA comes from the study of zooplankton diapause resting stages. Zooplankton diapause resting stages come in many forms; bare diapause eggs in the case of copepods, diapause eggs encased in a protective ephippia in the case of cladocerans, and dessicated individuals or bare eggs in the case of the rotifers (Hairston & Fox 2009). These stages are produced cyclically, as well as in times of intense competition or stress (Cáceres 1998). Many lacustrine systems rely on the resting egg bank both as a source of new individuals each growing season, as well as a source of genetic diversity.

Kerfoot *et al.* (1999) and Hairston *et al.* (1999) were among the first studies to characterize cladoceran egg assemblages through resurrection ecology, where the organisms are revived from resting eggs and their DNA extracted to identify their population identity. More recently DNA-based approaches for studying resting eggs in the paleoecological record have expanded such that the DNA is being extracted directly from the resting stage, thus avoiding the somewhat tricky work of resurrecting the organism. For example, Mergeay *et al.* (2007) extracted DNA from *Daphnia* eggs to identify changes in population dynamics following a large water level change in Lake Navaisha. In addition, genetic analyses can be useful to identify the species present within the broader cladoceran assemblage because there are only a small number of references for the taxonomy of cladoceran ephippia (Vandekerkhove *et al.* 2004). The combined use of subfossil morphological identifications and sequenced 12S rRNA from *Daphnia* ephippial cases was shown to be quite powerful in a study of *Daphnia* dynamics in an alpine lake (Hamrová *et al.* 2010). In this case, the authors found that the presence of a long tail spine on ephippia is not a species-specific trait, and thus the genetic analysis provided greater taxonomic resolution. In particular, they found that the ephippia in question were from the taxa *D. longispina* and not *D. lacustris* as previously thought.

New methods are also expanding horizons in terms of the sample and organism type that can be targeted. For example, Ishida *et al.* (2012) has described a method for the extraction and amplification of DNA from just the empty cladoceran ephippial cases, in addition to those with viable diapause eggs from sediments cores. However, the success of these methods in subsequent application has been quite variable (Briski *et al.* 2011; Ishida, Pers. Comm.). If analytical methods can be improved such that investigators can reliably extract and amplify DNA from just the ephippial cases, then we have the potential to evaluate whether there are any genetic differences between individuals that hatched (and thus just leave a case in the sediment records) and those that did not hatch (and thus leave an intact ephippia in the sediment records). Recent progress has also been shown in sequencing copepod eggs from sediments (Xu *et al.* 2011a; 2011b, Makino *et al.* 2013), which is particularly exciting as the carapaces of these organisms do not preserve well in lake sediments and the eggs are very difficult to identify morphologically.



### **VIII.4.3. Targeting specific functions in the ecosystem**

Sedimentary DNA studies also offer the opportunity to target specific genes and genotypes, which are linked to specific functions in an ecosystem. For example, Belle et al. (2014, 2015) used ancient DNA from the bacterial methanotroph community to provide information about past methane oxidation in two freshwater lakes. By pairing the analysis of methane-oxidizing bacteria DNA with carbon isotopes analyses of both chironomid head capsules and bulk sedimentary organic matter from the same sediment intervals, Belle and colleagues (2014, 2015) were able to show very convincingly that these bacteria were an important source of carbon for the chironomids and that the contribution of methanogenesis pathway to benthic production varies considerably over time.

Similarly, the dynamics of toxic cyanobacteria (e.g. *Planktothrix*) have been examined by studying a portion of the gene encoding for microcystins (*McyA*). For example, Savichtcheva et al. (2011) used the *McyA* gene as well as a more general cyanobacterial gene to quantify total and microcystin-producing *Planktothrix* populations from sediment cores spanning the past 70 years in three lakes. Each of these study lakes have long-term water column monitoring data (spanning > 20 years), and thus comparisons were drawn between the coherence of the sedimentary DNA and contemporaneous water column data. Overall, Savichtcheva et al. (2011, 2014) demonstrated strong congruence among the time series from any one study lake and the lake with the greatest amount of *Planktothrix* in the water column also had the greatest amount of *McyA* copies in the sediments. Further applications of this methodology to a wide range of sites will advance our understanding of the spatio-temporal dynamics of toxic and non-toxic cyanobacteria. When combined with other paleolimnological indicators (both genetic and classical tools), we can also gain greater insight into the impact of cyanobacteria on lake ecology.

## **VIII.5. Emerging topics & challenges in applying sedimentary DNA to the field of lacustrine ecology**

### **VIII.5.1. DNA approaches to track the dynamics of invasive species**

Given the high interest in invasive aquatic species, this is a ripe area for paleolimnological study. To date, there are a few examples of where classical paleolimnological methods have been

applied to address questions related to invasive species (e.g. the study of *Bythotrephes* invasion in a Canadian lake; Hall & Yan 1997). However, the application of sedimentary DNA methods offers an opportunity to study enlarged list of invasive taxa that could be targeted, as well as the possibility to study the consequences of the invasion. Molecular techniques have indeed been successfully employed for the detection of a growing number of invasive organisms comprising vertebrates, invertebrates, plants and pathogens (Jerde *et al.* 2011, Mahon *et al.* 2013, Goldberg *et al.* 2015) from the analysis of water samples; consequently some public agencies are now implementing eDNA analyses as part of their routine surveillance of invasive species (e.g. asian carp survey in USA and Canada). Paleolimnological studies have the potential to further advance our understanding of invasion dynamics as it can help answer the following questions: when was the species introduced, and which taxa were concurrent with this exotic species. Already, genetic analysis of sediment egg banks have proven to be a key source of information regarding the successful invasive *Daphnia* populations (e.g. Möst *et al.* 2015). Ultimately, we think that performing genetic and classical paleolimnological studies across many sites in a landscape is a promising approach to investigate biogeographical patterns of invasive taxa through time.

#### **VIII.5.2. DNA approaches to reveal the presence of fish populations over time**

Fishes are often of key interest to society because they provide numerous ecosystem services and are key components of the lacustrine food webs. As such, it would be ideal to be able to quantify past changes in fish populations or communities from the analysis of lake sediments, but in many cases this has been challenging based on classical methods alone. For one, the morphological elements of fish (i.e. scales and bones) are not frequently preserved (but see Davidson *et al.* 2003 as an exception). The temporal dynamics of fish and their role as predators, however, have been studied indirectly by measuring the size of key prey (e.g. *Daphnia*; Jeppesen *et al.* 2002) and by considering changes in zooplankton community structure (e.g. Palm *et al.* 2005, Palm & Svensson 2010). In some instances, migratory fishes may have distinct geochemical signatures and return in enough large abundances that the fish themselves substantially change a lake's nutrient budget. The reconstruction of past Sockeye Salmon population abundances from analyses of stable nitrogen isotopes has been shown to accurately track past historical records, and have been complemented by other classical indicators such as diatoms and cladocerans (Finney *et al.* 2000; reviewed by Gregory-Eaves *et al.* 2009). The possibility of expanding this area of study has been recently demonstrated by the analysis of environmental DNA in both water (Takahara *et al.* 2012,, Lacoursière-Roussel *et al.* 2016) and sediments (Eichmiller *et al.*

2014, Smith *et al.* 2015, Stager *et al.* 2015, Turner *et al.* 2015). For example, Eichmiller *et al.* (2015) were able to track the presence and composition of fish assemblages from environmental samples. Furthermore, Eichmiller *et al.* (2014) reported that the concentrations of fish eDNA in sediment samples were several orders of magnitude greater than water on a per mass basis. Nonetheless, Eichmiller *et al.* (2014) reported low detection rates for fish.

### **VIII.5.3. Inferring ecological networks from molecular inventories obtained for complex assemblages**

High Throughput Sequencing (HTS) approaches are becoming more and more popular tools to estimate plankton diversity in both marine (Massana *et al.* 2015, de Vargas *et al.* 2015) and lacustrine samples (Nolte *et al.* 2010, Debroas *et al.* 2015, Kammerlander *et al.* 2015, Filker *et al.* 2016). In lacustrine plankton, numerous authors have reported a previously undescribed diversity in both eukaryotic and prokaryotic communities (e.g. Eiler *et al.* 2012, Mangot *et al.* 2013, Oikonomou *et al.* 2015). To date, most of the HTS inventories performed on lakes have focused on water samples (rarely on sediments) and this body of work has shown a high amount of eukaryotic diversity, including poorly studied biological groups, such as Apicomplexa, Chytridiomycota, and Cryptomycota. The taxa detected in these biodiversity inventories can then be assigned to different functional traits (e.g. phototrophs, parasites, phagotrophs, etc). This trait-based approach facilitates the inference of ecological dynamics from molecular inventories of the thousands of taxa/OTU detected (e.g. de Vargas *et al.* 2015), and may be further explored through network analyses. Recent analyses of HTS data from water column samples have provided evidence of the kinds of insights that can be gained through association networks, which can be indicative of food web structure (Chaffron *et al.* 2010; Chow *et al.* 2014; Kammerlander *et al.* 2015; Lima-Mendez *et al.* 2015). In particular, network analyses reveal co-occurrence patterns between taxa and environmental factors, identify taxa that have a central role in the network (important connectors of the network) as well as allow for the description of emergent properties of the network such as the interconnectedness and average path length (Kara *et al.* 2013). To our knowledge, this type of analyses have not yet been applied to sedimentary DNA records but methods are certainly available to deal with time series data (recently review by Faust *et al.* 2015). Overall, the application of statistical analysis (such as alpha, beta and network analysis) to molecular inventories generated from sedimentary DNA records offers a unique opportunity to study community processes over very long time scales (decades to

centuries), which is crucial for understanding the impact of multiple stressors during the Anthropocene.

#### **VIII.5.4. Coupling spatial and temporal approach: temporal biogeography**

Paleobiogeography is another important area of research, and the contribution of phylogenetic studies to the field of paleobiogeography has been significant over the last two decades (Lieberman 2003, Neubauer *et al.* 2016). However, temporal considerations in biogeographic analyses have been generally restricted to macro-organisms. As such, the description of planktonic biodiversity patterns has generally been performed without temporal framework. The emergence of DNA-based methods to study the global lacustrine diversity from dated sediments could help to make key advances in plankton biogeography (e.g. Stoof-Leichsenring *et al.* 2015 for diatoms lineages).

#### **VIII.5.5. The use of molecular data for paleoecological transfer functions**

Transfer functions are an important tool in paleolimnology (Smol 2008, Smol & Stoermer 2010). In a transfer function, contemporary relationships between organisms preserved in surface sediments and environmental data ('calibration sets') are developed and applied to sediment core records to infer past environmental conditions based on the organisms present (Smol 2008). In paleolimnology, the most common organisms used for these models are the diatoms because of their adequate preservation in lake sediments, as well as their clearly delineated environmental tolerances for many environmental variables. Returning to Seddon *et al.* (2014), an additional question highlighted was "When using modern analogues, what measures can be taken to be sure that the training set is sufficient to reconstruct the full range of likely past conditions, and if not, what else should be used to supplement these methods?". We suggest that the DNA-based approaches, in addition to traditional paleoecological indicators, have the potential to address this question as one could evaluate if both the genetic and morphological diversity in the core is well represented in the calibration set. In addition, DNA based approaches could allow investigators to consider a broader lacustrine community than the traditional approach of developing calibration sets which focus on a single indicator at a time. While this research direction is not yet being actively pursued in the literature (to our knowledge), this is another fruitful area for research.

### **VIII.5.6. Integration of sedimentary DNA in biological collections and long-term environmental observatory: sediment core repository as a conservation tool for biodiversity**

Because sediment cores have long been recognized as a fundamental source of information, there has been a movement towards storing sediment cores in repositories and making them available to the scientific community. These repositories are incredibly valuable as often questions are thought of long after the sediments are collected and can be addressed by re-sampling of archived cores. However, many of the existing sediment core repositories do not explicitly routinely store sediments under ideal conditions for DNA analysis (often storage conditions are a 4°C fridge, where sediments are held in moisture-sealed containers). Certainly, sedimentary DNA analyses could benefit past coring expeditions like initiatives led by the International Continental Scientific Drilling Program (ICDP), but would require better storage conditions (e.g. -20 °C).

## **VIII.6. Conclusion**

Although early sedimentary DNA approaches were initiated in the late 1990s, this field is experiencing rapid growth as new tools are coming online. This growth is facilitating the study of a generally wider range of taxa and in a more efficient manner. As with any paleolimnological proxy, careful studies need to be conducted to evaluate preservation biases and appropriate lab methods need to be followed. However, the handful of calibration studies that have been performed to date where morphological or geochemical signatures have been compared with the results of sedimentary DNA analyses are very encouraging (e.g. Boere *et al.* 2011, Stoof-Leichsenring *et al.* 2012, Pal *et al.* 2015). There are many questions ripe for application of sedimentary DNA including the study of invasive species, fish population dynamics and interactions within lacustrine communities.

# Glossaire

**Autotrophie** : capacité d'un organisme à produire des composés organiques en utilisant comme source de carbone du CO<sub>2</sub>.

**Cils** : extension du cytoplasme permettant notamment le déplacement des organismes.

**Commensalisme** : capacité d'un organisme à vivre en interaction avec d'autres espèces en utilisant les ressources alimentaires qu'elles produisent, mais sans leur nuire.

**Cytospharynx** : invagination de la membrane chez les protozoaires pour l'alimentation.

**Endosymbiose** : théorie selon laquelle les chloroplastes sont issus d'une association entre un eucaryote ancestral et une cyanobactérie (endosymbiose primaire). Les endosymbioses secondaire et tertiaire correspondent à l'association d'eucaryotes avec des organismes issus d'endosymbioses primaire ou secondaire.

**Hétérotrophie** : capacité d'un organisme à utiliser des composés organiques comme source de carbone.

**Hydrogénosomes** : organelle de cellules eucaryotes qui dérivent probablement des mitochondries.

**Mastignomènes** : fines digitations de la membrane d'un flagelle.

**Mitosomes** : organelle de cellules eucaryotes qui dérivent probablement des mitochondries.

**Mixotrophie** : capacité d'un organisme à utiliser des sources de nutriments autotrophes et hétérotrophes.

**Mutualisme** : interaction bénéfique entre deux ou plusieurs organismes.

**Osmotrophie** : capacité d'ingérer des éléments nutritifs dissous par diffusion à travers la membrane pour l'alimentation.

