

MICROBIAL COMMUNITY ASSEMBLY AND  
BIOGEOGRAPHY IN THE PYRENEAN  
LACUSTRINE DISTRICT

**Rüdiger Ortiz Álvarez**

Per citar o enllaçar aquest document:  
Para citar o enlazar este documento:  
Use this url to cite or link to this publication:  
<http://hdl.handle.net/10803/663398>



<http://creativecommons.org/licenses/by-nc-nd/4.0/deed.ca>

Aquesta obra està subjecta a una llicència Creative Commons Reconeixement-NoComercial-SenseObraDerivada

Esta obra está bajo una licencia Creative Commons Reconocimiento-NoComercial-SinObraDerivada

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives licence



PHD THESIS

Microbial community assembly and  
biogeography in the Pyrenean  
lacustrine district

**Rüdiger Ortiz Álvarez**

2018

Doctoral program in Water Science and Technology

*Tutor UdG: Helena Guasch Padro*

*Supervised by Emilio O. Casamayor*

Centre d'estudis avançats de Blanes (CEAB-CSIC)



CENTRE D'ESTUDIS AVANÇATS DE BLANES

Thesis submitted in fulfillment of the requirements to obtain the doctoral  
degree at the University of Girona



This PhD thesis is presented under the compendium of the following publications:

#### CHAPTER 1

**Ortiz-Álvarez R**, Fierer N, De los Ríos A, Casamayor EO, Barberán A (2018) Consistent changes in the taxonomy and functional attributes of bacterial communities during primary succession. ISME-J. DOI: 10.1038/s41396-018-0076-2

#### CHAPTER 3

**Ortiz-Alvarez R**, Casamayor, E. (2016) High occurrence of Paecearchaeota and Woesearchaeota (Archaea superphylum DPANN) in the surface waters of oligotrophic high-altitude lakes. Environmental Microbiology Reports. DOI :10.1111/1758-2229.12370

#### CHAPTER 4

**Ortiz-Álvarez, R.**, Triadó-Margarit, X., Camarero, L., Casamayor, E.O. & Catalán, J. (2018) High planktonic diversity in mountain lakes contains similar richness of autotrophic, heterotrophic and parasitic eukaryotic lifeforms. Scientific reports. DOI:10.1038/s41598-018-22835-3

Besides these three publications, CHAPTER 2 and CHAPTER 5 correspond to two articles *in preparation* that will be submitted shortly to journals of high relevance within the Microbial Ecology area.

## LIST OF ACRONYMS

Vol: lake volume (m<sup>3</sup>)  
FOND: Lake depth (m)  
RWT: Water renewal time (days)  
ALC: Alkalinity ( $\mu\text{eq L}^{-1}$ )  
COND: Conductivity ( $\mu\text{S cm}^{-1}$ )  
DIC: Dissolved inorganic carbon (mg L<sup>-1</sup>)  
DOC: Dissolved organic carbon (mg L<sup>-1</sup>)  
SRP: Reactive species of phosphorus ( $\mu\text{g L}^{-1}$ )  
TDP: Total dissolved phosphorus ( $\mu\text{g L}^{-1}$ )  
TP: Total phosphorus ( $\mu\text{g L}^{-1}$ )  
DOP: Dissolved organic phosphorus ( $\mu\text{g L}^{-1}$ )  
PP: Particulated phosphorus ( $\mu\text{g L}^{-1}$ )  
NH<sub>4</sub>: Ammonium ( $\mu\text{eq L}^{-1}$ )  
NO<sub>2</sub>: Nitrites ( $\mu\text{eq L}^{-1}$ )  
NO<sub>3</sub>: Nitrates ( $\mu\text{eq L}^{-1}$ )  
TDN: Total dissolved nitrogen (mg L<sup>-1</sup>)  
TN: Total nitrogen ( $\mu\text{g L}^{-1}$ )  
DON: Dissolved organic nitrogen ( $\mu\text{g L}^{-1}$ )  
PN: Particulate nitrogen ( $\mu\text{eq L}^{-1}$ )  
DRSi: Dissolved reactive silicium (mg L<sup>-1</sup>)  
Cl: Chloride ( $\mu\text{eq L}^{-1}$ )  
SO<sub>4</sub>: Sulphates ( $\mu\text{eq L}^{-1}$ )  
Na: Sodium ( $\mu\text{eq L}^{-1}$ )  
K: Potassium ( $\mu\text{eq L}^{-1}$ )  
Ca: Calcium ( $\mu\text{eq L}^{-1}$ )  
Mg: Magnesium ( $\mu\text{eq L}^{-1}$ )  
rRNA: ribosomal ribonucleic acid  
PCR: polymerase chain reaction  
IMG: integred microbial database  
KEGG: Kyoto encyclopedia of genes and genomes  
Pst: high efficiency phosphate assimilation operon  
H': Shannon diversity  
PD: Phylogenetic diversity  
nMDS: non-metric multidimensional scaling  
PCA: principal components analysis  
RC: rotated component  
PC: principal component  
EnvO: environmental ontology  
OTU: operational taxonomic unit  
zOTU: zero-radius operational taxonomic unit

WS: Watts-Strogatz

ER: Erdos-Renyi

AIC: Akaike information criterium

AOA: Ammonia-oxidation archaea

HNF: heterotrophic nanoflagellates

## LIST OF FIGURES

### GENERAL INTRODUCTION

**Figure 1.** Lakes from the Pyrenees mountain range: an alpine biome. The photographs show (a) Rovinets, (b) Xic de Subenuix and (c) Baborte, during summer of 2015. (p12)

**Figure 2.** Number of limnological papers published per year through searches in 'web of science' (WoS) according to different keywords: high altitude (alpine or sub-alpine), low altitude (lowland or montane), prokaryotes (Bacteria, Archaea or Bacterioplankton) and eukaryotes (Protists, Algae or Phytoplankton). (p13)

**Figure 3.** Summary of processes that potentially drive community assembly and ecological succession, therefore, diversity. (p15)

**Figure 4.** (a) Mountains viewed under the lens of primary succession: microbial colonization through air and transport through habitat connectivity downstream. (b) Seasonal processes in high mountain lakes (Catalan *et al.*, 2006): ice cover formation, growth and spring ablation linked to water fluxes and particle movements. (p18)

### GENERAL METHODS

**Figure 1.** Location of sampled lakes across the Pyrenees mountain range. Lakes are colored by the analysis performed and results retrieved. (p31)

**Figure 2.** Principal components analysis (PCA) performed with the 30 environmental variables (log-transformed and standardized) and the geographic coordinates of 320 lakes. The diagram describes the lineal combination of variables showing the main environmental gradients (components PC1 and PC2 represented 37% of the total environmental variance. With the third component, PC3, the analysis accumulates 49% of the environmental variance). The different colors indicate the geographic position. (p33)

**Figure 3.** Rarefaction curves of lake samples for eukarya (upper panel) and bacteria (lower panel). (p34)

### CHAPTER 1

**Figure 1.** Comparison of early and late stage successional samples in non-

metric multidimensional scaling ordinations based on Bray-Curtis dissimilarities of taxonomic (a), and functional (b) matrices. (p42)

**Figure 2.** Changes between late and early stages of diversity metrics. Taxonomic (a) and functional (c) Shannon index ( $H'$ ), and beta diversity based on Bray Curtis dissimilarities of taxonomy (b) and functions (d). A discontinuous line represents the overall mean change. The abbreviations 'Ph.' and 'H.' indicate the *Phormidium* river biofilm' and '*Hydra*' samples respectively. (p44)

**Figure 3.** Changes in rRNA operon copy number (a) and high efficient phosphate transport (b) community weighted functional traits. A discontinuous line represents the overall mean change. (p45)

## CHAPTER 2

**Figure 1.** The proportion of the dominant (a, left) and more rare (b, right) phyla in the Pyrenees sorted by the most abundant in each case. (p56)

**Figure 2.** Means and standard deviations of Richness and Shannon index for the main EnvO terms after the intensive SeqEnv database search (barplot panels). Community beta dispersion of subsampled communities (boxplot panel). All metrics showed significant differences with ANOVA, and therefore significance groups based on *post-hoc* HSD test are indicated. (p57)

**Figure 3.** Environmental term affiliation of zOTUs within most abundant phyla, sorted by the 'Lake' EnvO term. (p57)

**Figure 4:** (a) Non-metric multidimensional scaling (nMDS) ordination of lake bacterial communities based on Bray-Curtis dissimilarities colored by the strongest correlated variable (pH) and dot-shapes indicating considering two Shannon index groups (above and below the median value) (b) Comparison of the beta dispersion between communities with high and low  $H'$  diversity, and classified in five groups according to pH (0.5 intervals). (p59)