**Phototrophie** : capacité d'un organisme à tirer de son énergie à partir des rayons lumineux.

**Parasitisme** : interaction avec un effet positif pour un organisme (parasite) et négatif pour l'autre (hôte) par utilisation de sa masse et de son énergie.

**Phagotrophie** : capacité d'un organisme à ingérer de la matière organique particulière pour l'utiliser pour l'alimentation.

**Prédation** : interaction au cours de laquelle le prédateur capture et consomme sa proie.

**Pseudopode** : extension de la membrane plasmique utilisée par les cellules pour le déplacement et l'alimentation.

**Saprotrophie** : capacité d'un organisme à se nourrir de matière organique.

**Synapomorphie** : caractère dérivé partagé par deux ou plusieurs unités taxonomiques. Les taxons, définis par un caractère synapomorphique, forment un groupe monophylétique (un ancêtre et tous ses descendants).

**Symbioses** : interaction durable bénéfique entre deux ou plusieurs espèces.

**Undulipodium** : extension extracellulaire composée de microtubules et impliquée dans le mouvement des cellules eucaryotes. Synonyme de flagelles et cils.



# Références

- Abram NJ, McGregor H V., Tierney JE *et al.* (2016) Early onset of industrial-era warming across the oceans and continents. *Nature*, **536**, 411–418.
- Adl SM, Habura A, Eglit Y (2014) Amplification primers of SSU rDNA for soil protists. *Soil Biology and Biochemistry*, **69**, 328–342.
- Adrian R, O'Reilly CM, Zagarese H *et al.* (2009) Lakes as sentinels of climate change. *Limnology and Oceanography*, **54**, 2283–2297.
- Allison SD, Martiny JBH (2008) Resistance, resilience, and redundancy in microbial communities. *Proceedings of the National Academy of Sciences*, **105**, 11512–11519.
- Alric B, Jenny J-P, Berthon V *et al.* (2013) Local forcings affect lake zooplankton vulnerability and response to climate warming. *Ecology*, **94**, 2767–2780.
- Alric B, Perga M-E (2011) Effects of production, sedimentation and taphonomic processes on the composition and size structure of sedimenting cladoceran remains in a large deep subalpine lake: paleo-ecological implications. *Hydrobiologia*, **676**, 101–116.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *Journal of Molecular Biology*, **215**, 403–10.
- 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.
- Amblard C, Boisson JC, Bourdier G *et al.* (1998) Écologie Microbienne En Milieu Aquatique : Des Virus Aux Protozoaires. *Revue des sciences de l'eau*, **11**, 145.
- Anderson-Carpenter LL, McLachlan JS, Jackson ST *et al.* (2011) Ancient DNA from lake sediments: bridging the gap between paleoecology and genetics. *BMC evolutionary biology*, **11**, 30.
- Anneville O, Gammeter S, Straile D (2005) Phosphorus decrease and climate variability: mediators of synchrony in phytoplankton changes among European peri-alpine lakes. *Freshwater Biology*, **50**, 1731–1746.
- Anneville O, Souissi S, Gammeter S, Straile D (2004) Seasonal and inter-annual scales of variability in phytoplankton assemblages: comparison of phytoplankton dynamics in three peri-alpine lakes over a period of 28 years. *Freshwater Biology*, **49**, 98–115.
- Arnaud F, Révillon S, Debret M *et al.* (2012) Lake Bourget regional erosion patterns reconstruction reveals Holocene NW European Alps soil evolution and paleohydrology. *Quaternary Science Reviews*, **51**, 81–92.
- Arneborg J, Lynnerup N, Heinemeier J *et al.* (2011) Norse Greenland Dietary Economy ca. AD 980-ca. AD 1450: Introduction. *Journal of the North Atlantic*, **3**, 1–39.



- Auer I, Böhm R, Jurkovic A *et al.* (2007) HISTALP—historical instrumental climatological surface time series of the Greater Alpine Region. *International Journal of Climatology*, **27**, 17–46.
- Austrheim G, Asheim LJ, Bjarnason G *et al.* (2008) Sheep Grazing in the North-Atlantic Region: A Long Term Perspective on Management, Resource Economy and Ecology. *Vitenskapsmuseet, Rapport Zoology Series*, **3**, 1–82.
- Aylward FO, Eppley JM, Smith JM, Chavez FP, Scholin CA (2015) Microbial community transcriptional networks are conserved in three domains at ocean basin scales. *Proceedings of the National Academy of Sciences*, **112**, 5443–5448.
- Azam F, Graf JS (1983) The Ecological Role of Water-Column Microbes in the Sea. *Marine Ecology Progress Series*, **10**, 257–263.
- Baltar F, Palovaara J, Vila-costa M *et al.* (2015) Response of rare, common and abundant bacterioplankton to anthropogenic perturbations in a Mediterranean coastal site. *FEMS Microbiology Ecology*, **91**, 1–12.
- Barberán A, Bates ST, Casamayor EO, Fierer N (2012) Using network analysis to explore co-occurrence patterns in soil microbial communities. *The ISME Journal*, **6**, 343–51.
- Barberán A, Henley J, Fierer N, Casamayor EO (2014) Structure, inter-annual recurrence, and global-scale connectivity of airborne microbial communities. *The Science of the Total Environment*, **487C**, 187–195.
- Barnosky AD, Koch PL, Feranec RS, Wing SL, Shabel AB (2004) Assessing the causes of late Pleistocene extinctions on the continents. *Science*, **306**, 70–5.
- Bastian M, Heymann S (2009) Gephi: An Open Source Software for Exploring and Manipulating Networks.
- Bates ST, Clemente JC, Flores GE *et al.* (2013) Global biogeography of highly diverse protistan communities in soil. *The ISME Journal*, **7**, 652–659.
- Battarbee RW (2000) Palaeolimnological approaches to climate change, with special regard to the biological record. *Quaternary Science Reviews*, **19**, 107–124.
- Battarbee R, Anderson J, Jeppesen E, Leavitt PR (2005) Combining palaeolimnological and limnological approaches in assessing lake ecosystem response to nutrient reduction. *Freshwater Biology*, **50**, 1772–1780.
- Battarbee RW, Carvalho L, Jones VJ *et al.* (2001) Diatoms. In: *Tracking Environmental Change Using Lake Sediments Volume 3: Terrestrial, Algal, and Siliceous Indicators* (eds Smol JP, Last WM, Birks HJB). Kluwer Academic Publishers, Dordrecht.
- Battarbee R, Charles D, Dixit S, Renberg I (1999) Diatoms as indicators of surface water acidity. In: *The Diatoms: Applications for the Environmental and Earth Sciences*, pp. 85–127. Cambridge University Press.

- Battarbee RW, Grytnes J, Thompson R *et al.* (2002) Comparing palaeolimnological and instrumental evidence of climate change for remote mountain lakes over the last 200 years. *Journal of Paleolimnology*, **28**, 161–179.
- Battarbee RW, Morley D, Bennion H *et al.* (2010) A palaeolimnological meta-database for assessing the ecological status of lakes. *Journal of Paleolimnology*, **45**, 405–414.
- Battarbee, RW, Anderson, NJ, Bennion, H, & Simpson G (2012) Combining limnological and palaeolimnological data to disentangle the effects of nutrient pollution and climate change on lake ecosystems: problems and potential. *Freshwater Biology*, **57**, 2091–2106.
- Beaver JR, Crisman TL (1982) The trophic response of ciliated protozoans in freshwater lakes. *Limnology and Oceanography*, **27**, 246–253.
- Beaver JR, Crisman TL (1989) Analysis of the community structure of planktonic ciliated protozoa relative to trophic state in Florida lakes. *Hydrobiologia*, **177**, 177–184.
- Behrensmeyer AK, Kidwell SM (1985) Paleontological Society Taphonomy ' s Contributions to Paleobiology. *Paleobiology*, **11**, 105–119.
- Behrensmeyer AK, Kidwell SM, Gastaldo RA (2000) Taphonomy and paleobiology. *Paleobiology*.
- Belle S, Parent C, Frossard V *et al.* (2014) Temporal changes in the contribution of methane-oxidizing bacteria to the biomass of chironomid larvae determined using stable carbon isotopes and ancient DNA. *Journal of Paleolimnology*, **52**, 215–228.
- Belle S, Verneaux V, Millet L, Parent C, Magny M (2015) A case study of the past CH<sub>4</sub> cycle in lakes by the combined use of dual isotopes (carbon and hydrogen) and ancient DNA of methane-oxidizing bacteria: rearing experiment and application to Lake Remoray (eastern France). *Aquatic Ecology*, **49**, 279–291.
- Bent SJ, Forney LJ (2008) The tragedy of the uncommon: understanding limitations in the analysis of microbial diversity. *The ISME Journal*, **2**, 689–695.
- Berkeley A, Perry CT, Smithers SG, Horton BP, Taylor KG (2007) A review of the ecological and taphonomic controls on foraminiferal assemblage development in intertidal environments. *Earth-Science Reviews*, **83**, 205–230.
- Berthon V, Alric B, Rimet F, Perga M-E (2014) Sensitivity and responses of diatoms to climate warming in lakes heavily influenced by humans. *Freshwater Biology*, **59**, 1755–1767.
- Berthon V, Marchetto A, Rimet F *et al.* (2013) Trophic history of French sub-alpine lakes over the last ~150 years: phosphorus reconstruction and assessment of taphonomic biases. *Journal of Limnology*, **72**, e34.
- Bertrand J-C, Lebaron P, Normand P, Caumette P, Matheron R (2011) *Écologie microbienne : microbiologie des milieux naturels et anthropisés*. Presses universitaires

de Pau et des Pays de l'Adour.

- Bik HM, Sung W, De Ley P *et al.* (2012) Metagenetic community analysis of microbial eukaryotes illuminates biogeographic patterns in deep-sea and shallow water sediments. *Molecular Ecology*, **21**, 1048–1059.
- Billard E, Domaizon I, Tissot N, Arnaud F, Lyautey E (2015) Multi-scale phylogenetic heterogeneity of archaea, bacteria, methanogens and methanotrophs in lake sediments. *Hydrobiologia*, **751**, 159–173.
- Bissett A, Gibson J, Jarman S, Swadling K, Cromer L (2005) Isolation, amplification, and identification of ancient copepod DNA from lake sediments. *Limnology and Oceanography: Methods*.
- 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.
- Blenckner T, Adrian R, Livingstone D *et al.* (2007) Large-scale climatic signatures in lakes across Europe: a meta-analysis. *Global Change Biology*, **13**, 1314–1326.
- Boere AC, Abbas B, Rijpstra WIC *et al.* (2009) Late-Holocene succession of dinoflagellates in an Antarctic fjord using a multi-proxy approach: paleoenvironmental genomics, lipid biomarkers and palynomorphs. *Geobiology*, **7**, 265–81.
- Boere AC, Rijpstra WIC, De Lange GJ, Sinninghe Damsté JS, Coolen MJL (2011a) Preservation potential of ancient plankton DNA in Pleistocene marine sediments. *Geobiology*, **9**, 377–93.
- Boere AC, Sinninghe Damsté JS, Rijpstra WIC, Volkman JK, Coolen MJL (2011b) Source-specific variability in post-depositional DNA preservation with potential implications for DNA based paleoecological records. *Organic Geochemistry*, **42**, 1216–1225.
- Borcard D, Gillet F, Legendre P (2011) *Numerical Ecology with R* (Springer, Ed.). New York, NY.
- Bradley IM, Pinto AJ, Guest JS (2016) Gene-Specific Primers for Improved Characterization of Mixed Phototrophic Communities. *Applied and Environmental Microbiology*, **82**, 5878–5891.
- Bråte J, Logares R, Berney C *et al.* (2010) Freshwater Perkinsea and marine-freshwater colonizations revealed by pyrosequencing and phylogeny of environmental rDNA. *The ISME Journal*, **4**, 1144–1153.
- Briski E, Cristescu ME, Bailey SA, MacIsaac HJ (2011) Use of DNA barcoding to detect invertebrate invasive species from diapausing eggs. *Biological Invasions*, **13**, 1325–1340.
- Brown M V, Lauro FM, DeMaere MZ *et al.* (2012) Global biogeography of SAR11 marine

- bacteria. *Molecular systems biology*, **8**, 595.
- Bryant JA, Aylward FO, Eppley JM *et al.* (2016) Wind and sunlight shape microbial diversity in surface waters of the North Pacific Subtropical Gyre. *The ISME Journal*, **10**, 1308–1322.
- Büntgen U, Tegel W, Nicolussi K *et al.* (2011) 2500 Years of European Climate Variability and Human Susceptibility. *Science*, **331**, 578–83.
- Burki F (2014) The eukaryotic tree of life from a global phylogenomic perspective. *Cold Spring Harbor Perspectives in Biology*, **6**, a016147.
- Buttigieg PL, Ramette A (2014) A Guide to Statistical Analysis in Microbial Ecology: a community-focused, living review of multivariate data analyses. *FEMS Microbiology Ecology*, **90**, 543–550.
- Cáceres CE (1998) Interspecific variation in the abundance, production, and emergence of *Daphnia* diapausin eggs. *Ecology*, **79**, 1699–1710.
- Callieri C, Stockner JG (2002) Freshwater autotrophic picoplankton: a review. *Journal of Limnology*, **61**, 1.
- del Campo J, Sieracki ME, Molestina R *et al.* (2014) The others: our biased perspective of eukaryotic genomes. *Trends in Ecology & Evolution*, **29**, 252–9.
- Capo E, Debroas D, Arnaud F *et al.* Long-term dynamics in microbial eukaryotes communities: a paleolimnological view based on sedimentary DNA. *Molecular Ecology*.
- Capo E, Debroas D, Arnaud F, Domaizon I (2015) Is Planktonic Diversity Well Recorded in Sedimentary DNA? Toward the Reconstruction of Past Protistan Diversity. *Microbial Ecology*, **70**, 865–875.
- Caporaso JG, Lauber CL, Walters WA *et al.* (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME Journal*, **6**, 1621–4.
- Caron DA, Countway PD, Savai P *et al.* (2009a) Defining DNA-Based Operational Taxonomic Units for Microbial-Eukaryote Ecology. *Applied and Environmental Microbiology*, **75**, 5797–5808.
- Caron DA, Hutchins DA (2013) The effects of changing climate on microzooplankton grazing and community structure: drivers, predictions and knowledge gaps. *Journal of Plankton Research*, **35**, 235–252.
- Caron D a., Peele ER, Lim EL, Dennett MR (1999) Picoplankton and nanoplankton and their trophic coupling in surface waters of the Sargasso Sea south of Bermuda. *Limnology and Oceanography*, **44**, 259–272.
- Caron DA, Worden AZ, Countway PD, Demir E, Heidelberg KB (2009b) Protists are microbes too: a perspective. *The ISME Journal*, **3**, 4–12.

- Carpenter SR, Bennett EM (2011) Reconsideration of the planetary boundary for phosphorus. *Environmental Research Letters*, **6**, 014009.
- Chaffron S, Rehrauer H, Pernthaler J *et al.* (2010) A global network of coexisting microbes from environmental and whole-genome sequence data. *Genome Research*, **20**, 947–59.
- Chambouvet A, Berney C, Romac S *et al.* (2014) Diverse molecular signatures for ribosomally “active” Perkinsea in marine sediments. *BMC Microbiology*, **14**, 110.
- Chao A, Gotelli NJ, Hsieh TC *et al.* (2014) Rarefaction and extrapolation with Hill numbers: A framework for sampling and estimation in species diversity studies. *Ecological Monographs*, **84**, 45–67.
- Charvet S, Vincent WF, Comeau A, Lovejoy C (2012) Pyrosequencing analysis of the protist communities in a High Arctic meromictic lake : DNA preservation and change. *Frontiers in Microbiology*, **3**, 1–14.
- 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.
- Chen M, Chen F, Yu Y, Ji J, Kong F (2008) Genetic diversity of eukaryotic microorganisms in Lake Taihu, a large shallow subtropical lake in China. *Microbial Ecology*, **56**, 572–83.
- Chen L, Zonneveld KAF, Versteegh GJM (2011) Short term climate variability during “Roman Classical Period” in the eastern Mediterranean. *Quaternary Science Reviews*, **30**, 3880–3891.
- Chèvre N (2014) Pharmaceuticals in surface waters: sources, behavior, ecological risk, and possible solutions. Case study of Lake Geneva, Switzerland. *Wiley Interdisciplinary Reviews: Water*, **1**, 69–86.
- Chipman ML, Kling GW, Lundstrom CC, Hu FS (2016) Multiple thermo-erosional episodes during the past six millennia: Implications for the response of Arctic permafrost to climate change. *Geology*, **44**, G37693.1.
- Chow C-ET, Kim DY, Sachdeva R, Caron DA, Fuhrman JA (2014) Top-down controls on bacterial community structure: microbial network analysis of bacteria, T4-like viruses and protists. *The ISME Journal*, **8**, 816–29.
- Chow C-ET, Sachdeva R, Cram J a *et al.* (2013) Temporal variability and coherence of euphotic zone bacterial communities over a decade in the Southern California Bight. *The ISME Journal*, **7**, 2259–2273.
- Chow C-ET, Suttle CA (2015) Biogeography of Viruses in the Sea. *Annual Review of Virology*, **2**, 150902154948005.
- Christner BC, Priscu JC, Achberger AM *et al.* (2014) A microbial ecosystem beneath the West Antarctic ice sheet. *Nature*, **512**, 310–313.

- Coolen MJL, Gibson JAE (2009) Ancient DNA in lake sediment records. *PAGES news*, **17**, 104–106.
- Coolen MJL, Muyzer G, Rijpstra WIC *et al.* (2004) Combined DNA and lipid analyses of sediments reveal changes in Holocene haptophyte and diatom populations in an Antarctic lake. *Earth and Planetary Science Letters*, **223**, 225–239.
- Coolen MJL, Orsi WD, Balkema C *et al.* (2013) Evolution of the plankton paleome in the Black Sea from the Deglacial to Anthropocene. *Proceedings of the National Academy of Sciences*, **110**, 8609–14.
- Coolen MJ, Overmann J (1998) Analysis of subfossil molecular remains of purple sulfur bacteria in a lake sediment. *Applied and Environmental Microbiology*, **64**, 4513–21.
- Coolen MJL, Saenz JP, Giosan L *et al.* (2009) DNA and lipid molecular stratigraphic records of haptophyte succession in the Black Sea during the Holocene. *Earth and Planetary Science Letters*, **284**, 610–621.
- Coolen MJL, Shtereva G (2009) Vertical distribution of metabolically active eukaryotes in the water column and sediments of the Black Sea. *FEMS microbiology ecology*, **70**, 525–39.
- Coolen MJL, Talbot HM, Abbas BA *et al.* (2008) Sources for sedimentary bacterioplanepolyols as revealed by 16S rDNA stratigraphy. *Environmental Microbiology*, **10**, 1783–803.
- Corinaldesi C, Barucca M, Luna GM, Dell'Anno A (2011) Preservation, origin and genetic imprint of extracellular DNA in permanently anoxic deep-sea sediments. *Molecular ecology*, **20**, 642–54.
- Corinaldesi C, Beolchini F, Dell'Anno A (2008) Damage and degradation rates of extracellular DNA in marine sediments: implications for the preservation of gene sequences. *Molecular Ecology*, **17**, 3939–51.
- Corinaldesi C, Danovaro R, Dell'Anno A (2005) Simultaneous recovery of extracellular and intracellular DNA suitable for molecular studies from marine sediments. *Applied and Environmental Microbiology*, **71**, 46–50.
- Coveley S, Elshahed MS, Youssef NH (2015) Response of the rare biosphere to environmental stressors in a highly diverse ecosystem (Zodletone spring, OK, USA). *PeerJ*, **3**, e1182.
- Cram JA, Chow C-ET, Sachdeva R *et al.* (2015) Seasonal and interannual variability of the marine bacterioplankton community throughout the water column over ten years. *The ISME Journal*, **9**, 563–580.
- D'Amore R, Ijaz UZ, Schirmer M *et al.* (2016) A comprehensive benchmarking study of protocols and sequencing platforms for 16S rRNA community profiling. *BMC Genomics*, **17**, 55.

- Davidson TA, Jeppesen E (2013) The role of palaeolimnology in assessing eutrophication and its impact on lakes. *Journal of Paleolimnology*, **49**, 391–410.
- Davidson T, Sayer C, Perrow M, Tomlinson M (2003) Representation of fish communities by scale sub-fossils in shallow lakes: implications for inferring percid–cyprinid shifts. *Journal of Paleolimnology*, **30**, 441–449.
- De'ath G (2002) Multivariate regression trees: a new technique: a new technique for modeling species - environment relationships. *Ecology*, **83**, 1105–1117.
- Debroas D, Hugoni M, Domaizon I (2015) Evidence for an active rare biosphere within freshwater protists community. *Molecular Ecology*, **24**, 1236–1247.
- Dell'Anno A, Bompadre S, Danovaro R (2002) Quantification, base composition, and fate of extracellular DNA in marine sediments. *Limnology and Oceanography*, **47**, 899–905.
- Dell'Anno A, Corinaldesi C (2004) Degradation and Turnover of Extracellular DNA in Marine Sediments : Ecological and Methodological Considerations Degradation and Turnover of Extracellular DNA in Marine Sediments : Ecological and Methodological Considerations. *Applied and Environmental Microbiology*, **70**, 4384–4386.
- Dell'Anno A, Danovaro R (2005) Extracellular DNA plays a key role in deep-sea ecosystem functioning. *Science*, **309**, 2179.
- Dittrich M, Kurz P, Wehrli B (2004) The Role of Autotrophic Picocyanobacteria in Calcite Precipitation in an Oligotrophic Lake. *Geomicrobiology Journal*, **21**, 45–53.
- Dokulil MT (2014) Impact of climate warming on European inland waters. *Inland Waters*, **4**, 27–40.
- Domaizon I, Savichtcheva O, Debroas D *et al.* (2013) DNA from lake sediments reveals the long-term dynamics and diversity of *Synechococcus* assemblages. *Biogeosciences Discussions*, **10**, 2515–2564.
- Dornelas M, Magurran AE, Buckland ST *et al.* (2013) Quantifying temporal change in biodiversity : challenges and opportunities. *Proceedings of the Royal Society B*, **280**.
- Downing JA, McCauley E (1992) The nitrogen : phosphorus relationship in lakes. *Limnology and Oceanography*, **37**, 936–945.
- Downing JA, Prairie YT, Cole JJ *et al.* (2006) The global abundance and size distribution of lakes, ponds and impoundments. *Limnology and Oceanography*, **51**, 2388–2397.
- Dunthorn M, Otto J, Berger S a *et al.* (2014) Placing environmental next-generation sequencing amplicons from microbial eukaryotes into a phylogenetic context. *Molecular biology and evolution*, **31**, 993–1009.
- Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics (Oxford, England)*, **26**, 2460–1.
- Edgar RC (2013) UPARSE: highly accurate OTU sequences from microbial amplicon reads.

*Nature methods*, **10**, 996–8.

- Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011) UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics (Oxford, England)*, **27**, 2194–200.
- Edgcomb VP, Beaudoin D, Gast R, Biddle JF, Teske A (2011) Marine subsurface eukaryotes: the fungal majority. *Environmental Microbiology*, **13**, 172–83.
- Egge E, Johannessen T, Andersen T *et al.* (2015) Seasonal diversity and dynamics of haptophytes in the Skagerrak, Norway, explored by high-throughput sequencing. *Molecular Ecology*, **24**, 3026–3042.
- Eichmiller JJ, Bajer PG, Sorensen PW (2014) The Relationship between the Distribution of Common Carp and Their Environmental DNA in a Small Lake (AC Anil, Ed.). *PLoS ONE*, **9**, e112611.
- Eiler A, Heinrich F, Bertilsson S (2012) Coherent dynamics and association networks among lake bacterioplankton taxa. *The ISME Journal*, **6**, 330–42.
- Eland LE, Davenport R, Mota CR (2012) Evaluation of DNA extraction methods for freshwater eukaryotic microalgae. *Water Research*, **46**, 5355–5364.
- Elliott J (2012) Predicting the impact of changing nutrient load and temperature on the phytoplankton of England's largest lake, Windermere. *Freshwater Biology*, **57**, 400–413.
- Elliott J, May L (2008) The sensitivity of phytoplankton in Loch Leven (U.K.) to changes in nutrient load and water temperature. *Freshwater Biology*, **53**, 32–41.
- Engstrom DR, Schottler SP, Leavitt PR, Havens KE (2006) A reevaluation of the cultural eutrophication of Lake Okeechobee using multiproxy sediment records. *Ecological Applications*, **16**, 1194–1206.
- Epp LS, Gussarova G, Boessenkool S *et al.* (2015) Lake sediment multi-taxon DNA from North Greenland records early post-glacial appearance of vascular plants and accurately tracks environmental changes. *Quaternary Science Reviews*, **117**, 152–163.
- Epp LS, Stoof KR, Trauth MH, Tiedemann R (2010) Historical genetics on a sediment core from a Kenyan lake: intraspecific genotype turnover in a tropical rotifer is related to past environmental changes. *Journal of Paleolimnology*, **43**, 939–954.
- Epp LS, Stoof-Leichsenring KR, Trauth MH, Tiedemann R (2011) Molecular profiling of diatom assemblages in tropical lake sediments using taxon-specific PCR and Denaturing High-Performance Liquid Chromatography (PCR-DHPLC). *Molecular Ecology Resources*, **11**, 842–53.
- Erdman C, Emerson JW (2007) bcp: An R Package for Performing a Bayesian Analysis of Change Point Problems. *Journal of Statistical Software*, **23**.
- Esling P, Lejzerowicz F, Pawlowski J (2015) Accurate multiplexing and filtering for high-throughput amplicon-sequencing. *Nucleic Acids Research*, **43**, 2513–2524.