**Figure 5:** Network of correlations between nMDS axis, environmental variables, environmental sources, diversity metrics, and RC components (gephi layout: Forceatlas2). Node colors indicate three different modules connected by links displayed by size (Spearman's  $r$ ) and color (red: negative, grey: positive). Node size represents the number of links of that node. The network was drawn with Gephi. (p60)



**Figure 6:** Relationships between geography and community features. (a) Altitude along the longitudinal gradient colored by pH. Dashed lines indicate western and eastern geographic sub-regions. (b) Community betadispersion per Pyrenean sub-regions. (c) Sliding means of z-OTUs alpha diversity (PD) (rollmeans function in the R package 'zoo') across the altitudinal gradient. (d) Sliding means of alpha diversity (H') starting from the geographic centroid separating the gradient to the east and to the west. Linear models and loess functions were used to draw the geographic trends observed. (p61)

### CHAPTER 3

**Figure 1.** nMDS ordination on genetic distances of the OTUs assigned to each archaeal lineage. A line connects each point in the two-dimensional space and the centroid. (p77)

**Figure 2.** High rank archaeal lineages comparisons after rarefaction to 35 sequences per group. A: Nucleotidic diversity vs % occurrence. B: Nucleotidic diversity vs. Mean scaled weighted variances of environmental preferences. Point size indicates number of OTUs. (p78)

**Figure 3.** Relationship between relative abundance of Paeearchaeota and Woesearchaeota and bacterial phylogenetic diversity (PD). Both axis are in logarithmic scale. (p78)

**Figure 4.** Percentage of coding DNA sequence (CDS%) in the closest genomes versus mean scaled variances of environmental preferences as an indicator of environmental/habitat breadth. Circle size indicates proportion of genes not identified in KEGG orthology. \*\*\*16S rRNA gene identity >95% with genome, \*\*16S rRNA gene identity <91.5%, \*16S rRNA gene identity <83%, see Table 1. (p79)

### CHAPTER 4

**Figure 1.** Distribution of 227 lakes surveyed and sequenced across the Pyrenees. The main river basins are indicated. Dots of nearby lakes may overlap. Generated with ArcGIS 10.1 (©ESRI). (p86)

**Figure 2.** OTU richness and cumulative relative abundance of the main groups found in the plankton of Pyrenean lakes in logarithmic scale. Colours indicate dominant nutrition type while shapes indicate their main habitats. Autotrophs are split into two groups: those purely osmotrophs and those also able of phagotrophic ingestion (mixotrophs). A fitted linear model is drawn. (p88)

**Figure 3.** Novelty of OTUs as indicated by their closest cultured and environmental matches identity values. The highest novelty area is indicated, and falls below the 97% thresholds. Color displays the main groups of study, leaving other OTUs of the high taxonomic ranks in black. Dot size indicates OTU occurrence in the dataset. (p89)

**Figure 4.** Rank abundance plot (log scale) of OTUs by taxonomic groups of interest. Cumulative sum of relative abundances determines OTU rank per group. (p90)

**Figure 5.** Rarefaction curves per taxonomic group of the Pyrenean pool (227 samples rarefied as a single pool) versus the Mediterranean Sea pool from the Tara Oceans project (11 samples per size fraction). Despite the variable number of samples, we equalized both pools by rarefying at equivalent sequencing depths per pool (see Supplementary information 2). (p91)

## CHAPTER 5

**Figure 1.** The proportion of co-occurrence and co-exclusion pairs estimated by the probabilistic method, separated by bacteria-bacteria, bacteria-eukarya and eukarya-eukarya type of pair, and explained by the environment and the geography following Blois *et al.*, 2014. Environmental variables were ranked based on the cumulative proportion of links not previously explained by any other environmental variable (black line). (p106)

**Figure 2.** The proportion of negative links through the complete network comparing between modules and within modules. (p108)

**Figure 3.** Histogram and frequency distribution of node weighted degree, calculated for the network of co-occurrences (a) and the network of co-exclusions (b). Distribution and mean value (vertical dashed line) was drawn for Bacteria (blue) and Eukarya (black). (p109)

**Figure 4.** The relationship between weighted degree and node mean abundances, separating nodes per modules of the co-occurrence network (all the positive interactions). Linear models are shown in the modules for which the relationship was significant. (p110)

**Figure 5.** Network topology (a) module profiles colored by ecological categories (EnvO terms and trophic strategies) (b) and degree distribution by ecological categories (c) for the three networks based on the domains interacting: Bacteria-Bacteria (1), Eukarya-Eukarya (2) and Bacteria-Eukarya (3). (p112)

**Figure 6.** Co-exclusion networks of (a) Bacteria-Bacteria, (b) Bacteria-Eukarya and (c) Eukarya-Eukarya. Node size indicates the weighted degree and black and white scale the authority property. (p114)

**Figure 7.** Examples of co-excluding organisms: the relationship between their proportions per lake and the number of interactions they present in each lake. **(a)** HgcI bacteria (zOTU 1) displays a decreasing upper limit (quantile 0.99) and a threshold after ten co-exclusions. **(b)** Chytridiomycota (zOTU 645) exhibits an increase in the lower limit (quantile 0.01) and upper limit (quantile 0.75) with a potential threshold between one and two interactions. (p116)

**Figure 8.** Fungi co-occurrence network colored by the estimated modules (10). Edge size indicates the weight of co-occurrences. (p117)

**Figure 9.** **(a)** Relationships between local diversity and the number of interactions, **(b)** number of bacterial co-occurrences per geographic location and **(c)** number of bacterial co-exclusions per altitude groups. (p118)

**Figure 10.** Mean regional abundances and regional occurrence between inferred interacting zOTUs (834 zOTUs with environment-independent interactions) or non-interacting zOTUs (83 zOTUs with no environmentindependent interactions). (p121)

**Figure 11.** Mean zOTU abundance grouped by the annotated dominant EnvO affinities, and the assigned modules (using the B-B positive network). (122)

## GENERAL DISCUSSION

**Figure 1.** Eukaryotic pyrenees metacommunity displayed with the main environmental drivers. Also, relationships with alpha diversity, and beta-diversity based on differential phosphate/nitrogen limitation. Betadispersion is based on diversity categories and in alkalinity (strongest env. correlation) quartiles. (p136)

**Figure 2.** Relationships of distance metrics of communities, immigration sources and water chemistry. Mantel tests with the Spearman method were carried out to assess the strength of links between communities and community descriptors (dispersal and environmental filtering). (p138)

## LIST OF TABLES

### GENERAL METHODS

**Table 1.** Environmental descriptors of the Pyrenean dataset. Mean values, standard deviation (SD), maximum and minimum values for 30 chemical and morphometric variables.

### CHAPTER 2

**Table 1.** Summary of principal components analysis with rotation 'varimax'. The weights of variables associated to the rotated components are shown, as the percentage of variation explained by each component. (p55)

**Table 2.** Relationships of alpha diversity ( $H'$ ) measured in each environmental source and the geographic context (measured as distance to the Pyrenees center and by regional clusters). Significance levels: \*\* ( $p < 0.01$ ), \* ( $p < 0.05$ ). (p62)

### CHAPTER 3

**Table 1.** Archaeal sequences in the 16S rRNA gene PCR amplicon mixture (prokaryote 'universal' 515f-806r primers pair), OTUs at 97% identity and number of lakes (% occurrence) where each lineage was detected. (p77)

### CHAPTER 4

**Table 1.** Regional (lacustrine district) and local (per lake) diversity indexes. The mean closest match identity and OTU occurrence, Alpha and Gamma diversities estimated by the total number of OTUs (Richness, OTU Nr), Shannon diversity index ( $H'$ ), and Berger-Parker dominance index (BP dominance) are shown. (p87)

**Table 2.** Geographic relationships statistics among group communities. Mantel test results between group communities and geographic distances (\* $p < 0.01$ , \*\* $p < 0.001$ ). Beta dispersion comparison per basin and sub-basin. Beta-dispersion was calculated per sub-basin when there were at least two lakes to compare. Statistical significance was assessed through post-hoc

Tukey HSD for sub-basins: groups a–d are groups of analogous betadispersion patterns. (p91)

## CHAPTER 5

**Table 1.** Classification of zOTUs pairs by type (positive and negative), by domain (Bacteria and Eukarya), and by geographic scale (regional or local). (p105)

**Table 2.** Average values of node properties in the complete network grouped by modules. (p107)

**Table 3.** Comparison of network properties of the complete network and specific subnetworks based on interaction types (Bacteria/Eukarya - positive/negative). \*: Indicates that solely properties of the main component are shown (i.e., components of two nodes are removed). (p108)

**Table 4.** ML estimates for the ER and WS models. The log-likelihood has been reported as well. The ER model is selected because it has a smaller degree of complexity (a single parameter vs two parameters). The log-likelihood values for the selected model are marked in boldface. (p110)