- Falkowski PG, Fenchel T, Delong EF (2008) The microbial engines that drive Earth's biogeochemical cycles. *Science*, **320**, 1034–9.
- Falkowski PG, Oliver MJ (2007) Mix and match: how climate selects phytoplankton. *Nature reviews. Microbiology*, **5**, 813–819.
- Faure F, Demars C, Wieser O *et al.* (2015) Plastic pollution in Swiss surface waters: nature and concentrations, interaction with pollutants. *Environmental Chemistry*, **12**, 582.
- Faust K, Lahti L, Gonze D, de Vos WM, Raes J (2015) Metagenomics meets time series analysis: Unraveling microbial community dynamics. *Current Opinion in Microbiology*, **25**, 56–66.
- Faust K, Raes J (2012) Microbial interactions: from networks to models. *Nature reviews. Microbiology*, **10**, 538–50.
- Faust K, Sathirapongsasuti JF, Izard J *et al.* (2012) Microbial co-occurrence relationships in the human microbiome. (CA Ouzounis, Ed.). *PLoS computational biology*, **8**, e1002606.
- Fee EJ (1979) A relation between lake morphometry and primary productivity and its use in interpreting whole-lake eutrophication experiments. *Limnology and Oceanography*, **24**, 401–416.
- Fenchel T, Finlay BJ (1991) Synchronous Division of an Endosymbiotic Methanogenic Bacterium In the Anaerobic Ciliate Plagiopyla Frontata Kahl. *The Journal of Protozoology*, **38**, 22–28.
- Fernandez-Carazo R, Verleyen E, Hodgson DA *et al.* (2013) Late Holocene changes in cyanobacterial community structure in maritime Antarctic lakes. *Journal of Paleolimnology*, **50**, 15–31.
- Filker S, Sommaruga R, Vila I, Stoeck T (2016) Microbial eukaryote plankton communities of high-mountain lakes from three continents exhibit strong biogeographic patterns. *Molecular Ecology*, **25**, 2286–2301.
- Finlay BJ (2002) Global dispersal of free-living microbial eukaryote species. *Science*, **296**, 1061–3.
- Finlay BJ, Fenchel T (2004) Cosmopolitan Metapopulations of Free-Living Microbial Eukaryotes. *Protist*, **155**, 237–244.
- Finney BP, Gregory-Eaves I, Sweetman J, Douglas MS, Smol JP (2000) Impacts of climatic change and fishing on Pacific salmon abundance over the past 300 years. *Science*, **290**, 795–9.
- Fischer H, Sachse A, Steinberg CEW, Pusch M (2002) Differential retention and utilization of dissolved organic carbon by bacteria in river sediments. *Limnology and Oceanography*, **47**, 1702–1711.
- Foissner W (1988) Taxonomic and nomenclatural revision of Sládeček's list of ciliates

- (Protozoa: Ciliophora) as indicators of water quality. *Hydrobiologia*, **166**, 1–64.
- Foissner W (1999) Protist Diversity: Estimates of the Near-Imponderable. *Protist*, **150**, 363–368.
- Foissner W (2006) Biogeography and dispersal of micro-organisms : A review emphasizing protists. *Acta Protozoologica*, **45**, 111–136.
- Foissner W (2008) Protist diversity and distribution: some basic considerations. *Biodiversity and Conservation*, **17**, 235–242.
- Foissner W, Chao A, Katz LA (2008) Diversity and geographic distribution of ciliates (Protista: Ciliophora). *Biodiversity and Conservation*, 345–363.
- Frey KG, Herrera-Galeano J, Redden CL *et al.* (2014) Comparison of three next-generation sequencing platforms for metagenomic sequencing and identification of pathogens in blood. *BMC Genomics*, **15**, 96.
- Frontier S, Pischod-Viale D, Leprêtre A, Dominique D, Luczak C (2008) *Écosystèmes : structure, fonctionnement, évolution*. Dunod.
- Frossard V, Millet L, Verneaux V *et al.* (2013) Chironomid assemblages in cores from multiple water depths reflect oxygen-driven changes in a deep French lake over the last 150 years. *Journal of Paleolimnology*, **50**, 257–273.
- Fuhrman J, Cram J, Needham D (2015) Marine microbial community dynamics and their ecological interpretation. *Nature Reviews Microbiology*, **13**, 133–146.
- Gallina N, Salmaso N, Morabito G, Beniston M (2013) Phytoplankton configuration in six deep lakes in the peri-Alpine region : are the key drivers related to eutrophication and climate ? *Aquatic Ecology*, **47**, 177–193.
- Gälman V, Petterson G, Renberg I (2006) A Comparison of Sediment Varves (1950–2003 AD) in Two Adjacent Lakes in Northern Sweden. *Journal of Paleolimnology*, **35**, 837–853.
- Gälman V, Rydberg J, Bigler C (2009a) Decadal diagenetic effects on d13C and d15N studied in varved lake sediment. *Limnology and Oceanography*, **54**, 917–924.
- Gälman V, Rydberg J, De-Luna SS, Bindler R, Renberg I (2008) Carbon and nitrogen loss rates during aging of lake sediment : Changes over 27 years studied in varved lake sediment. *Limnology and oceanography*, **53**, 1076–1082.
- Gälman V, Rydberg J, Shchukarev A *et al.* (2009b) The role of iron and sulfur in the visual appearance of lake sediment varves. *Journal of Paleolimnology*, **42**, 141–153.
- Gast RJ, Dennett MR, Caron DA (2004) Characterization of Protistan Assemblages in the Ross Sea , Antarctica , by Denaturing Gradient Gel Electrophoresis. *Applied and Environmental Microbiology*, **70**, 2028–2037.
- Gauthier E, Bichet V, Massa C *et al.* (2010) Pollen and non-pollen palynomorph evidence of

- medieval farming activities in southwestern Greenland. *Vegetation History and Archaeobotany*, **19**, 427–438.
- Gibbons SM, Caporaso JG, Pirrung M *et al.* (2013) Evidence for a persistent microbial seed bank throughout the global ocean. *Proceedings of the National Academy of Sciences*, **110**, 4651–4655.
- Gibbons SM, Gilbert JA (2015) Microbial diversity - exploration of natural ecosystems and microbiomes. *Current Opinion in Genetics & Development*, **35**, 66–72.
- Giguet-Covex C, Arnaud F, Poulencard J *et al.* (2010) Sedimentological and geochemical records of past trophic state and hypolimnetic anoxia in large, hard-water Lake Bourget, French Alps. *Journal of Paleolimnology*, **43**, 171–190.
- Gilbert MTP, Bandelt H-J, Hofreiter M, Barnes I (2005) Assessing ancient DNA studies. *Trends in ecology & evolution*, **20**, 541–4.
- Gilbert J, Steele J, Caporaso JG *et al.* (2012) Defining seasonal marine microbial community dynamics. *The ISME Journal*, **6**, 298–308.
- Giner CR, Forn I, Romac S *et al.* (2016) Environmental sequencing provides reasonable estimates of the relative abundance of specific picoeukaryotes. *Applied and Environmental Microbiology*, **82**, AEM.00560–16.
- Giongo A, Crabb DB, Davis-Richardson AG *et al.* (2010) PANGEA: pipeline for analysis of next generation amplicons. *The ISME Journal*, **4**, 852–61.
- Giovannoni SJ, Vergin KL (2012) Seasonality in Ocean Microbial Communities. *Science*, **335**, 671–676.
- Girvan MS, Campbell CD, Killham K, Prosser JI, Glover LA (2005) Bacterial diversity promotes community stability and functional resilience after perturbation. *Environmental Microbiology*, **7**, 301–313.
- Godhe A, Asplund ME, Hårnström K *et al.* (2008) Quantification of diatom and dinoflagellate biomasses in coastal marine seawater samples by real-time PCR. *Applied and Environmental Microbiology*, **74**, 7174–82.
- Goldberg CS, Strickler KM, Pilliod DS (2015) Moving environmental DNA methods from concept to practice for monitoring aquatic macroorganisms. *Biological Conservation*, **183**, 1–3.
- Gotelli NJ, Chao A (2013) Measuring and Estimating Species Richness, Species Diversity, and Biotic Similarity from Sampling Data. In: *Encyclopedia of Biodiversity*, pp. 195–211. Elsevier.
- Grattepanche J-DD, Santoferrara LFLF, McManus GB *et al.* (2014) Diversity of diversity: Conceptual and methodological differences in biodiversity estimates of eukaryotic microbes as compared to bacteria. *Trends in Microbiology*, **22**, 432–437.

- Gregory-Eaves I, Beisner B (2011) Palaeolimnological insights for biodiversity science: an emerging field. *Freshwater Biology*, **56**, 2653–2661.
- Guidi L, Chaffron S, Bittner L *et al.* (2016) Plankton networks driving carbon export in the oligotrophic ocean. *Nature*, 1–16.
- Guillemot T, Zocatelli R, Bichet V *et al.* (2015) Evolution of pastoralism in Southern Greenland during the last two millennia reconstructed from bile acids and coprophilous fungal spores in lacustrine sediments. *Organic Geochemistry*, **81**, 40–44.
- Haegeman B, Hamelin J, Moriarty J *et al.* (2013) Robust estimation of microbial diversity in theory and in practice. *The ISME Journal*, **7**, 1092–1101.
- Hagelberg E, Hofreiter M, Keyser C (2015) Ancient DNA : the first three decades.
- Haglund A-L, Lantz P, Törnblom E, Tranvik L (2003) Depth distribution of active bacteria and bacterial activity in lake sediment. *FEMS Microbiology Ecology*, **46**, 31–38.
- Haile J, Holdaway R, Oliver K *et al.* (2007) Ancient DNA Chronology within Sediment Deposits: Are Paleobiological Reconstructions Possible and Is DNA Leaching a Factor? *Molecular Biology and Evolution*, **24**, 982–989.
- Hairston N, Fox J (2009) Egg banks. In: *Encyclopedia of Inland Waters*, p. vol. 3, pp. 659–666.
- Hairston NG, Lampert W, Cáceres CE *et al.* (1999) Lake ecosystems: Rapid evolution revealed by dormant eggs. *Nature*, **401**, 446–446.
- Hammer Ø, Harper DAT, Ryan PD (2001) PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica*, **4**, 9pp.
- Hamrová E, Goliáš V, Petrusek A (2010) Identifying century-old long-spined Daphnia: species replacement in a mountain lake characterised by paleogenetic methods. *Hydrobiologia*, **643**, 97–106.
- Harnstrom K, Ellegaard M, Andersen TJ, Godhe A (2011) Hundred years of genetic structure in a sediment revived diatom population. *Proceedings of the National Academy of Sciences*, **108**, 4252–4257.
- Hastie T, Tibshirani R (1986) Generalized Additive Models. *Statistical Science*, **1**, 297–310.
- Haukka K, Kolmonen E, Hyder R *et al.* (2006) Effect of nutrient loading on bacterioplankton community composition in lake mesocosms. *Microbial Ecology*, **51**, 137–146.
- Hebsgaard MB, Phillips MJ, Willerslev E (2005) Geologically ancient DNA: fact or artefact? *Trends in microbiology*, **13**, 212–20.
- Hill MO (1973) Diversity and Evenness: A Unifying Notation and Its Consequences. *Ecology*, **54**, 427–432.
- Hofreiter M, Serre D, Poinar HN, Kuch M, Pääbo S (2001) Ancient DNA. *Nature Reviews*

- Genetics*, **2**, 353–359.
- Hou W, Dong H, Li G *et al.* (2014) Identification of Photosynthetic Plankton Communities Using Sedimentary Ancient DNA and Their Response to late-Holocene Climate Change on the Tibetan Plateau. *Scientific Reports*, **4**, 6648.
- Hsieh TC, Ma KH, Chao A (2016) iNEXT: an R package for interpolation and extrapolation of species diversity (Hill numbers). *Methods in Ecology and Evolution*, 1–32.
- Hugoni M, Taib N, Debroas D *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 of the United States of America*, **110**, 6004–9.
- Ibelings BW, Portielje R, Lammens EHRR *et al.* (2007) Resilience of alternative stable states during the recovery of shallow lakes from eutrophication: Lake Veluwe as a case study. *Ecosystems*, **10**, 4–16.
- Imam SH, Buchanan MJ, Shin HC, Snell WJ (1985) The Chlamydomonas cell wall: Characterization of the wall framework. *Journal of Cell Biology*, **101**, 1599–1607.
- IPCC (2013) Climate Change 2013: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. In: Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
- Ishida S, Ohtsuki H, Awano T *et al.* (2012) DNA extraction and amplification methods for ephippial cases of Daphnia resting eggs in lake sediments: a novel approach for reconstructing zooplankton population structure from the past. *Limnology*, **13**, 261–267.
- Jackson DA, Harvey HH (1993) Fish and Benthic Invertebrates: Community Concordance and Community–Environment Relationships. *Canadian Journal of Fisheries and Aquatic Sciences*, **50**, 2641–2651.
- Jacquet S, Domaizon I, Anneville O (2014) The need for ecological monitoring of freshwaters in a changing world: A case study of Lakes Annecy, Bourget, and Geneva. *Environmental Monitoring and Assessment*, **186**, 3455–3476.
- Jenny JP, Arnaud F, Alric B *et al.* (2014) Inherited hypoxia: A new challenge for reoligotrophicated lakes under global warming. *Global Biogeochemical Cycles*, **28**, 1413–1423.
- Jenny J-P, Arnaud F, Dorioz J-M *et al.* (2013) A spatiotemporal investigation of varved sediments highlights the dynamics of hypolimnetic hypoxia in a large hard-water lake over the last 150 years. *Limnology and Oceanography*, **58**, 1395–1408.
- Jensen JP, Jeppesen E, Olrik K, Kristensen P (1994) Impact of Nutrients and Physical Factors on the Shift from Cyanobacterial to Chlorophyte Dominance in Shallow Danish Lakes. *Canadian Journal of Fisheries and Aquatic Sciences*, **51**, 1692–1699.
- Jeppesen E, Jensen JP, Søndergaard M (2002) Response of phytoplankton, zooplankton, and

- fish to re-oligotrophication: an 11 year study of 23 Danish lakes. *Aquatic Ecosystem Health Management*, **5**, 31–43.
- Jeppesen E, Jensen JP, Søndergaard M, Lauridsen T, Landkildehus F (2000) Trophic structure, species richness and biodiversity in Danish lakes: Changes along a phosphorus gradient. *Freshwater Biology*, **45**, 201–218.
- Jeppesen E, Søndergaard M, Jensen JP *et al.* (2005) Lake responses to reduced nutrient loading - An analysis of contemporary long-term data from 35 case studies. *Freshwater Biology*, **50**, 1747–1771.
- Jeppesen E, Søndergaard M, Meerhoff M, Lauridsen TL, Jensen JP (2007) Shallow lake restoration by nutrient loading reduction - Some recent findings and challenges ahead. *Hydrobiologia*, **584**, 239–252.
- Jerde CL, Mahon AR, Chadderton WL, Lodge DM (2011) “Sight-unseen” detection of rare aquatic species using environmental DNA. *Conservation Letters*, **4**, 150–157.
- Jessup CM, Kassen R, Forde SE *et al.* (2004) Big questions, small worlds: Microbial model systems in ecology. *Trends in Ecology and Evolution*, **19**, 189–197.
- Jiang Y, Xu H, Hu X, Warren A, Song W (2013) Functional groups of marine ciliated protozoa and their relationships to water quality. *Environmental science and pollution research international*, **20**, 5272–80.
- Jones E, Hyde K, Pang K (2014) Freshwater fungi: and fungal-like organisms.
- Jones SE, Lennon JT (2010) Dormancy contributes to the maintenance of microbial diversity. *Proceedings of the National Academy of Sciences*, **107**, 5881–5886.
- Jones AC, Liao TSV, Najjar FZ *et al.* (2013) Seasonality and disturbance: annual pattern and response of the bacterial and microbial eukaryotic assemblages in a freshwater ecosystem. *Environmental Microbiology*, **15**, 2557–72.
- Jørgensen T, Haile J, Möller P *et al.* (2012) A comparative study of ancient sedimentary DNA, pollen and microfossils from permafrost sediments of northern Siberia reveals long-term vegetational stability. *Molecular Ecology*, **21**, 1989–2003.
- Jost L (2007) Partitioning diversity into independent alpha and beta components. *Ecology*, **88**, 2427–2439.
- Kagami M, De Bruin A, Ibelings BW, Van Donk E (2007) Parasitic chytrids: Their effects on phytoplankton communities and food-web dynamics. *Hydrobiologia*, **578**, 113–129.
- Kammerlander B, Breiner H-W, Filker S *et al.* (2015) High diversity of protistan plankton communities in remote high mountain lakes in the European Alps and the Himalayan mountains. *FEMS Microbiology Ecology*, **91**, 1–10.
- Kara EL, Hanson PC, Hu YH, Winslow L, McMahan KD (2013) A decade of seasonal dynamics and co-occurrences within freshwater bacterioplankton communities from