**Table 5.** List of zOTUs with facilitation potential based on their node properties. (p113)

**Table 6.** List of zOTUs with active exclusion potential based on their node properties. (p115)

## INDEX

RESUMEN / RESUM / SUMMARY	1
GENERAL INTRODUCTION	9
<b>1. Microbial ecology in freshwater ecosystems</b>	
1.1 <i>Water masses in the alpine biome: a remote extreme environment</i>	12
1.2 <i>Putting freshwater microbes in the spotlight</i>	14
<b>2. Microbial community assembly in the alpine biome</b>	
2.1 <i>The Pyrenean metacommunity and community assembly rules</i>	15
2.2 <i>Scale dependent processes of community assembly</i>	16
2.3 <i>Spatial trends of temporal processes in the Pyrenees</i>	20
MAIN OBJECTIVES	23
GENERAL METHODS	29
RESULTS	
CHAPTER 1	37
<i>Consistent changes in the taxonomic structure and functional attributes of bacterial communities during primary succession</i>	
CHAPTER 2	47
<i>A complex integration of immigration and local pH shapes surface lake bacterial community assembly in the Pyrenean lacustrine district</i>	
CHAPTER 3	73
<i>High occurrence of Pacearchaeota and Woesearchaeota (Archaea superphylum DPANN) in the surface waters of oligotrophic high-altitude lakes</i>	
CHAPTER 4	83
<i>High planktonic diversity in mountain lakes contains similar contributions of autotrophic, heterotrophic and parasitic eukaryotic lifeforms</i>	
CHAPTER 5	97
<i>Microbial interactions based on co-occurrence and co-exclusion probabilities lie behind the ecology of Pyrenean freshwater lakes</i>	

## GENERAL DISCUSSION

### **1. The Pyrenean lake district microbiota**

- 1.1. *Core and transient organisms in the community ecology of lake surfaces* 133
- 1.2. *Unveiling new diversity of interacting microbiota and their ecological roles* 134

### **2. Environmental filtering in bacterial and eukaryotic communities**

- 2.1 *The process of environmental filtering on environmental preferences* 135
- 2.2. *Spatial trends of environmental filtering in eukaryotes* 136
- 2.3 *Environmental filtering and dispersal in bacteria vs. Eukaryotes* 137

### **3. Microbial community assembly and biogeography on the Pyrenees**

- 3.1. *Links of environmental and biotic preferences in community assembly* 138
- 3.2. *Implication of biotic interactions on community resilience and resistance* 140
- 3.3 *Trends in space of temporal succession* 141

GENERAL CONCLUSIONS 143

REFERENCES 147

ANNEX: SUPPLEMENTARY INFORMATION 157

## AGRADECIMIENTOS

Cuando iniciaba la tesis doctoral allá por el uno de enero de 2014 resultaba imposible evaluar la magnitud que cuatro años más tarde esta sección iba a alcanzar. Los agradecimientos de la tesis son, de alguna manera, una mirada hacia atrás y un recuerdo a quién me ha acompañado durante este (duro) camino que es el doctorado. El tiempo compartido con otros nos hace quien somos, y los recuerdos de esta época dejan una huella que sin duda definirá quienes seremos. Pues más allá del contenido científico, mi tiempo en Blanes ha supuesto un crecimiento personal de mucho más valor.

Inicio estos agradecimientos con el punto de partida tras un significativo viaje en solitario por Tailandia cuando me decidí a probar con un doctorado. A unos emails enviados en 2013 con quien ha sido mi director de tesis este tiempo, Emilio, que me terminarían de convencer para emprender este camino de cuatro años (mas un extra). Cabe destacar que dije que no a su primer ofrecimiento, y fue la visión que me planteó la que finalmente me atrajo a solicitar la beca FPI. De no haber insistido me habría perdido la oportunidad de conocer a multitud de científicos, compañeros y amigos de inigualable calidad profesional y sobretodo, humana.

Ya en Blanes, tuve la suerte de aterrizar en el *Piso Power*. Un piso cuya historia ceabina se remonta (al menos) a 2007 y espero que continúe tras mi partida. He tenido el extraordinario honor de compartir este piso con personas maravillosas que durante más o menos tiempo han compartido conmigo esta experiencia en Blanes. Mención especial a mis primeros compañeros, Guillem Roca y Anna Sans, que acogieron a este madrileño despistado en su casa y me dieron a conocer Catalunya de un modo muy personal. Excursiones al Montseny, *rogaines de orientació* en Pirineos, cenas compartidas, *tuppers* olvidados, *festes majors a pobles petits...* y que en definitiva, *fan caliu al llar*. Vuestra amistad es un tesoro. Pero la lista continúa. Por el piso han pasado Kathrin y su dialecto suizo-alemán, con Aina y Anna Thoran que descubrí Sense8 y las bondades de Netflix, Júlia que marcaba la *arribada de l'estiu*, Marcos que vino e hizo del piso su hogar, Mirco al que engañé para ir a Crossfit y él sigue y yo lo dejé, Nerea cuya actitud ante la vida resulta inspiradora, o Gerard cuyas plantas aromáticas me quedé tras su partida. Carmen, mi malagueña favorita del mundo entero, artista, paleoecóloga y brillante. Su envidiable resiliencia sólo es equiparable a la fortaleza de sus convicciones. Y Mateu, la última incorporación y mi heredero FPI cuya actitud encajó desde el minuto uno en un cóctel desenfrenado de conversacio-



nes en la cocina, plantas, *firstdates* y R a las 11 de la noche. Por muchas más gráficas, *plot-brother*. *Bonus track* a Sergio, empático compañero errante con quien he compartido desayunos ilegales más de un lunes por la mañana. A todos: ha sido un placer compartir *plannings* de limpieza, *tapergüers*, espacio, y un tiempo valioso que me guardo para siempre.

De casa al trabajo y del trabajo a casa, nada habría sido lo mismo sin los ceabinos, de quienes seguro me quedo corto en palabras. Sin ánimo de construir la casa por el tejado, mención súper especial al club de los gordacos y las meriendas de las cinco. Quién iba a decir que en los 30 minutos que corre un gel de PCR podrían generarse amistades como las que vinieron. En primer lugar a los co-fundadores, Magda, persona extraordinaria de quien siempre admiraré su valentía y su decisión, cuya integridad siempre ha estado por encima de convencionalismos. Y a Leire. Pieza fundamental. No sé cuántos téis hemos compartido, pero sé que esto habría sido muy diferente si no hubiese tenido el honor de una amistad como la tuya. Mi taza siempre estará preparada para que toques en mi puerta a las cinco de la tarde. Y ¿cuántas personas han pasado por nuestras meriendas? Cèlia, cuando viniste de visita durante mi segunda estancia en San Diego me diste un aire que me faltaba. Y de nuevo fue así en noviembre durante nuestra aventura en Borneo. Ojalá el futuro nos aguarde más sorpresas y *brunchs* en la playa. Sara, mi contrincante favorita de duelos de miradas, *grace* y bultos a partes iguales pero ante todo, un ejemplo de sobreponerse a las dificultades y resurgir. Compañeros y amigos que vienen y van, como Carlos Palacín, con quien buenas excursiones montamos al Pirineo y a anillar pájaros. Ibor, ser libre que salta por las montañas y los bares con la misma fluidez. Vicente, co-propietario de la báscula del despacho que compartimos brevemente y brillante teórico. María López cuya energía y alegría se transmite con su sola presencia. Clúster, siempre recordaremos aquella película. Marta García, cuyo talento audiovisual y bailes en las cenas de navidad son maravillosos. Luisfran y Nayeli, que son unos *hosts* y vecinos de 10 en las partidas de *boardgames* que suelen convocar. Y qué decir de barra\_baixa, mítico grupo musical que pasará a la historia por sus versiones de Katy Perry, Cohen, los Beatles o Sau durante las barbacoas y comidas de navidad en el CEAB. Ha sido un placer cantar con vosotros, Xevi, Joan G., Fede e Ibor. A los crossfitters, Thor, Ambra, Miquel, Mirco, Mateu y Jorge, cuya compañía durante horas de sudor nos han hecho más sanos y fuertes. Y a muchos más, Eliot, Héctor, Xavi Torras, Maria Elena, Marc Sala, Carla, Aitziber, Mireya, Dani B, Eneko, Jana, Jordi B, Jordi P, Núria, Rubén, Víctor, Jenny, Angela, Chiara, Francesco, Marta Royo y Mikel Aguirre (no ceabinos pero doctorandos

coetáneos) y seguro que me dejo a más de uno.