- eutrophic Lake Mendota, WI, USA. *The ISME Journal*, **7**, 680–684.
- Kaufman DS, Schneider DP, McKay NP *et al.* (2009) Recent Warming Reverses Long-Term Arctic Cooling. *Science*.
- Kembel SW, Cowan PD, Helmus MR *et al.* (2010) Picante: R tools for integrating phylogenies and ecology. *Bioinformatics (Oxford, England)*, **26**, 1463–4.
- Kerfoot WC, Robbins JA, Weider LJ (1999) A new approach to historical reconstruction: Combining descriptive and experimental paleolimnology. *Limnology and Oceanography*, **44**, 1232–1247.
- Kirkpatrick JB, Walsh EA, D'Hondt S (2016) Fossil DNA persistence and decay in marine sediment over hundred-thousand-year to million-year time scales. *Geology*, G37933.1.
- Kisand V, Zingel P (2000) Dominance of ciliate grazing on bacteria during spring in a shallow eutrophic lake. *Aquatic Microbial Ecology*, **22**, 135–142.
- Klouch KZ, Schmidt S, Andrieux-Loyer F *et al.* (2016) Historical records from dated sediment cores reveal the multidecadal dynamic of the toxic dinoflagellate *Alexandrium minutum* in the Bay of Brest (France). *FEMS Microbiology Ecology*, 1–16.
- Koid A, Nelson WC, Mraz A, Heidelberg KB (2012) Comparative analysis of eukaryotic marine microbial assemblages from 18s rna gene and gene transcript clone libraries by using different methods of extraction. *Applied and Environmental Microbiology*, **78**, 3958–3965.
- Kong F-X, Gao G (2005) Hypothesis on cyanobacteria bloom-forming mechanism in large shallow eutrophic lakes. *Acta Ecologica Sinica*, **03**.
- Kopylova E, Navas-Molina JA, Mercier C *et al.* (2016) Open-Source Sequence Clustering Methods Improve the State Of the Art. *mSystems*, **1**, e00003–15.
- Kosten S, Huszar VLM, Bécares E *et al.* (2012) Warmer climates boost cyanobacterial dominance in shallow lakes. *Global Change Biology*, **18**, 118–126.
- Kraemer BM, Chandra S, Dix M *et al.* (2016) Global patterns in aquatic ecosystem responses to warming based on activation energies. *Proceedings of the National Academy of Sciences of the United States of America*, 1–10.
- Kristensen E (2000) Organic matter diagenesis at the oxic/anoxic interface in coastal marine sediments, with emphasis on the role of burrowing animals. *Hydrobiologia*, **426**, 1–24.
- Kyle M, Haande S, Ostermaier V *et al.* (2015a) The Red Queen Race between Parasitic Chytrids and Their Host, *Planktothrix*: A Test Using a Time Series Reconstructed from Sediment DNA (S Ban, Ed.). *PLOS ONE*, **10**, e0118738.
- Kyle M, Haande S, Sønstebo J, Rohrlack T (2015b) Amplification of DNA in sediment cores to detect historic *Planktothrix* occurrence in three Norwegian lakes. *Journal of Paleolimnology*, **53**, 61–72.

- Lacoursière-Roussel A, Côté G, Leclerc V, Bernatchez L (2016) Quantifying relative fish abundance with eDNA: a promising tool for fisheries management (M Cadotte, Ed.). *Journal of Applied Ecology*, **53**, 1148–1157.
- Lange A, Jost S, Heider D *et al.* (2015) AmpliconDuo: A Split-Sample Filtering Protocol for High-Throughput Amplicon Sequencing of Microbial Communities (M Casiraghi, Ed.). *PLoS One*, **10**, e0141590.
- Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*, **9**, 559.
- Lara E, Seppely CVW, Garraza GG *et al.* (2015) Planktonic eukaryote molecular diversity: discrimination of minerotrophic and ombrotrophic peatland pools in Tierra del Fuego (Argentina). *Journal of Plankton Research*, **37**, 645–655.
- Leavitt PR, Fritz SC, Anderson NJ *et al.* (2009) Paleolimnological evidence of the effects on lakes of energy and mass transfer from climate and humans. *Limnology and Oceanography*, **54**, 2330–2348.
- Leavitt PR, Hodgson DA (2002) Sedimentary Pigments. In: *Tracking Environmental Change Using Lake Sediments*, pp. 295–325. Kluwer Academic Publishers, Dordrecht.
- Lefèvre E, Roussel B, Amblard C, Sime-Ngando T (2008) The molecular diversity of freshwater picoeukaryotes reveals high occurrence of putative parasitoids in the plankton. *PLoS One*, **3**, e2324.
- Lefranc M, Thénot A, Lepère C, Debroas D (2005a) Genetic Diversity of Small Eukaryotes in Lakes Differing by Their Trophic Status Genetic Diversity of Small Eukaryotes in Lakes Differing by Their Trophic Status. *Applied and Environmental Microbiology*, **71**, 5935–5942.
- Lefranc M, Thénot A, Lepère C, Debroas D (2005b) Genetic diversity of small eukaryotes in lakes differing by their trophic status. *Applied and Environmental Microbiology*, **71**, 5935–42.
- Legendre P, Legendre L (2012) *Numerical ecology* (Elsevier, Ed.).
- Lehmann MF, Bernasconi SM, Barbieri A, McKenzie JA (2002) Preservation of organic matter and alteration of its carbon and nitrogen isotope composition during simulated and in situ early sedimentary diagenesis. *Geochimica et Cosmochimica Acta*, **66**, 3573–3584.
- Lejzerowicz F, Esling P, Majewski W *et al.* (2013) Ancient DNA complements microfossil record in deep-sea subsurface sediments. *Biology letters*, **9**, 20130283.
- Lentendu G, Wubet T, Chatzinotas A *et al.* (2014) Effects of long-term differential fertilization on eukaryotic microbial communities in an arable soil: A multiple barcoding approach. *Molecular Ecology*, **23**, 3341–3355.



- Lepère C, Boucher D, Jardillier L, Domaizon I, Debroas D (2006) Succession and regulation factors of small eukaryote community composition in a lacustrine ecosystem (Lake Pavin). *Applied and Environmental Microbiology*, **72**, 2971–81.
- Lepère C, Domaizon I, Debroas D (2008) Unexpected importance of potential parasites in the composition of the freshwater small-eukaryote community. *Applied and Environmental Microbiology*, **74**, 2940–9.
- Lepère C, Domaizon I, Taïb N *et al.* (2013) Geographic distance and ecosystem size determine the distribution of smallest protists in lacustrine ecosystems. *FEMS Microbiology Ecology*, **85**, 85–94.
- Letunic I, Bork P (2011) Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. *Nucleic acids research*, **39**, W475–8.
- Lever MA, Torti A, Eickenbusch P *et al.* (2015) A modular method for the extraction of DNA and RNA, and the separation of DNA pools from diverse environmental sample types. *Frontiers in Microbiology*, **6**, 476.
- Li G, Dong H, Hou W *et al.* (2016) Temporal Succession of Ancient Phytoplankton Community in Qinghai Lake and Implication for Paleo-environmental Change. *Scientific Reports*, **6**, 19769.
- Lieberman BS (2003) Paleobiogeography: The Relevance of Fossils to Biogeography. *Annual Review of Ecology, Evolution, and Systematics*, **34**, 51–69.
- Lima-Mendez G, Faust K, Henry N *et al.* (2015) Determinants of community structure in the global plankton interactome. *Science*, **348**, 1262073\_1–1262073\_9.
- Lloyd KG, Macgregor BJ, Teske A *et al.* (2010) Quantitative PCR methods for RNA and DNA in marine sediments: maximizing yield while overcoming inhibition. *FEMS microbiology ecology*, **72**, 143–51.
- Logares R, Audic S, Bass D *et al.* (2014) Patterns of Rare and Abundant Marine Microbial Eukaryotes. *Current Biology*, **24**, 1–9.
- Logares R, Mangot J-F, Massana R (2015) Rarity in aquatic microbes: placing protists on the map. *Research in Microbiology*, **166**, 831–841.
- Logares R, Sunagawa S, Salazar G *et al.* (2013) Metagenomic 16S rDNA Illumina tags are a powerful alternative to amplicon sequencing to explore diversity and structure of microbial communities. *Environmental microbiology*, **16**, 2659–2671.
- Lom J (2002) A catalogue of described genera and species of microsporidians parasitic in fish. *Systematic Parasitology*, **53**, 81–99.
- Londeix L, Herreyre Y, Turon JL, Fletcher W (2009) Last Glacial to Holocene hydrology of the Marmara Sea inferred from a dinoflagellate cyst record. *Review of Palaeobotany and Palynology*, **158**, 52–71.

- Lynch MDJ, Neufeld JD (2015) Ecology and exploration of the rare biosphere. *Nature Reviews Microbiology*, **13**, 217–229.
- Madeja J, Wacnik A, Wypasek E, Chandran A, Stanjiewicz E (2010) Integrated palynological and molecular analyses of late Holocene deposits from Lake Miłkowskie (NE Poland): Verification of local human impact on environment. *Quaternary international*, **220**, 147–152.
- Madoni P (2005) Ciliated protozoan communities and saprobic evaluation of water quality in the hilly zone of some tributaries of the Po River (northern Italy). *Hydrobiologia*, **541**, 55–69.
- Madoni P, Bassanini N (1999) Longitudinal changes in the ciliated protozoa communities along a fluvial system polluted by organic matter. *European Journal of Protistology*, **35**, 391–402.
- Mahon AR, Jerde CL, Galaska M *et al.* (2013) Validation of eDNA Surveillance Sensitivity for Detection of Asian Carps in Controlled and Field Experiments (MR Liles, Ed.). *PLoS ONE*, **8**, e58316.
- Maier DB, Rydberg J, Bigler C, Renberg I (2013) Compaction of recent varved lake sediments. *GFF*, 1–6.
- Majaneva M, Hyytiainen K, Varvio SL, Nagai S, Blomster J (2015) Bioinformatic amplicon read processing strategies strongly affect eukaryotic diversity and the taxonomic composition of communities. *PLoS One*, **10**, 1–18.
- Makino W, Ohtsuki H, Urabe J (2013) Finding copepod footprints: a protocol for molecular identification of diapausing eggs in lake sediments. *Limnology*, **14**, 269–282.
- Mangot JF, Debroas D, Domaizon I (2011) Perkinsozoa, a well-known marine protozoan flagellate parasite group, newly identified in lacustrine systems: A review. *Hydrobiologia*, **659**, 37–48.
- Mangot J-F, Domaizon I, Taib N *et al.* (2013) Short-term dynamics of diversity patterns: evidence of continual reassembly within lacustrine small eukaryotes. *Environmental Microbiology*, **15**, 1745–58.
- Mangot J-F, Lepère C, Bouvier C, Debroas D, Domaizon I (2009) Community structure and dynamics of small eukaryotes targeted by new oligonucleotide probes: new insight into the lacustrine microbial food web. *Applied and environmental microbiology*, **75**, 6373–81.
- Mann ME (2003) Global surface temperatures over the past two millennia. *Geophysical Research Letters*, **30**, 15–18.
- Massa C, Perren BB, Gauthier É *et al.* (2012) A multiproxy evaluation of Holocene environmental change from Lake Igaliku, South Greenland. *Journal of Paleolimnology*, **48**, 241–258.

- Massana R, Del Campo J, Sieracki ME *et al.* (2014) Exploring the uncultured microeukaryote majority in the oceans: reevaluation of ribogroups within stramenopiles. *The ISME Journal*, **8**, 854–66.
- Massana R, Gobet A, Audic S *et al.* (2015) Marine protist diversity in European coastal waters and sediments as revealed by high-throughput sequencing. *Environmental Microbiology*, **17**, 4035–4049.
- McElroy K, Thomas T, Luciani F (2014) Deep sequencing of evolving pathogen populations: applications, errors, and bioinformatic solutions. *Microbial Informatics and Experimentation*, **4**, 1.
- Mckay NP, Kaufman DS (2014) An extended Arctic proxy temperature database for the past 2,000 years. *Scientific data*, 1–10.
- Meyers PA, Lallier-vergés E (1999) Lacustrine Sedimentary Organic Matter Records of Late Quaternary Paleoclimates. *Journal of Paleolimnology*, **21**, 345–372.
- Millet L, Arnaud F, Heiri O *et al.* (2009) Late-Holocene summer temperature reconstruction from chironomid assemblages of Lake Anterne, northern French Alps. *The Holocene*, **19**, 317–328.
- Millet L, Giguët-Covex C, Verneaux V *et al.* (2010) Reconstruction of the recent history of a large deep prealpine lake (Lake Bourget, France) using subfossil chironomids, diatoms, and organic matter analysis: towards the definition of a lake-specific reference state. *Journal of Paleolimnology*, **44**, 963–978.
- Millet L, Massa C, Bichet V *et al.* (2014) Anthropogenic versus climatic control in a high-resolution 1500-year chironomid stratigraphy from a southwestern Greenland lake. *Quaternary Research*, **81**, 193–202.
- Miyazono A, Nagai S, Kudo I, Tanizawa K (2012) Viability of *Alexandrium tamarense* cysts in the sediment of Funka Bay, Hokkaido, Japan: Over a hundred year survival times for cysts. *Harmful Algae*, **16**, 81–88.
- Mooij WM, Hülsmann S, De Senerpont Domis LN *et al.* (2005) The impact of climate change on lakes in the Netherlands: A review. *Aquatic Ecology*, **39**, 381–400.
- 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**, 255–266.
- Morton PK, Frisch D, Jeyasingh PD, Weider LJ (2015) Out with the old, in with the new? Younger *Daphnia* clones are competitively superior over centuries-old ancestors. *Hydrobiologia*, **749**, 43–52.
- Moss B (2012) Cogs in the endless machine: Lakes, climate change and nutrient cycles: A review. *Science of the Total Environment*, **434**, 130–142.

- Moss B, Mckee D, Atkinson D *et al.* (2003) How important is climate? Effects of warming, nutrient addition and fish on phytoplankton in shallow lake microcosms. *Journal of Applied Ecology*, **40**, 782–792.
- Möst M, Oexle S, Marková S *et al.* (2015) Population genetic dynamics of an invasion reconstructed from the sediment egg bank. *Molecular Ecology*, **24**, 4074–4093.
- Müller H (1989) The relative importance of different ciliate taxa in the pelagic food web of lake constance. *Microbial Ecology*, **18**, 261–273.
- Müller H, Schöne A, Pinto-Coelho RM, Schweizer A, Weisse T (1991) Seasonal succession of ciliates in lake constance. *Microbial Ecology*, **21**, 119–138.
- Munn C (2011) *Marine Microbiology: Ecology and Applications*. Taylor & Francis Group.
- Naselli-Flores L, Barone R (2000) Phytoplankton dynamics and structure: a comparative analysis in natural and man-made water bodies of different trophic state. *Hydrobiologia*, **438**, 65–74.
- Nelson DM, Ohene-Adjei S, Hu FS, Cann IKO, Mackie RI (2007) Bacterial Diversity and Distribution in the Holocene Sediments of a Northern Temperate Lake. *Microbial Ecology*, **54**, 252–263.
- Nemergut DR, Shade A, Violle C (2014) When, where and how does microbial community composition matter? *Frontiers in Microbiology*, **5**, 497.
- Nesme J, Achouak W, Agathos SN *et al.* (2016) Back to the future of soil metagenomics. *Frontiers in Microbiology*, **7**, 1–5.
- Neubauer TA, Harzhauser M, Mandic O, Georgopoulou E, Kroh A (2016) Paleobiogeography and historical biogeography of the non-marine caenogastropod family Melanopsidae. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **444**, 124–143.
- Newbold T, Hudson LN, Arnell AP *et al.* (2016) Has land use pushed terrestrial biodiversity beyond the planetary boundary? A global assessment. *Science*, **353**, 288–91.
- Nguyen TNM, Berzano M, Gualerzi CO, Spurio R (2011) Development of molecular tools for the detection of freshwater diatoms. *Journal of Microbiological Methods*, **84**, 33–40.
- Nogales B, Lanfranconi MP, Piña-Villalonga JM, Bosch R (2011) Anthropogenic perturbations in marine microbial communities. *FEMS Microbiology Reviews*, **35**, 275–298.
- Nolte V, Pandey RV, Jost S *et al.* (2010) Contrasting seasonal niche separation between rare and abundant taxa conceals the extent of protist diversity. *Molecular Ecology*, **19**, 2908–15.
- Oikonomou A, Filker S, Breiner H-W, Stoeck T (2015) Protistan diversity in a permanently stratified meromictic lake (Lake Alatzee, SW Germany). *Environmental Microbiology*, **17**, 2144–2157.

- Oksanen AJ, Blanchet FG, Kindt R *et al.* (2015) The vegan package. Community ecology package.
- Orsi W, Biddle JF, Edgcomb V (2013) Deep sequencing of subseafloor eukaryotic rRNA reveals active Fungi across marine subsurface provinces. *PLoS One*, **8**, e56335.
- Ottesen E a., Young CR, Gifford SM *et al.* (2014) Multispecies diel transcriptional oscillations in open ocean heterotrophic bacterial assemblages. *Science*, **345**, 207–212.
- Pace ML (1982) Planktonic Ciliates: Their Distribution, Abundance, and Relationship to Microbial Resources in a Monomictic Lake. *Canadian Journal of Fisheries and Aquatic Sciences*, **39**, 1106–1116.
- Paerl HW (2006) Assessing and managing nutrient-enhanced eutrophication in estuarine and coastal waters: Interactive effects of human and climatic perturbations. *Ecological Engineering*, **26**, 40–54.
- Paerl HW, Huisman J (2008) Climate. Blooms like it hot. *Science*, **320**, 57–8.
- Paerl HW, Xu H, McCarthy MJ *et al.* (2011) Controlling harmful cyanobacterial blooms in a hyper-eutrophic lake (Lake Taihu, China): The need for a dual nutrient (N & P) management strategy. *Water Research*, **45**, 1973–1983.
- Pal S, Gregory-Eaves I, Pick FR (2015) Temporal trends in cyanobacteria revealed through DNA and pigment analyses of temperate lake sediment cores. *Journal of Paleolimnology*, **54**, 87–101.
- Palm F, Stenson J, Lagergren R (2005) Which paleolimnological zooplankton records can indicate changes in planktivorous fish predation? *Verh Internat Verein Limnol*, **29**, 661–666.
- Palm F, Svensson J-E (2010) Subfossil <i>Chaoborus</i> mandibles confirm historical fish decline in two acidified lakes, SW Sweden. *Fundamental and Applied Limnology / Archiv für Hydrobiologie*, **177**, 313–320.
- Pansu J, Giguet-Covex C, Ficetola GF *et al.* (2015) Reconstructing long-term human impacts on plant communities: An ecological approach based on lake sediment DNA. *Molecular Ecology*, **24**, 1485–1498.
- Paradis E, Claude J, Strimmer K (2004) APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics (Oxford, England)*, **20**, 289–90.
- Pascoal C, Cássio F, Marcotegui A, Sanz B, Gomes P (2005) Role of fungi, bacteria, and invertebrates in leaf litter breakdown in a polluted river. *Journal of the North American Benthological Society*, **24**, 784–797.
- Pawlowski J (2013) The new micro-kingdoms of eukaryotes. *BMC biology*, **11**, 40.
- Pawlowski J, Audic S, Adl SS *et al.* (2012) CBOL protist working group: barcoding eukaryotic richness beyond the animal, plant, and fungal kingdoms. *PLoS biology*, **10**,

e1001419.

- Pawlowski J, Christen R, Lecroq B *et al.* (2011) Eukaryotic Richness in the Abyss: Insights from Pyrotag Sequencing (J Gilbert, Ed.). *PLoS ONE*, **6**, e18169.
- Pawlowski J, Lejzerowicz F, Apotheloz-Perret-Gentil L, Visco J, Esling P (2016) Protist metabarcoding and environmental biomonitoring: Time for change. *European Journal of Protistology*, In press.
- Pedersen MW, Overballe-Petersen SS, Ermini L *et al.* (2015) Ancient and modern environmental DNA. *Philosophical Transactions of the Royal Society Biological Sciences*, **370**, 20130383–20130383.
- Pedrós-Alió C (2006) Marine microbial diversity: can it be determined? *Trends in Microbiology*, **14**, 257–63.
- Pedrós-Alió C (2012) The Rare Bacterial Biosphere. *Annual Review of Marine Science*, **4**, 449–466.
- Van de Peer Y, Rijk P De, Wuyts J, Winkelmans T, De Wachter R (2000) The European Small Subunit Ribosomal RNA database. *Nucleic Acids Research*, **28**, 175–176.
- Peres-Neto P, Jackson D (2001) How well do multivariate data sets match? The advantages of a Procrustean superimposition approach over the Mantel test. *Oecologia*, **129**, 169–178.
- Perga M-E, Frossard V, Jenny J-P *et al.* (2015) High-resolution paleolimnology opens new management perspectives for lakes adaptation to climate warming. *Frontiers in Ecology and Evolution*, **3**, 1–17.
- Perren BB, Massa C, Bichet V *et al.* (2012) A paleoecological perspective on 1450 years of human impacts from a lake in southern Greenland. *The Holocene*, **22**, 1025–1034.
- Perry JJ, Staley JT, Lory S (2004) *Microbiologie : cours et questions de révision*. Dunod.
- Piscia R, Colombini M, Ponti B *et al.* (2015) Lifetime Response of Contemporary Versus Resurrected *Daphnia galeata* Sars (Crustacea, Cladocera) to Cu(II) Chronic Exposure. *Bulletin of Environmental Contamination and Toxicology*, **94**, 46–51.
- Poinar H, Cooper A (2000) Ancient DNA: do it right or not at all. *Science*.
- Pollard HG, Colbourne JK, Keller W (2003) Reconstruction of centuries-old *Daphnia* communities in a lake recovering from acidification and metal contamination. *Ambio*, **32**, 214–8.
- Pomeroy LR (1974) The Ocean's Food Web, A changing Paradigm. *BioScience*, **24**, 499–504.
- Posch T, Eugster B, Pomati F *et al.* (2015) Network of Interactions Between Ciliates and Phytoplankton During Spring. *Frontiers in Microbiology*, **6**, 1289.
- Posch T, Köster O, Salcher MM, Pernthaler J (2012) Harmful filamentous cyanobacteria favoured by reduced water turnover with lake warming. *Nature Climate Change*, **2**, 809–

- Price MN, Dehal PS, Arkin AP (2010) FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS One*, **5**, e9490.
- Props R, Kerckhof F-M, Rubbens P *et al.* (2016) Absolute quantification of microbial taxon abundances. *The ISME Journal*, 1–4.
- Prosser JI, Bohannan BJM, Curtis TP *et al.* (2007) The role of ecological theory in microbial ecology. *Nature Reviews Microbiology*, **5**, 384–92.
- Pruesse E, Quast C, Knittel K *et al.* (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Research*, **35**, 7188–96.
- Pyro VS, Roesch LFW, Morais DK *et al.* (2014) Data analysis for 16S microbial profiling from different benchtop sequencing platforms. *Journal of Microbiological Methods*, **107**, 30–37.
- Raghukumar C, Raghukumar S, Sheelu G, Gupta SM (2004) Buried in time : culturable fungi in a deep-sea sediment core from the Chagos Trench , Indian Ocean. *Deep Sea Research Part I*, **51**, 1759–1768.
- Ramette A (2007) Multivariate analyses in microbial ecology. *FEMS Microbiology Ecology*, **62**, 142–60.
- Randlett M-È, Coolen MJL, Stockhecke M *et al.* (2014) Alkenone distribution in Lake Van sediment over the last 270 ka: influence of temperature and haptophyte species composition. *Quaternary Science Reviews*, **104**, 53–62.
- Randsalu-Wendrup L, Conley DJ, Carstensen J, Fritz SC (2016) Paleolimnological records of regime shifts in lakes in response to climate change and anthropogenic activities. *Journal of Paleolimnology*, **56**, 1–14.
- Rasconi S, Niquil N, Sime-Ngando T (2012) Phytoplankton chytridiomycosis: Community structure and infectivity of fungal parasites in aquatic ecosystems. *Environmental Microbiology*, **14**, 2151–2170.
- Renberg I (1981) Improved methods for sampling, photographing and varve-counting of varved lake sediments. *Boreas*, **10**, 255–258.
- Renberg I (1986) Photographic demonstration of the annual nature of a varve type common in N. Swedish lake sediments. *Hydrobiologia*, **140**, 93–95.
- de Reviers B (2002) *Biologie et phylogénie des algues. Tome 1*. Belin.
- Reynolds CS, Padisák J, Sommer U (1993) Intermediate disturbance in the ecology of phytoplankton and the maintenance of species diversity: a synthesis. *Hydrobiologia*, **249**, 183–188.
- Richards T a, Vepritskiy A a, Gouliamova DE, Nierzwicki-Bauer S a (2005) The molecular

- diversity of freshwater picoeukaryotes from an oligotrophic lake reveals diverse, distinctive and globally dispersed lineages. *Environmental microbiology*, **7**, 1413–25.
- Rigosi A, Carey CC, Ibelings BW, Brookes JD (2014) The interaction between climate warming and eutrophication to promote cyanobacteria is dependent on trophic state and varies among taxa. *Limnology and Oceanography*, **59**, 99–114.
- Rizzi E, Lari M, Gigli E, De Bellis G, Caramelli D (2012) Ancient DNA studies: new perspectives on old samples. *Genetics, selection, evolution : GSE*, **44**, 21.
- Rockström J, Steffen W, Noone K *et al.* (2009) Planetary boundaries: Exploring the safe operating space for humanity. *Ecology and Society*, **14**.
- Rogalski MA (2015) Tainted resurrection: metal pollution is linked with reduced hatching and high juvenile mortality in *Daphnia* egg banks. *Ecology*, **96**, 1166–1173.
- Rose JM, Feng Y, Gobler CJ *et al.* (2009) Effects of increased pCO<sub>2</sub> and temperature on the North Atlantic spring bloom. II. Microzooplankton abundance and grazing. *Marine Ecology Progress Series*, **388**, 27–40.
- Rubin BER, Sanders JG, Hampton-Marcell J *et al.* (2014) DNA extraction protocols cause differences in 16S rRNA amplicon sequencing efficiency but not in community profile composition or structure. *MicrobiologyOpen*, **3**, 910–21.
- Rull V (2014) Time continuum and true long-term ecology: from theory to practice. *Frontiers in Ecology and Evolution*, **2**, 75.
- Rydberg J, Gälman V, Renberg I *et al.* (2008) Assessing the Stability of Mercury and Methylmercury in a Varved Lake Sediment Deposit. *Environmental Science & Technology*, **42**, 4391–4396.
- Rydberg J, Martinez-Cortizas A (2014) Geochemical assessment of an annually laminated lake sediment record from northern Sweden: A multi-core, multi-element approach. *Journal of Paleolimnology*, **51**, 499–514.
- Savichtcheva O, Debroas D, Kurmayer R *et al.* (2011) Quantitative PCR enumeration of total/toxic *Planktothrix rubescens* and total cyanobacteria in preserved DNA isolated from lake sediments. *Applied and environmental microbiology*, **77**, 8744–53.
- Savichtcheva O, Debroas D, Perga M-EE *et al.* (2015) Effects of nutrients and warming on *Planktothrix* dynamics and diversity: a palaeolimnological view based on sedimentary DNA and RNA. *Freshwater Biology*, **60**, 31–49.
- Schindler DW (2006) Recent advances in the understanding and management of eutrophication. *Limnology and Oceanography*, **51**, 356–363.
- Schloss PD, Handelsman J (2005) Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. *Applied and environmental microbiology*, **71**, 1501–6.