Si hay alguien a quien deba agradecer que esta tesis haya salido adelante es sin duda a mis compañeros de grupo. Xavier Triadó, *Xevi, ets molt crack*, tu buen hacer, tu temple, incluso tu castellano de libro. Joan Cáliz, *an altre crack*, que me enseñó a analizar secuencias cuando no sabía hacer la o con un canuto. A vosotros dos os tengo que agradecer vuestras enseñanzas y consejos en el terreno profesional, y vuestra paciencia para escucharme en el personal cuando llegaba al despacho con un día regular. Lo mismo a Tomàs, que me dejó en herencia su mesa y algunos scripts. Buen científico, buen corredor y un *blanenc autèntic*. El futuro queda en manos de Gerard, que a buen seguro destacará en todo en lo que se proponga. Bueno y de Mateu, pero como es el becario del bridges lo petará seguro (buena suerte con las fórmulas (J)). Del bridges tengo que destacar a Capi, que cada vez que viene revoluciona (¿Y la europea?) y tiene la gran habilidad de hacer entender mates y teoría a cualquiera.

Como el CEAB tiene presencia internacional, he tenido la gran oportunidad de irme de estancia en dos ocasiones. En primer lugar, en 2015 a Boulder (Colorado) donde tuve el gusto de trabajar con Albert Barberán, ex-ceab. Sin estos tres meses esta tesis no existiría. Mis habilidades con R se incrementaron exponencialmente gracias a Albert, y he de agradecer además numerosos ratos de charlas y debates sobre la vida y la situación política y del mundo. También te debo mi actual ‘afición’ al café. *Furthermore, thanks to Noah Fierer for his knowledge, help, and welcoming attitude. I learnt a lot from you preparing the succession manuscript. During my time in Boulder I was hosted under the roof of María, Rob and Taos. I have to thank you incredibly because you made me feel at home. Also, the projector in my basement was amazing, as the aerial classes we shared.* ¡Ah! Y descubrí que el Parkour es muy divertido. En mi segunda estancia, en 2016 marché a San Diego (California), para colaborar en el grupo de Christopher Dupont. *Chris, these three months working with you were amongst the most motivating ones of my PhD. You encouraged my creativity and valued it, while building-up upon the ideas I shared. Also, thanks to Drishti, Josh and Alex for their time. I also have to thank to Jonathan, who drove me here and there, sharing Yoga lessons, meaningful conversations and quality time despite I was always busy at work. Finally, I want to thank Goody and Brian, delightful travelers who hosted me and made me feel at home, a truly American home during the American elections. I am such a tourist sometimes.*

Y como somos de dónde venimos, no puedo en estos agradecimientos

dejar de hablar de personas fundamentales no blanenses. Gracias a Asun y a Sergio por introducirme en el mundo científico y demostrar que además de jefes sois amigos. Gracias también a los colegas de esta etapa, Mariajo, MiguelÁngel e Isaac. Gracias a *los feeders*, Carlos, Pablo, Adri, y Gon. Y a Jose Ramón y a Paloma. La prueba de que la ciencia y la música pueden ir de la mano. Os quiero. Miles de emails, wasaps y hasta un disco común nos avalan. También a amistades de la época de la carrera que sobreviven el paso del tiempo, y son muy valiosas para mí: Javi, Isra, Pau, Selín, Anadeory, Adriana, Iris, Bea Sánchez, Ceci o Aurosmith. También del año de Erasmus, en especial a Carmen y a Gema, el póster que me regalasteis cuando dejé Madrid me ha acompañado todo este tiempo en Blanes. Y por los reencuentros, a Iraia y a Fer. Y si hablamos de amistades relevantes, este párrafo es para vosotros. Porque estuvisteis, estáis, y estaréis. Ana, Dani, Miguel y Ruth. Os adoro. Los tiempos cambian, las personas evolucionan y la vida continúa, pero nosotros estaremos ahí aprendiendo.

A mi madre, que me inculcó el amor a la naturaleza en todas sus facetas y que es, en definitiva, la gran culpable de que esta tesis exista. Y no sólo eso, gracias porque siempre me has dado libertad y porque siempre has confiado en mi y en mi criterio. Porque siempre me has apoyado en todo. Y a mi tía, que siempre ha sido el contrapunto perfecto, manteniéndome atento para hacer las cosas bien, salvándome los exámenes de matemáticas e inundar mi mente de imaginación y creatividad desde que tengo memoria. A las dos sólo os puedo decir, que gracias por todo y que os quiero.

Para rebajar el impacto emocional antes de acabar añado unos agradecimientos aleatorios. Gracias a gusanito (mi gato), a M. Rajoy por sus frases y las risas que nos da, a las pepis por suministrarme frutas y verduras con tan buen humor, a Netflix, a mis plantas, y sobre todo a la compañía que no me ha abandonado nunca y sin quien no puedo vivir: gracias Spotify por toda la música que me das cada día. Y acabo estos agradecimientos con una reflexión. Una tesis es el resultado final de una etapa formativa. No es fácil hacer una tesis, muchas de las personas aquí mencionadas lo saben. Y nunca se sabe cuándo una simple frase o una breve conversación pueden calar y convertirse en un anclaje que te permita avanzar. Estos agradecimientos intentan reflejarlo. Somos las decisiones que tomamos, las personas que hemos conocido, y las personas que conoceremos. No creo en los finales pero estos agradecimientos finalizan una etapa. Una etapa que es el inicio de lo que está por venir. Y los recuerdos de esta historia formarán parte de la siguiente. Gracias.

*A mi madre Adela  
y a mi tía Carmen.*

*(...) hasta que en el horizonte al final de la calle sólo queda un espacio infinito de color azul en el que alguien ha pintado estrellas tan brillantes como las notas altas de un piano, esperando un acorde final que nunca llega. (...). R.O.A. (2017).*



Resumen / Resum / Summary



## RESUMEN

Durante décadas, el componente microbiano de los ecosistemas ha sido considerado una caja negra debido a las complejidades de su estudio en favor de organismos macroscópicos. Hoy los avances tecnológicos han desvelado la altísima diversidad y complejidad de arqueas, bacterias y eucariotas y han permitido un análisis ecológico a la altura de los estudios desarrollados en macroecología. Así, los marcos teóricos de sucesión ecológica y conceptos de metacomunidades, permiten abordar la comprensión del ensamblaje ecológico de las comunidades. Esta tesis ha abordado estas cuestiones centrándose en las comunidades lacustres de una región geológicamente diferenciada: la cordillera montañosa de los Pirineos. Pese a ser uno de los enclaves más estudiados dentro de la limnología de alta montaña, el conocimiento actual es parcial al centrarse en grupos poblacionales concretos o áreas geográficas reducidas. Por ello esta tesis, busca describir la diversidad taxonómica y revelar los procesos ecológicos (bióticos y abióticos) que pueden estar tras la distribución espacial y la heterogeneidad de las comunidades superficiales de lagos alpinos. Mediante un muestreo de más de 300 lagos durante el verano de 2011 y técnicas de secuenciación masiva de genes ribosómicos (16S y 18S rRNA) se obtuvieron los perfiles microbianos de comunidades de arqueas, bacterias y eucariotas. Mediante un meta-análisis de estudios previamente publicados se comprobó que las comunidades microbianas tienen patrones predecibles en cuanto a sus cambios taxonómicos y funcionales a lo largo del tiempo, en concreto un aumento generalizado en el número de especies y funciones, y que la composición de estas especies y funciones suele ser menos variable en comunidades más maduras mostrando tendencia a la oligotrofia. Estos rasgos son extrapolables a un muestreo espacial. En este sentido los patrones de las comunidades de bacterias en Pirineos muestran una integración compleja de la dispersión de fuentes ambientales a nivel de cuenca y filtros ambientales, procesos que a lo largo del tiempo parecen dejar una huella espacial que recuerda a lo observado en el meta-análisis. Además, en este estudio se ha profundizado en el estudio de grupos microbianos que no era posible estudiar con técnicas tradicionales o en los que los primeros estudios moleculares quedaban fuera del límite de detección. En la región pirenaica hemos detectado un alto número de arqueas de linajes nuevos (como *Pacearchaeota*, *Woearchaeota*) que parecen dominar en términos de ocurrencia y que podrían presentar interacciones con otros organismos. Respecto a los eucariotas microscópicos, esta tesis resalta una dominancia de algas crisofíceas en términos de riqueza, abundancia y



también en novedad. Asimismo, destaca una inesperada riqueza de grupos heterótrofos fagótrofos y parásitos. En particular, cercozoos y hongos quitridios que presentan patrones de diversidad análogos a los grupos autótrofos y que sugiere reconsiderar las cadenas tróficas lacustres. La prominente beta dispersión mostrada por parásitos sugiere interacciones altamente específicas y un rol muy relevante en las cadenas tróficas. Curiosamente, la metacomunidad de Pirineos parece contener poblaciones más específicas que su homólogo oligotrófico marino, presentando una mayor diversidad beta. En términos de diversidad, las particularidades de los Pirineos parecen fomentar una alta diversificación genética de formas de vida eucarióticas y de arqueas. El estudio de coocurrencias y coexclusiones de especies muestra un filtrado ambiental en base a sus afinidades y rangos de tolerancia ambiental, pero también muestra las interacciones bióticas que pueden establecerse entre los distintos componentes de la comunidad. A partir de análisis de redes de estas interacciones bióticas, se han establecido relaciones concretas entre parásitos y mixótrofos al margen de sus nichos ambientales. Las bacterias dependen menos del ambiente que los eucariotas, y se dividen en aquellas que establecen relaciones más duraderas típicas de sistemas acuáticos, y otras bacterias asociadas a ambientes terrestres que utilizan los lagos como medio para su dispersión y colonización. A nivel general, se observa que la competencia en el conjunto de especies parece enlazada a una estrategia de compartir recursos, mediante la limitación de abundancias en la comunidad. Asimismo, se ha intentado identificar organismos clave mediante sus propiedades en la red de interacciones. Las propiedades conjuntas de las distintas redes de interacciones biológicas permiten extraer conclusiones sobre la resistencia y resiliencia frente a perturbaciones ambientales de las distintas comunidades. Tanto la composición, alta diversidad y potenciales reglas de ensamblaje de comunidades del distrito lacustre de Pirineos aquí presentadas buscan inspirar más estudios biogeográficos para evaluar si las ideas aquí descritas son una excepción o una regla en los biomas alpinos de todo el planeta.