- Schmidt TSB, Matias Rodrigues JF, von Mering C (2014) Ecological Consistency of SSU rRNA-Based Operational Taxonomic Units at a Global Scale. *PLoS Computational Biology*, **10**, e1003594.
- Scupham AJ, Presley LL, Wei B *et al.* (2006) Abundant and diverse fungal microbiota in the murine intestine. *Applied and environmental microbiology*, **72**, 793–801.
- Seddon AWR, Mackay AW, Baker AG *et al.* (2014) Looking forward through the past: Identification of 50 priority research questions in palaeoecology. *Journal of Ecology*, **102**, 256–267.
- De Senerpont Domis LN, Elser JJ, Gsell AS *et al.* (2013) Plankton dynamics under different climatic conditions in space and time. *Freshwater Biology*, **58**, 463–482.
- De Senerpont Domis LN, Van De Waal DB, Helmsing NR, Van Donk E, Mooij WM (2014) Community stoichiometry in a changing world: Combined effects of warming and eutrophication on phytoplankton dynamics. *Ecology*, **95**, 1485–1495.
- Shade A (2016) Diversity is the question, not the answer. *The ISME Journal*, 1–6.
- Shade A, Caporaso JG, Handelsman J, Knight R, Fierer N (2013) A meta-analysis of changes in bacterial and archaeal communities with time. *The ISME Journal*, **7**, 1493–506.
- Shade A, Gilbert JA (2015) Temporal patterns of rarity provide a more complete view of microbial diversity. *Trends in Microbiology*, 1–6.
- Shade A, Jones SE, Caporaso JG *et al.* (2014) Conditionally Rare Taxa Disproportionately Contribute to Temporal Changes in Microbial Diversity. *mBio*, **5**, 1–9.
- Shade A, Peter H, Allison SD *et al.* (2012) Fundamentals of microbial community resistance and resilience. *Frontiers in Microbiology*, **3**, 417.
- Sherr EB, Sherr BF (2002) Significance of predation by protists in aquatic microbial food webs. *Antonie van Leeuwenhoek*, **81**, 293–308.
- Sherr E, Sherr B (2007) Heterotrophic dinoflagellates: a significant component of microzooplankton biomass and major grazers of diatoms in the sea. *Marine Ecology Progress Series*, **352**, 187–197.
- Shimoda Y, Azim ME, Perhar G *et al.* (2011) Our current understanding of lake ecosystem response to climate change: What have we really learned from the north temperate deep lakes? *Journal of Great Lakes Research*, **37**, 173–193.
- Sicre MA, Jacob J, Ezat U *et al.* (2008) Decadal variability of sea surface temperatures off North Iceland over the last 2000 years. *Earth and Planetary Science Letters*, **268**, 137–142.
- Šimek K, Kasalický V, Jezbera J *et al.* (2013) Differential freshwater flagellate community response to bacterial food quality with a focus on Limnohabitans bacteria. *The ISME Journal*, **7**, 1519–30.

- Simek K, Nedoma J, Znachor P *et al.* (2014) A finely tuned symphony of factors modulates the microbial food web of a freshwater reservoir in spring. *Limnology and Oceanography*, **59**, 1477–1492.
- Simon M, López-García P, Deschamps P *et al.* (2015) Marked seasonality and high spatial variability of protist communities in shallow freshwater systems. *The ISME Journal*, **9**, 1941–1953.
- Simpson G, Anderson N (2009) Deciphering the effect of climate change and separating the influence of confounding factors in sediment core records using additive models. *Limnology and Oceanography*, **54**, 2529–2541.
- Sjöstedt J, Koch-schmidt P, Pontarp M *et al.* (2012) Recruitment of Members from the Rare Biosphere of Marine Bacterioplankton Communities after an Environmental Disturbance. *Applied and Environmental Microbiology*, 1361–1369.
- Smith VH (2003) Eutrophication of Freshwater and Coastal Marine Ecosystems A Global Problem. *Environmental Science and Pollution Research*, **10**, 126–139.
- Smol JP (1992) Paleolimnology: an important tool for effective ecosystem management. *Journal of Aquatic Ecosystem Health*, **1**, 49–58.
- Smol J (2008) *Pollution of lakes and rivers a paleoenvironmental perspective*. Blackwell Pub.
- Smol J (2010) *The Diatoms: Applications for the Environmental and Earth Sciences- 2nd edition*. Cambridge University Press, Cambridge.
- Smol JP, Cumming BF (2000) Review tracking long-term changes in climate using algal indicators in lake sediments. *Journal of Phycology*, **1011**, 986–1011.
- Smol JP, Wolfe AP, Birks HJB *et al.* (2005) Climate-driven regime shifts in the biological communities of arctic lakes. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 4397–402.
- Soares EM, Figueredo CC, Gücker B, Boëchat IG (2013) Effects of growth condition on succession patterns in tropical phytoplankton assemblages subjected to experimental eutrophication. *Journal of Plankton Research*, **35**, 1141–1153.
- Sogin ML, Morrison HG, Huber JA *et al.* (2006) Microbial diversity in the deep sea and the underexplored “rare biosphere”. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 12115–20.
- Sommer U, Gliwicz ZM, Lampert W, Duncan A (1986) The PEG-model of seasonal succession of planktonic events in fresh waters. *Archiv für Hydrobiologie*, **106**, 433–471.
- Sommer U, Padisák J, Reynolds CS, Juhász-Nagy P (1993) Hutchinson’s heritage: the diversity-disturbance relationship in phytoplankton. *Hydrobiologia*, **249**, 1–7.
- Sondergaard M, Jeppesen E, Lauridsen TL *et al.* (2007) Lake restoration: Successes, failures and long-term effects. *Journal of Applied Ecology*, **44**, 1095–1105.

- Sparrow F (2013) Ecology of freshwater fungi. *The fungi*.
- Stager JC, Sporn LA, Johnson M *et al.* (2015) Of Paleo-Genes and Perch: What if an “Alien” Is Actually a Native? (H Wang, Ed.). *PLOS ONE*, **10**, e0119071.
- Steele JA, Countway PD, Xia L *et al.* (2011) Marine bacterial, archaeal and protistan association networks reveal ecological linkages. *The ISME Journal*, **5**, 1414–25.
- Steffen W, Richardson K, Rockström J *et al.* (2015) Planetary boundaries: Guiding human development on a changing planet. *Science*, **348**, 1217.
- Stelzer RS, Heffernan J, Likens GE (2003) The influence of dissolved nutrients and particulate organic matter quality on microbial respiration and biomass in a forest stream. *Freshwater Biology*, **48**, 1925–1937.
- Stoof-Leichsenring KR, Epp LS, Trauth MH, Tiedemann R (2012) Hidden diversity in diatoms of Kenyan Lake Naivasha: a genetic approach detects temporal variation. *Molecular Ecology*, **21**, 1918–30.
- Stoof-Leichsenring KR, Herzsuh U, Pestryakova L *et al.* (2015) Genetic data from algae sedimentary DNA reflect the influence of environment over geography. *Scientific Reports*, **5**, 12924.
- Strickland MS, Lauber C, Fierer N, Bradford MA (2009) Testing the functional significance of microbial community composition. *Ecology*, **90**, 441–451.
- Sunagawa S, Coelho LP, Chaffron S *et al.* (2015) Structure and function of the global ocean microbiome. *Science*, **348**, 1–10.
- Taberlet P, Prud’Homme SM, Campione E *et al.* (2012) Soil sampling and isolation of extracellular DNA from large amount of starting material suitable for metabarcoding studies. *Molecular Ecology*, **21**, 1816–1820.
- Taib N, Mangot J-F, Domaizon I, Bronner G, Debroas D (2013) Phylogenetic affiliation of SSU rRNA genes generated by massively parallel sequencing: new insights into the freshwater protist diversity. *PLoS One*, **8**, e58950.
- Takahara T, Minamoto T, Yamanaka H *et al.* (2012) Estimation of Fish Biomass Using Environmental DNA (JA Gilbert, Ed.). *PLoS ONE*, **7**, e35868.
- Takishita K, Tsuchiya M, Kawato M *et al.* (2007) Genetic Diversity of Microbial Eukaryotes in Anoxic Sediment of the Saline Meromictic Lake Namako-ike (Japan): On the Detection of Anaerobic or Anoxic-tolerant Lineages of Eukaryotes. *Protist*, **158**, 51–64.
- Tank JL, Rosi-Marshall EJ, Griffiths NA, Entekin SA, Stephen ML (2010) A review of allochthonous organic matter dynamics and metabolism in streams. *Journal of the North American Benthological Society*, **29**, 118–146.
- Taranu Z, Gregory-Eaves I, Leavitt P *et al.* (2015) Acceleration of cyanobacterial dominance in north temperate- subarctic lakes during the Anthropocene. *Ecology Letters*, **18**, 375–

- Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic acids research*, **22**, 4673–80.
- Thomsen PF, Willerslev E (2015) Environmental DNA - An emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation*, **183**, 4–18.
- Torti A, Alexander M, Barker B (2015) Marine Genomics Origin , dynamics, and implications of extracellular DNA pools in marine sediments. *Marine Genomics*, **24**, 185–196.
- Triadó-Margarit X, Casamayor EO (2012) Genetic diversity of planktonic eukaryotes in high mountain lakes (Central Pyrenees, Spain). *Environmental microbiology*, **14**, 2445–56.
- Triadó-Margarit X, Casamayor EO (2015) High protists diversity in the plankton of sulfurous lakes and lagoons examined by 18S rRNA gene sequence analyses. *Environmental Microbiology Reports*, **7**, 908–917.
- Tsugeki NK, Urabe J, Hayami Y, Kuwae M, Nakanishi M (2009) Phytoplankton dynamics in Lake Biwa during the 20th century: complex responses to climate variation and changes in nutrient status. *Journal of Paleolimnology*, **44**, 69–83.
- Turner CR, Uy KL, Everhart RC (2015) Fish environmental DNA is more concentrated in aquatic sediments than surface water. *Biological Conservation*, **183**, 93–102.
- Turon V, Trably E, Fayet A, Fouilland E, Steyer JP (2015) Raw dark fermentation effluent to support heterotrophic microalgae growth: Microalgae successfully outcompete bacteria for acetate. *Algal Research*, **12**, 119–125.
- Tveit AT, Urich T, Frenzel P, Svenning MM (2015) Metabolic and trophic interactions modulate methane production by Arctic peat microbiota in response to warming. *Proceedings of the National Academy of Sciences of the United States of America*, **112**, E2507–16.
- Unrein F, Gasol JM, Not F, Forn I, Massana R (2014) Mixotrophic haptophytes are key bacterial grazers in oligotrophic coastal waters. *The ISME Journal*, **8**, 164–76.
- Vandekerkhove J, Declerk S, Vanhove M *et al.* (2004) Use of ehippial morphology to assess richness of anomopods: potentials and pitfalls. *Journal of Limnology*, **63**, 75.
- VanWinkle-Swift KP, Rickoll WL (1997) The zygo-spore wall of *Chlamydomonas monoica* (Chlorophyceae): morphogenesis and evidence for the presence of sporopollenin. *Journal of Phycology*, **33**, 655–665.
- De Vargas C, Engelen S, Hingamp P *et al.* (2015) Eukaryotic plankton diversity in the sunlit ocean. *Science*, **348**, 1–12.
- Vergin KL, Beszteri B, Monier A *et al.* (2013) High-resolution SAR11 ecotype dynamics at

- the Bermuda Atlantic Time-series Study site by phylogenetic placement of pyrosequences. *The ISME Journal*, **7**, 1322–32.
- Vernal A de, Henry M, Matthiessen J *et al.* (2001) Dinoflagellate cyst assemblages as tracers of sea-surface conditions in the Northern North Atlantic, Arctic and sub-Arctic seas: The new “n = 677” data base and its application for quantitative palaeoceanographic reconstruction. *Journal of Quaternary Science*, **16**, 681–698.
- Vick-Majors TJ, Priscu JC, Amaral-Zettler L (2014) Modular community structure suggests metabolic plasticity during the transition to polar night in ice-covered Antarctic lakes. *The ISME Journal*, 1–12.
- Vincent WF, Laurion I, Pienitz R, Anthony KMW (2013) Climate Impacts on Arctic Lake Ecosystems. In: *Climatic Change and Global Warming of Inland Waters: Impacts and Mitigation for Ecosystems and Societies* (ed Wiley J)
- Vitousek PM (1997) Human Domination of Earth’s Ecosystems. *Science*, **277**, 494–499.
- Vuillemin A, Ariztegui D, Leavitt PR (2015) Recording of climate and diagenesis through fossil pigments and sedimentary DNA at Laguna Potrok Aike , Argentina.
- Wagner Mackenzie B, Waite DW, Taylor MW (2015) Evaluating variation in human gut microbiota profiles due to DNA extraction method and inter-subject differences. *Frontiers in Microbiology*, **6**, 130.
- Walther G-R, Post E, Convey P *et al.* (2002) Ecological responses to recent climate change. *Nature*, **416**, 389–95.
- Wesolowska-Andersen A, Bahl M, Carvalho V *et al.* (2014) Choice of bacterial DNA extraction method from fecal material influences community structure as evaluated by metagenomic analysis. *Microbiome*, **2**, 19.
- Wetzel RG (2001) *Limnology: Lake and River Ecosystems*.
- Wilhelm L, Besemer K, Fasching C *et al.* (2014) Rare but active taxa contribute to community dynamics of benthic biofilms in glacier-fed streams. *Environmental Microbiology*, **16**, 2514–24.
- Wilkinson DM, Koumoutsaris S, Mitchell EAD, Bey I (2012) Modelling the effect of size on the aerial dispersal of microorganisms. *Journal of Biogeography*, **39**, 89–97.
- Willerslev E, Cooper A (2005) Ancient DNA. *Proceedings of the Royal Society B*, **272**, 3–16.
- Williamson CE, Saros JE, Vincent WF, Smol JP (2009) Lakes and reservoirs as sentinels, integrators, and regulators of climate change. *Limnology and Oceanography*, **54**, 2273–2282.
- Willner D, Daly J, Whiley D *et al.* (2012) Comparison of DNA Extraction Methods for Microbial Community Profiling with an Application to Pediatric Bronchoalveolar Lavage Samples. *PLoS One*, **7**, e34605.