## RESUM

Durant dècades, el component microbià dels ecosistemes ha estat considerat una caixa negra degut a la complexitat del seu estudi en comparació als organismes macroscòpics. A dia d'avui els avançaments tecnològics han revelat una altíssima diversitat i complexitat d'arquees, bacteris i eucariotes, i han permès un anàlisi ecològic del nivell dels estudis desenvolupats en la macroecologia. Així, els marcs teòrics de successió ecològica i els conceptes de metacomunitats permeten abordar la comprensió de l'acoblament ecològic de les comunitats. Aquesta tesi ha abordat aquestes qüestions centrant-se en les comunitats lacustres d'una regió geològicament diferenciada: la serrelada dels Pirineus. Tot i ser un dels enclavaments més estudiats de la limnologia d'alta muntanya, el coneixement actual és parcial si ens centrem en grups poblacionals concrets o àrees geogràfiques reduïdes. És per això que aquesta tesi busca descriure la diversitat taxonòmica i revelar els processos ecològics (biòtics i abiòtics) que es troben darrere la distribució espacial i la heterogeneïtat de les comunitats superficials dels llacs alpins. Mitjançant un mostreig de més de 300 llacs durant l'estiu de 2011 i tècniques de seqüenciació massiva de gens ribosomals (16S i 18S rRNA) es va obtenir els perfils microbians de les comunitats d'arquees, bacteries i eucariotes. A través d'un meta-anàlisi d'estudis prèviament publicats, es va comprovar que les comunitats microbianes tenen patrons predictibles pel que fa als canvis taxonòmics i funcionals al llarg del temps. Concretament, s'ha constatat un augment generalitzat en el número d'espècies i de funcions, i que la composició d'aquestes espècies i funcions sol ser menys variable en aquelles comunitats més madures que mostren una tendència a l'oligotròfia. Aquests trets són extrapolables a un mostreig espacial. En aquest sentit els patrons de les comunitats de bacteris en els Pirineus mostren un integració complexa de la dispersió de fonts ambientals a nivell de conca i de filtres ambientals, processos que al llarg del temps semblen deixar una empremta espacial que recorda al que s'ha observat en el meta-anàlisi. A més, en aquest treball s'ha aprofunditzat en l'estudi de grups microbians que no s'havien pogut estudiar abans amb tècniques tradicionals, o que en els primers estudis moleculars quedaven fora del límit de detecció. En la regió pirinenca hem detectat un alt nombre d'arquees de nous llinatges (com *Pacearchaeota*, *Woesearchaeota*) que semblen dominar en termes d'ocurrència i que podrien presentar interaccions amb altres organismes. Pel que fa als eucariotes microscòpics, aquesta tesi ressalta una dominància d'algues crisofícies en termes de riquesa, abundància i de novetat. Així mateix, destaca una inesperada riquesa de grups heteròtrofs fagotròfics i paràsits. Particularment, cercozous i fongs quitridis, que presenten patrons de diversitat anàlegs als grups autòtrofs, fet que suggereix reconsiderar les cadenes tròfiques lacustres. La prominent

beta dispersió mostrada per paràsits suggereix que les interaccions són altament específiques i que tenen un paper molt important en les cadenes tròfiques. Curiosament, la metacomunitat de Pirineus sembla tenir poblacions més específiques que el seu homòleg oligotròfic marí, presentant una major beta diversitat. En termes de diversitat, les particularitats dels Pirineus semblen fomentar una alta diversificació genètica de formes de vida d'eucariotes i d'arquees. L'estudi de concurrències i exclusions mútues d'espècies mostra un filtrat ambiental en base a les seves afinitats i rangs de tolerància ambiental, però també interaccions biòtiques que poden establir-se entre els diferents components de la comunitat. A partir de l'anàlisi de les xarxes d'aquestes interaccions biòtiques, s'han establert relacions concretes entre paràsits i mixòtrofs al marge dels seus nínxols ambientals. Els bacteris depenen menys de l'ambient que els eucariotes, i es divideixen en aquells que estableixen relacions més perdurables, típiques de sistemes aquàtics, i d'altres bacteris associats a ambients terrestres que utilitzen els llacs com a mitjà de dispersió i colonització. A nivell general, s'observa que la competència en el conjunt de les espècies sembla relacionar-se amb una estratègia de compartir recursos, a través de la limitació de les abundàncies en la comunitat. De la mateixa manera, s'ha intentat identificar organismes clau mitjançant les seves propietats en la xarxa d'interaccions. Les propietats conjuntes de les diferents xarxes d'interaccions biològiques permeten treure conclusions sobre la resistència i resiliència davant les perturbacions ambientals de les diferents comunitats. Tant la composició, com l'alta diversitat i les regles potencials d'acoblament de comunitats de la zona lacustre dels Pirineus aquí presentades volen inspirar més estudis biogeogràfics per avaluar si les idees descrites aquí són una excepció o més aviat una regla en altres biomes alpins del planeta.

## SUMMARY

For decades, the microbial component of ecosystems has been considered a black box due to the complexities of its study in favor of macroscopic organisms. Today technological advances have revealed the high diversity and complexity of archaea, bacteria and eukaryotes and have allowed the ecological analysis that matches the studies developed in macroecology. Thus, the theoretical frameworks of ecological succession and metacomunity concepts allow us to address the understanding of the ecological assembly of communities. This thesis has addressed these issues by focusing on the lake communities of a geologically differentiated region: the mountain range of the Pyrenees. In spite of being one of the most studied enclaves within high mountain limnology, the present knowledge is partial because it focuses on specific population groups or small geographical areas. Therefore, this thesis seeks to describe the taxonomic diversity and reveal the ecological processes (biotic and abiotic) that may be behind the spatial distribution and heterogeneity of the surface communities of alpine lakes. Through a sampling of more than 300 lakes during the summer of 2011 and massive sequencing techniques of ribosomal genes (16S and 18S rRNA) microbial profiles of communities of archaea, bacteria and eukaryotes were obtained. Through a meta-analysis of previously published studies, we found that microbial communities have predictable patterns in terms of their taxonomic and functional changes over time, specifically a generalized increase in the number of species and functions, and that the composition of these species and functions tend to be less variable in more mature communities showing trends towards oligotrophy. These features are extrapolated to the spatial dataset. In this sense, the patterns of bacterial communities in the Pyrenees show a complex integration of the dispersion of environmental sources at the basin level and environmental filters, processes that over time seem to leave a spatial trace that recalls what was observed in the meta-analysis. In addition, this study has deepened the study of microbial groups that it was not possible to describe with traditional techniques or that were outside the limit of detection in the first molecular studies. Furthermore, in the Pyrenean region we have detected a high number of archaea of new lineages (such as Pacearchaeota or Woesearchaeota) that seem to dominate in terms of occurrence and that could present interactions with other organisms. Regarding microscopic eukaryotes, this thesis highlights a dominance of Chrysophyceae algae in terms of richness, abundance and also in novelty. It also highlights an unexpected richness of heterotrophic phagotrophs and parasites. In particular, Cercozoa and Chytridiomycota exhibit diversity patterns

analogous to autotrophic groups which suggests a reconsideration of lacustrine food webs. The prominent beta dispersion shown by parasites suggests highly specific interactions and a relevant role in the food web. Interestingly, the metacommunity of the Pyrenees seems to contain more specific populations than its marine oligotrophic counterpart, presenting greater beta diversity. In terms of diversity, the particularities of the Pyrenees seem to promote a high genetic diversification of eukaryotic life forms and archaea. The study of co-occurrences and coexclusions of species shows an environmental filtering based on their affinities and ranges of environmental tolerance, but also shows the biotic interactions that can be established between the different components of the community. From the analysis of networks of these biotic interactions, concrete relationships between parasites and mixotrophs have been established regardless of their environmental niches. Bacteria depend less on the environment than eukaryotes, and are divided into those that establish more durable relationships typical of aquatic systems, and other bacteria associated with terrestrial environments that use lakes as a means of dispersal and colonization. At a general level, it is observed that the competition in the set of species seems linked to a strategy of resource sharing, by limiting abundances in the community. Likewise, an attempt has been made to identify key organisms through their properties in the network of interactions. The joint properties of the different networks of biological interactions allow drawing conclusions about the resistance and resilience to environmental disturbances of the different communities. The composition, high diversity and potential assembly rules of the Pyrenean lake district communities presented here seek to inspire more biogeographic studies to assess whether the observed patterns are an exception or a rule in other alpine areas of the planet.