- de Wit R, Bouvier T (2006) “Everything is everywhere, but, the environment selects”; what did Baas Becking and Beijerinck really say? *Environmental Microbiology*, **8**, 755–758.
- Wittebolle L, Marzorati M, Clement L *et al.* (2009) Initial community evenness favours functionality under selective stress. *Nature*, **458**, 623–626.
- Wolf C, Iversen M, Klaas C, Metfies K (2016) Limited sinking of Phaeocystis during a 12 days sediment trap study. *Molecular Ecology*, **25**, 3428–3435.
- Wolinska J, Giessler S, Koerner H (2009) Molecular identification and hidden diversity of novel Daphnia parasites from European lakes. *Applied and Environmental microbiology*, **75**, 7051–9.
- Wood S (2015) Package “mgcv.”
- Worden AZ, Not F (2008) Ecology and diversity of picoeukaryotes. In: *Microbial Ecology of the Oceans*, pp. 159–205.
- Wurzbacher C, Bärlocher F, Grossart H (2010) Fungi in lake ecosystems. *Aquatic Microbial Ecology*, **59**, 125–149.
- Wurzbacher C, Fuchs A, Attermeyer K *et al.* (2016) Vertical organization of freshwater sediment microbial communities: implications for microbial activities and burial processes. *bioRxiv*.
- Xiao X, Sogge H, Lagesen K *et al.* (2014) Use of High Throughput Sequencing and Light Microscopy Show Contrasting Results in a Study of Phytoplankton Occurrence in a Freshwater Environment (C Lovejoy, Ed.). *PLoS One*, **9**, e106510.
- Xu Z-H, Jiang X-D, Wang G-Z *et al.* (2011a) DNA extraction, amplification and analysis of the 28S rRNA portion in sediment-buried copepod DNA in the Great Wall Bay and Xihu Lake, Antarctica. *Journal of Plankton Research*, **33**, 917–925.
- Xu Y, Vick-Majors T, Morgan-Kiss R, Priscu JC, Amaral-Zettler L (2014) Ciliate Diversity, Community Structure, and Novel Taxa in Lakes of the McMurdo Dry Valleys, Antarctica. *Biological Bulletin*, **227**, 175–190.
- Xu Z-H, Wang G-Z, Mu Q, Wu L-S, Li S-J (2011b) An approach to the study of copepod egg banks based on efficient DNA extraction from individual copepod eggs. *Marine Biology Research*, **7**, 592–598.
- Zhang H, Huang T, Chen S *et al.* (2015) Ignored sediment fungal populations in water supply reservoirs are revealed by quantitative PCR and 454 pyrosequencing. *BMC Microbiology*, **15**, 44.
- Zhao B, Chen M, Sun Y, Yang J, Chen F (2011) Genetic diversity of picoeukaryotes in eight lakes differing in trophic status. *Canadian Journal of Microbiology*, **126**, 115–126.
- Zhu F, Massana R, Not F, Marie D, Vaultot D (2005) Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. *FEMS Microbiology Ecology*,

**52**, 79–92.

Zinger L, Gobet A, Pommier T (2012) Two decades of describing the unseen majority of aquatic microbial diversity. *Molecular ecology*, **21**, 1878–96.

Zonneveld KAF, Versteegh GJM, de Lange GJ (1997) Preservation of organic-walled dinoflagellate cysts in different oxygen regimes: a 10,000 year natural experiment. *Marine Micropaleontology*, **29**, 393–405.

# Resume

## Education & Positions

---

- 2013-2016 **PhD in paleo-microbial ecology**  
INRA UMR CARRTEL, Thonon-les-bains, France  
*Past microbial diversity dynamics in lacustrine systems*
- 2011-2013 **Master of Science in Microbial Ecology**, University Lyon 1, France
- 2009-2011 **Bachelor of Science in Microbiology**, University Lyon 1, France
- 2007-2009 **HND in Biological and Biochemical Analyzes**, IUT Tours, France
- 2004-2006 **Baccalauréat scientifique**, High School, Ingré, France

## Scientific skills

---

<b>Molecular Biology</b>	DNA extraction, PCR, clone libraries, primers design, qPCR amplicons preparation (454, MiSeq), fingerprinting (RISA, RFLP)
<b>Molecular data analysis</b>	$\alpha$ , $\beta$ , $\gamma$ diversity, network topology, temporal ecology
<b>Statistics &amp; data analysis tools R packages</b>	R software, Seaview, MEGA, BLASTn, Gephi, Cytoscape vegan, ade4, bcp, mvpart, FactoMineR, strucchange...
<b>Microbiology</b>	Microbial culture, species identification, API Gallery
<b>Chemical analysis</b>	CG-MS, SPIR, hydrocarbons and lipids extraction
<b>Languages</b>	French (Native Language) Scientific English (Good skills to speak, read and write)
<b>Mentoring</b>	1 Undergraduate student (Leandre Vasseur)

## Internships

---

- August 2014 **UNESCO International Summer School** in Limnology, Evian-les-bains, France  
*Lake ecology and temporal dynamics*
- July 2013 **Summer School** at LOV UPMC, Villefranche-sur-mer, France  
*Marine plankton diversity*
- June-July 2013 **Volunteer work** at ISARA, Lyon, France (with [J Robin](#))  
*Phytoplankton dynamics in ponds*
- Jan-May 2013 **Master 2 Internship**, EMG UMR 5005 Ecully, France (with [TM Vogel](#) & [C Larose](#))  
*Microbial degradation of hydrocarbons in snow*
- Apr-Aug 2012 **Volunteer work** in "Société des Amis de Navarrosse", Biscarrosse, France  
*Cyanobacterial diversity of coastal lakes*
- July 2012 **Summer school** at Marine Station of Arcachon, France  
*Plankton and macrofauna ecology*
- Jan-Apr 2012 **Master 1 Internship**, MAP UMR 5240, Villeurbanne, France (with [MA Mandrand](#))  
*Effect of an atpase gene, isolated from metagenomic DNA of mining spoils in New-Caledonia, on the resistance to metals*
- June-July 2011 **Internship**, IMBE UMR 6116, Marseille, France (with [C Charpy-Roubaud](#))  
*Evaluation of the bactericidal effect of *Ludwigia* spp.*



## Papers

---

- *Accepted*

**Capo E.**, Debroas D., Arnaud F., Guillemot T., Bichet V., Millet L., Gauthier E., Massa C., Develle A-L., Pignol C., Lejzerowicz F., Domaizon I (In press). Long-term dynamics in microbial eukaryotes communities: a paleolimnological view based on sedimentary DNA. *Molecular Ecology*

**Capo E.**, Debroas D., Arnaud F., Domaizon I (2015). Is Planktonic diversity well recorded in sedimentary DNA? Toward the reconstruction of past protistan diversity. *Microbial Ecology*, 70(4):865-75

- *In revision*

Domaizon I., Winegardner A., **Capo E.**, Gauthier J., Gregory-Eaves I., Application of DNA-based methods in paleolimnology: new opportunities for investigating long-term dynamics of lacustrine biodiversity; submitted in *Journal of Paleolimnology*

- *In preparation*

**Capo E.**, Debroas D., Arnaud F., Perga M-E., Chardon C., Domaizon I. A century of change in lacustrine microbial eukaryotes driven by eutrophication and climate warming: a paleoecological approach based on the high throughput sequencing of sedimentary DNA. In prep

**Capo E.**, Bigler C., Debroas D., Domaizon I. Preservation of microbial eukaryotic DNA signal during burying into lake sediments: case of Lake Nylandssjön (Sweden) over a 40 years time series. In prep

Winegardner A., **Capo E.**, Domaizon I., Debroas D., Hajibabei M., Shokralla S., White L., Wing B., Beissner B., Gregory-Eaves. Eukaryotic biodiversity dynamics during the Anthropocene from a mining region: an exploration using NGS. In prep

Figueroa D., **Capo E.**, Lindh M., Rowe O., Paczkowska J., Pinhassi J., Andersson A., Coupling between bacterial community composition and allochthonous organic matter in a sub-arctic estuary. In prep

## Other written contributions

---

- *Chapter of book*

Domaizon I., Arnaud F., **Capo E.**, Debroas D., (2015) L'ADN des archives sédimentaires : une opportunité unique de retracer les changements de biodiversité lacustre. In book: De la reconstitution des paysages à l'histoire des sociétés - 10000 ans d'archives sédimentaires en zones humides, Editors: Eds, Ph Barral, M Magny, M Thiviet

- *Chapter of thesis*

Winegardner A., **Capo E.**, Domaizon I., Debroas D., Hajibabei M., Shokralla S., White L., Wing B., Beissner B., Gregory-Eaves. Eukaryotic biodiversity dynamics during the Anthropocene from a mining region: an exploration using NGS. PhD Thesis of Amanda Winegardner. In thesis Aquatic biodiversity patterns along gradients of multiple stressors and disturbance histories: Integration of paleoecology and molecular techniques. 2016. McGill University, Department of Biology, Gregory-Eaves lab, Montréal, Canada

Figueroa D., **Capo E.**, Lindh M., Rowe O., Paczkowska J., Pinhassi J., Andersson A., Coupling between bacterial community composition and allochthonous organic matter in a sub-arctic estuary. PhD Thesis of Daniela Figueroa. In thesis "Bacterioplankton in the Baltic Sea: influence of allochthonous organic matter and salinity". 2016. Umeå University, Faculty of Science and Technology, Department of Ecology and Environmental Sciences. Umeå, Sweden

## Oral presentations

---

**Capo E.**, Debroas D., Arnaud F., Perga M-E., Chardon C., Domaizon I. Tracking a century of change in lacustrine microbial eukaryotes driven by local and global forcings. ISME16 Symposium. 21-26 august 2016. Montreal QC

Figueroa D., Lindh M., **Capo E.**, Lundin D., Pinhassi J., Andersson A. Bacterioplankton community response to allochthonous organic matter discharges in coastal areas of the Bothnian Bay. Swedish Society for Marine Sciences 18-20 november 2015, Lund, Sweden

Winegardner A., **Capo E.**, Domaizon I., Debroas D., Hajjibabei M., Shokralla S., White L., Wing B., Beisner B., Gregory-Eaves I. The use of Next-Generation Sequencing to characterize eukaryote biodiversity in lakes with historical metal pollution. Quebec Centre for Biodiversity Science Annual Symposium. 29-30 october 2015. Montreal QC

**Capo E.**, Arnaud F., Bichet V., Chardon C., Debroas D., Develle A-L, Guillemot T, Lejzerowicz F. & Domaizon I. Response of lacustrine protistan assemblages to climatic and anthropogenic pressures: a paleolimnological view based on sedimentary DNA. 6 July 2015, Symposium for European Fresh Water Sciences SEFS9, Geneva

**Capo E.**, Debroas D., Arnaud F. & Domaizon I. Paleo-ecological approach of lacustrine microbial diversity. 18 February 2015, Seminar at Department of Ecology and Environmental Science, University of Umeå, Sweden

Domaizon I., Debroas D., Bronner G., Taïb N., Lepère C., Mangot JF. , Savichtcheva O., **Capo E.**, Lyautey E., Chardon C., Jacas L., Arnaud F., Pignol C., Perga ME, Gregory Eaves I. Application of DNA-based methods for investigating modern and past diversity of planktonic taxa in lakes : Planktonic biodiversity and its forcing factors. In: DNAWatch2013, 4-5 December 2013, Frasnes, France

## Posters

---

**Capo E.**, Debroas D., Arnaud F., Keck F. & Domaizon I., Is planktonic diversity well recorded in sedimentary DNA ? Toward the reconstruction of past protistan diversity. In: Journée des doctorants ED SISEO, Annecy, 4 décembre 2014

**Capo E.**, Debroas D., Arnaud F., Pignol C. Taib N. Bronner G. Domaizon I. Paleoecological study of lake microbial diversity : long term dynamic of protists and cyanobacteria diversity related to environmental factors. In : DNAWatch2013, 4-5 December 2013, Frasnes, France

Domaizon I., Savichtcheva O., **Capo E.**, Arnaud F., Debroas D., Villar C., Pignol C., Jenny J.P., Alric B., Perga M.E. Dynamique et diversité à long terme des cyanobactéries révélées par l'ADN préservé dans les archives sédimentaires du lac du Bourget. In : ArcheoPal2013, 14-16 October 2013, Frasnes, France

Demuez M., Zhao C., Fall S., Roussel F., Effantin G., **Capo E.**, Héry M., Rodrigue A., Navarro E. and Mandrand-Berthelot M.A. Role in nickel homeostasis of genes isolated from the bacterial metagenome of an ultramafic soil. In: Metals Homeostasis 5th International IMBG Meeting. 17-21 September 2012, Autrans, France

## Other contributions

---

Member of [AFL](#) (French association of Limnology) since 2015

Part of the Local Organizing Committee of [SEFS9](#) (Geneva 2015)

Part of the organizing board of PhD student day (doctoral School SISEO, 2015)

Contribution to the science fairs at INRA UMR CARTELE (Thonon-les-bains, 2015)

Assistant of [René Capo](#) in Aquitaine coastal protection

Creator of the blog [Environmental Microbiology & Astrobiology](#)