# General introduction



*“A lake is to the naturalist a chapter out of the history of a primeval time, for the conditions of life there are primitive,—the forms of life are, as a whole, relatively low and ancient, and the system of organic interactions by which they influence and control each other has remained substantially unchanged from a remote geological period.” Stephen A. Forbes, 1887*

When Stephen A. Forbes, one of the earliest limnologists, wrote in 1887 his seminal paper “The lake as a microcosm”, he couldn’t know that his somewhat poetic words would anticipate a myriad of studies on food-webs, population and community ecology. He and many colleagues who followed, focused on small animals from concrete localities; but he already acknowledged that whatever affects any species must have an influence upon the whole assemblage. Algae, protozoa and even animalcules should have a ‘role impossible to describe within reasonable limits’. The smallest part of the microbial ecosystem, the ‘animalcules’ as it was coined by the father of microbiology Antonie van Leeuwenhoek in the 17<sup>th</sup> century, was particularly neglected for centuries in favor of the more accessible fraction of biological communities. Today, high throughput technologies have revealed an impressive diversity of Archaea, Bacteria and small Eukarya, so the ecological analysis of the microbial biosphere is not impossible to achieve nor imaginative anymore.

Ecological theory dwells into multiple potential processes that explain how communities establish and change (Margalef, 1963, 1968). The particularities of these processes along the spatial scale can be studied through metacommunity theory (Leibold *et al.*, 2004), although, this approach is typically underrepresented in the microbial literature. To comprehend metacommunities, the study of a system with partially isolated local communities within a regional context should be the aim, since the spatial structure may resemble the ecological processes taking place at different spatial scales. In this respect, the Pyrenees mountain range is especially suitable, since the region harbors one of the most intensive and extensively studied limnological areas in the alpine biome to date (Catalan *et al.*, 2006). However, because of the technical and logistic challenges, comprehensive studies in the Pyrenees within a community ecology perspective are yet based in a restricted number of lakes within small sampling areas, or considering solely partial components of the lake communities (i.e., phytoplankton, zooplankton, or animal groups (i.e., amphibians or fishes)). This PhD dissertation aims to unveil all the microbial components and to uncover the ecological processes (both biotic and abiotic) behind the spatial distribution and biological heterogeneity of lake surface communities in the Pyrenees

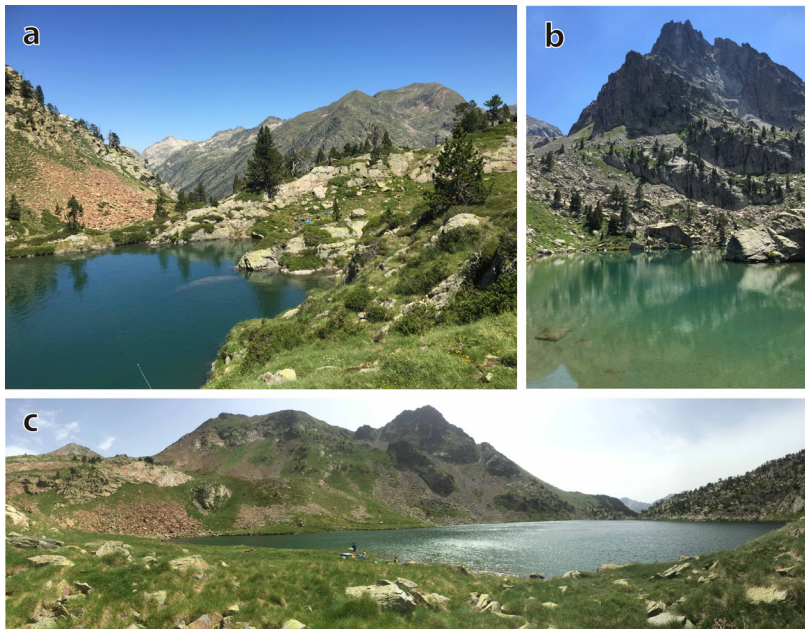


high mountain lacustrine district, circumventing some of the limitations that have hampered these studies in the past.

## 1. Microbial ecology in freshwater ecosystems

### 1.1 Water masses in the alpine biome: a remote extreme environment

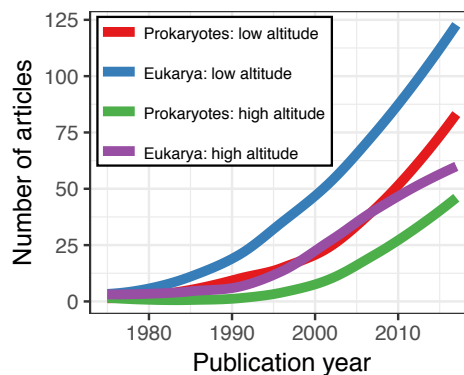
Earth's surface is covered by 70% of water (an estimated volume of  $1.38 \times 10^9 \text{ km}^3$ ), of which 0.0013% is non-frozen freshwater occurring within continental boundaries. Freshwater ecosystems are very diverse, and are often highly connected. Water can accumulate underground in subterranean aquifers nourishing wetlands, and flow through streams and rivers to ponds and lakes. Since the terrestrial environment surrounds continental water bodies, lakes are the most isolated of the different hydrographic compartments and are highly heterogeneous. Their geographic location, ranging from sea level to high altitudes and from tropical to polar or alpine biomes, confer lakes with widely heterogeneous properties, trophic dynamics and sizes (Downing *et al.*, 2006).



**Figure 1.** Lakes from the Pyrenees mountain range: an alpine biome. The photographs show (a) Rovinets, (b) Xic de Subenuix and (c) Baborte, during summer of 2015.

In the alpine biome, remote lakes are exposed to a cold, snowy and windy environment. Alpine lakes can be found in high mountain regions all around the world located above the tree line. That is the altitudinal edge of the habitat at which trees are capable of growing, which is limited by Latitude (Allen and Walsh, 1996; Malanson *et al.*, 2007). The alpine biome characterizes the Himalayan Mountains, the Alps, the Rocky Mountains, or the Pyrenees mountain range. Europe comprises many high mountain systems in its orography, especially in the colliding areas of the continental crust, in Southern Europe. The Alps is one of the largest mountain systems, followed by The Pyrenees and the Scandinavian mountains. These regions are affected by marked seasonality, with summer temperatures that can range from 10 to 15°C and winter temperatures falling below freezing. Also, exposure to UV wavelengths is stronger because of being up higher in the atmosphere and this has effects on the microbiota (Sommaruga *et al.*, 1996) but also the exposure of lake plankton to long and dark winters. Isolated life forms have had to adapt to these extreme conditions in these regions.

Historically, paleoecologists have studied plankton microfossils from lake sedimentary records (Lotter, 1989; Müller *et al.*, 1998; Catalan *et al.*, 2002), while limnologists have focused on present alive planktonic communities through morphology and microscopy (Pernthaler *et al.*, 1996; Wille *et al.*, 1999). More recently and with the onset of molecular techniques (Pernthaler and Glöckner, 1998; Alfreider *et al.*, 1996), the number of publications regarding microbiology in freshwater lakes has been increasing every year (Figure 2).



**Figure 2.** Number of limnological papers published per year through searches in ‘web of science’ (WoS) according to different keywords: high altitude (alpine or sub-alpine), low altitude (lowland or montane), prokaryotes (Bacteria, Archaea or Bacterioplankton) and eukaryotes (Protists, Algae or Phytoplankton).

Still, the comparison of the number of publications between the alpine biome and lower altitude biomes (i.e., lowland or montane) shows a large comparative gap in the current knowledge of the biology of high altitude lakes. As expected, the prokaryotic fraction is the group with the largest gap so far.

### 1.2 Putting freshwater microbes in the spotlight

Microorganisms root beneath the origin of life on planet earth, and have evolved to exist in every natural and anthropic system. The origin of life, regardless of different hypothesis and time estimates, is linked to aquatic environments with extreme conditions (Baross and Hoffman, 1985; Martin *et al.*, 2008). Not surprisingly, the three main life domains -Bacteria, Archaea and Eukarya- are always present and have evolved to inhabit all the variety of water masses across the planet even under euxinia or hyperhaline conditions for small eukaryotes (Triadó-Margarit and Casamayor, 2013, 2015). Diversity of the three branches is still far from being completely understood, in fact whole new lineages are still being unveiled nowadays (Hug *et al.*, 2016), such as the CPR (Candidate Phyla Radiation), DPANN (named after five archaea phyla), widespread in the Pyrenees (see chapter 3) or Asgard archaea (Spang *et al.*, 2015): a potential evolutionary link between the Archaea and Eukarya (Zaremba-Niedzwiedzka *et al.*, 2017). Since late 1980, microbial ecology has been on an ongoing revolution lately boosted by the development of the so-called 'Next generation sequencing' (NGS) technologies. Early molecular studies with only a few sequences from underrepresented communities tried to explain and understand the ecology of microbes in nature (Pace *et al.*, 1986; Alfreider *et al.*, 1996). These valuable contributions started to capture the immensity of the microbial biosphere, but the limitations on understanding its diversity and ecological dynamics have been evident (Zinger *et al.*, 2012).

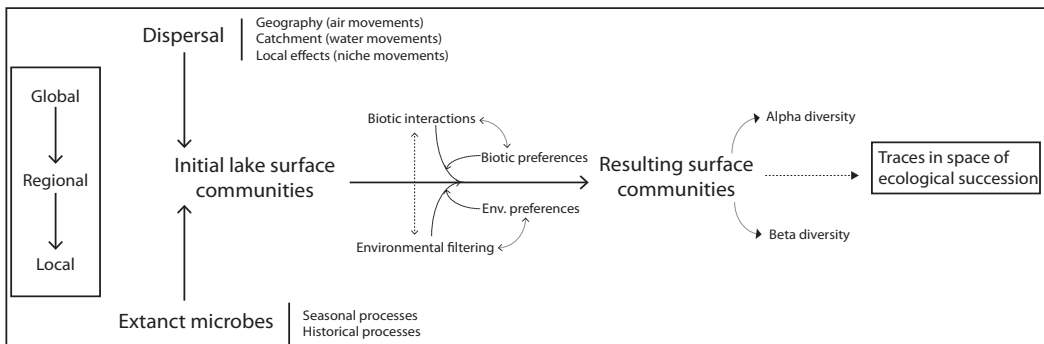
Given the aquatic origin of microorganisms, all life branches can thrive in the water, and their relevance nowadays in the high mountain freshwater environments is not anecdotic. Because of the low amount of nutrients available in high mountain lakes, fixation, mobilization and accumulation of nutrients are crucial for the whole ecosystem functioning (Grossmann *et al.*, 2016a). Nutrient cycling and mobilization by microbes in highly elevated lakes will have a fertilizing impact downhill through transport and accumulation of organic compounds (Ciccazzo *et al.*, 2016). Microbes capable of attaching to surfaces in oligotrophic areas are involved in

substrate bioweathering, creating further diverse habitats (i.e. Bartrons *et al.*, 2012) for other organisms to thrive and therefore increasing local diversity. These microbes, rather than simply adapting to external conditions they thrive together (Faust and Raes, 2012) and need to cope with oligotrophy, UV radiation and low temperatures together. In fact, these factors can have an impact not in the environmental preference of an organism, but in the interaction of many (Sommaruga *et al.*, 1996). Quantification and description of diversity, and understanding the general rules of community assembly in high mountain lakes would allow to predict vulnerability and stability of populations in the current climate change scenario (Cardinale *et al.*, 2012).

## 2. Microbial community assembly in the alpine biome

### 2.1 The Pyrenean metacommunity and community assembly rules

For years, dozens of general hypothesis have been proposed to explain the diversity of communities (Palmer, 1994). Potential processes shaping biological communities can be summarized into four: dispersal, environmental filtering, ecological drift and speciation (Vellend, 2010). Currently accepted ecological theory posits that communities assemble through dispersal from spatial pools of populations and species sorting processes to form local communities (Barberán *et al.*, 2014a) according to their environmental preferences (Martiny *et al.*, 2006; Lindström and Langenheder, 2012) or biotic preferences, that is, species interactions (Blois *et al.*, 2013) (Figure 3).



**Figure 3.** Summary of processes that potentially drive community assembly and ecological succession, therefore, diversity.

Local communities are comprised by all the individuals of populations that thrive within a single patch, or a local area of a given habitat. Together, sets of local communities that are linked by dispersal of multiple potentially interacting species form metacommunities (Hanski and Gilpin, 1991; Wilson, 1992). Metacommunities are studied under the lens of community ecology, which is concerned with explaining the patterns of distribution, abundance and interaction of species within a given area. The metacommunity concept is an important way to think about linkages between different spatial scales in ecology (Leibold *et al.*, 2004). Not all the processes may be equally important regarding spatial or temporal scales (Vellend, 2010), and their complex integration may result in different community patterns along the sampling region. In the Pyrenees, each lake acts as a local community, closely related and influenced by catchment processes, all within the total potential species of the region (regional meta-community) (Barberán *et al.*, 2014a)

The ecological processes and dynamics shape community assembly of a complex interacting planktonic microbiota, under the idiosyncratic characteristics of regional geography and geology leading to a specific community pattern. The microbial metacommunity in the Pyrenees is likely to be subject to the past and present regional characteristics. First, the mountain range is a unit formed through the same geological conditions and long term climatic forces that have shaped lake ontogeny (Delmas *et al.*, 2008), so spatial patterns should reflect, at least partially, this fact. Also, the Pyrenees lie between two big water masses: the Atlantic Ocean in the west and the Mediterranean Sea in the east. This circumstance makes lake communities, specially surface communities during summer, affected by northern winds with oceanic influence; although surface communities may also be affected by Saharan dust plumes driven by southern winds (Hervas and Casamayor, 2009). Therefore, different lakes may have received quantitatively different aerial inputs during the seasonal cycle (Vila-Costa *et al.*, 2013). Altogether, the Pyrenean meta-community is connected and mobilized through atmospheric deposition (Hervas and Casamayor, 2009) and water flows (Niño-García *et al.*, 2016) within the regional scale but also at the catchment level (Felip *et al.*, 1999b).

## *2.2 Scale dependent processes of community assembly*

### 2.2.1 Historical processes in the Pyrenean region

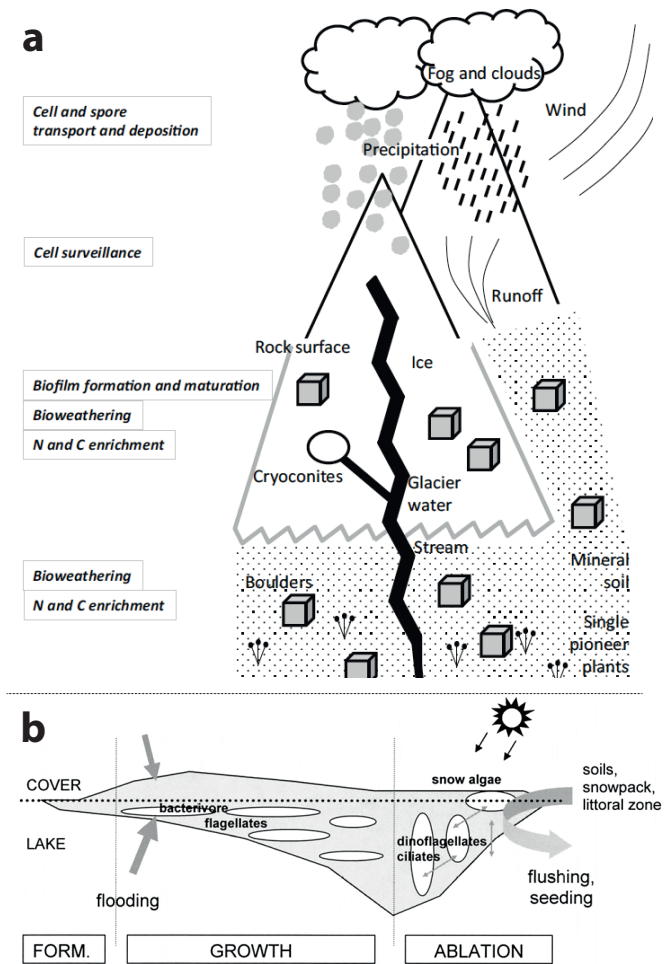
A lacustrine region is, in a way, the result of a scientific experiment

designed by the geological scale. High mountain lakes were formed during de-glaciation processes that occurred after the last glacial maximum, and during the onset of the Holocene epoch approximately 11,700 years before present (Walker *et al.*, 2009). Each lake can be understood as a replicate of the same driving processes, although each replicate has concrete characteristics that become the factors suitable of comparison. The regional scale allows taking a picture on a specific time of that geologic experiment. In this context the historical assembly processes are the broad research hypothesis, since it is not feasible to simulate in the lab the conditions of >10,000 years of lake dynamics and ecological processes such as ecological succession. Although the sedimentary record can track back some of the historical processes and past environmental situations, only a fraction of the microbial populations can be retrieved from the paleoecological record, mostly diatoms, dinoflagellates or chrysophytes (Lotter, 1989; Lotter and Bigler, 2000; Pla *et al.*, 2003), leaving a substantial part of the community untraceable. Present communities are at least partially the result of these historical processes. However the glaciation dynamics have differed in the different high mountain systems (Delmas, 2015). In the Pyrenees, de-glaciation was longer because it started earlier than recorded on a global scale, which surely has had an impact on the timing of colonization and evolution by pioneer communities. Lake formation starts after glacial retreat, with water flows having a strong role in the weathering of rock and in the movement of particles and nutrients (Ciccazzo *et al.*, 2016), and ultimately in early lake communities (Peter and Sommaruga, 2016).

### 2.2.2 Catchment and seasonal processes

Within the region, there might be important processes within smaller areas of higher connectivity, that is, within each catchment or each mountain. Because of the altitudinal succession of weathering and lake formation downhill, added to the temperature gradient, there are often patterns of communities related to the altitudinal gradient (Wang *et al.*, 2016); although these patterns may be different across mountains and therefore, have strong differences in regional or multiregional gradients (Li *et al.*, 2017). Community patterns linked to altitude may be the result of catchment processes, such as water flows from the relatively infertile oligotrophic high altitude lakes to the lowland lakes with more eutrophic waters, which adds directionality to microbial colonizers (Figure 4a). For instance, allochthonous organic carbon from the surrounding terrestrial environment has an important influence in the littoral community and these fluxes tend to be smaller in lakes above the tree line (Leavitt *et al.*, 1997). These catchment fluxes may be crucial through

the seasonal cycle. In winter, an ice cover isolates lakes for several months until spring, when the ablation of the ice cover connects the lake again to the hydrographic network (Felip *et al.*, 1999b). During the ablation period, which is dependent of the altitudinal gradient, water movements through the catchment will seed lake communities (Felip *et al.*, 1999a) (Figure 4b). Theory states that during a disturbance, established communities are more willing to accept immigrants (Shurin and Allen, 2001; Leibold *et al.*, 2004), suggesting that the seasonal swift between winter and spring may be key in the summer communities.



**Figure 4.** (a) Mountains viewed under the lens of primary succession: microbial colonization through air and transport through habitat connectivity downstream. (b) Seasonal processes in high mountain lakes (Catalan *et al.*, 2006): ice cover formation, growth and spring ablation linked to water fluxes and particle movements.

### 2.2.3 Local processes

There are processes acting within each lake that shape the resulting composition of the microbial communities (Figure 3). The process of species sorting locally considers the environmental idiosyncrasies of each lake to 'order' microbes in terms of their environmental preferences (Van der Gucht *et al.*, 2007). Most lakes in the Pyrenees lie within ultraoligotrophic/oligotrophic ranges (Camarero *et al.*, 2009), covering distinct environmental conditions compared to lakes in Tibet or the Swiss alps (Caliz and Casamayor, 2014). Alpine lakes in the Pyrenees are generally oligotrophic; therefore lake trophic status is not expected as a major segregator of communities as it may happen in eutrophized lakes (Catalan *et al.*, 2009). However, a threshold separating ultraoligotrophic and oligotrophic conditions may affect the amount and quality of DOC, which in turn affects their UV absorptive properties (Laurion *et al.*, 2000). Also, the balance of atmospheric deposition of phosphorus and nitrogen as limiting nutrients of the system (Camarero and Catalan, 2012) may be relevant regarding the resources and strategies required for the local communities to thrive. For instance, in the Alps phytoplankton and zooplankton are not solely explained by the climatology (Tolotti *et al.*, 2012) but also by the trophic status (Tolotti *et al.*, 2012; Grossmann *et al.*, 2016b) and alkalinity (Marchetto *et al.*, 2009; Grossmann *et al.*, 2016b).

Species sorting is traditionally linked to environmental filtering. However, biotic interactions (i.e. 'biotic filtering') should be considered as a similar process of ordering species within a community. Early studies showed the protozoan control on phytoplankton (Pernthaler *et al.*, 1996), or the bacterivory displayed by ciliates and flagellates (Simek *et al.*, 1995; Šimek *et al.*, 1997). Furthermore, it has been observed that the effect of some environmental factors such as high UV has an effect lowering bacterivory (Sommaruga *et al.*, 1996). Even the presence of viruses (Sommaruga *et al.*, 1995) can be explanatory of microbial population dynamics (Rodríguez-Brito *et al.*, 2010) observed with molecular methods (Pernthaler *et al.*, 1998) rather than environmental situations. Regarding the trophic webs of the Pyrenees, fish presence can alter the trophic chain dynamics by removing key zooplanktonic species, driving cascade effects (Ventura *et al.*, 2017). Still, the full trophic web of the Pyrenees is not fully understood even in terms of what are the main components and their roles, because of the high novelty of the microbial eukaryotes inhabiting the lakes (Triadó-Margarit and Casamayor, 2012), although significant advances have been recently gathered in studies of other freshwater systems (Frenken *et al.*, 2017).



In addition to species sorting and filters, ecological drift can change the abundances of populations of similar functions (McPeck and Gomulkiewicz, 2005) through stochastic fluctuations (Chesson and Warner, 1981), although there is little experimental evidence of its consequences in nature, where competitive forces modulate species abundances (Gilbert and Levine, 2017). Furthermore, the small microbial sizes, short generation time and genetic plasticity makes microorganisms capable of rapid adaptation to new situations generating specialized populations which could eventually lead to speciation (Caro-Quintero *et al.*, 2011; Ortiz-Álvarez *et al.*, 2015; Llorens-Marès *et al.*, 2017). Ultimately, speciation is the ultimate generator of species diversity, and because of the isolation of the Pyrenees metacommunity, speciation may generate microbial endemisms as observed in plant (Manzano *et al.*, 2017) or animal communities (Catalan *et al.*, 2006). In fact, using plant communities (Lauga *et al.*, 2009), the Pyrenees has been defined as one of the five centers of endemism of alpine biomes. There are many terrestrial species with more than one genetic lineage in the system, translating into two glacial refuge or differentiation. As a reference, the much larger Alps region has at least four (Schmitt, 2017). The effects of speciation and ecological drift, although acknowledged, lie beyond the scope of this dissertation.

### 2.3 Spatial trends of temporal processes in the Pyrenees

Traditionally, studies have focused on quantifying and taxonomically characterize how much diversity is present; this is valuable and needed, however under this spatial framework diversity estimates are the question not the answer. The ecological processes and dynamics shape community assembly of a complex interacting microbiota, under the idiosyncratic characteristics of Pyrenean geography and geology. Taken together all the processes described so far into a single ecological force, ecological succession is the process by which communities change through time in terms of diversity (Margalef, 1963; Odum, 1969). Species diversity is inevitably linked to ecosystem dynamics and processes shaping the assembly of communities, and therefore, studying the spatial distribution of local diversities is an opportunity to explore in the space processes that depend on time (Walker and del Moral, 2003). Recent observations in the Pyrenees have indicated an alkalinity gradient that drives community differentiation in bacterioplankton communities (Barberán and Casamayor, 2014), eukaryotic plankton (Triadó-Margarit and Casamayor, 2012) and Archaea (Auguet and Casamayor, 2013). However these studies focused

only in a few lakes, which were not fully representative of the whole region. The current technological possibilities available make suitable the analysis of large amounts of samples in an ecological context, comprising the whole regional extent and avoiding focusing solely in specific populations. This PhD dissertation adds to the intensive and extensive limnological history of the Pyrenees, through describing the microbial taxonomic components and revealing potential ecological processes behind the different assemblies at an unprecedented regional lake district level.



# Rationale and main objectives



The main objective of this PhD dissertation was to give a regional perspective on the microbial ecology of the high mountain lakes of the Pyrenees using high-performance tag-sequencing techniques on more than 300 lakes distributed along the region. Through the extensive description of taxonomic components (Archaea, Bacteria and Eukarya) and spatial patterns (regional scale) we aim to infer the ecological processes behind community assembly of lake surface communities. The temporal scales acting in the Pyrenees shall be considered, despite the sampling strategy was solely spatial. For this, we did first a succession meta-analysis to understand how diversity metrics can indicate successional trends (**Chapter 1**). This meta-analysis also justifies the strategy of sampling during the most stable lake situation, when stratification was possible (although the low lake depth in some cases prevented stratification). The assembly of idiosyncratic microbiota of such remote and oligotrophic waters will be studied through multiple different lenses. From the regional variability of local communities and habitat sources of aquatic populations (**Chapter 2**), or the variability in occurrence and environmental preferences of unknown groups (**Chapter 3**), to the differences between local and regional diversities (**Chapter 4**). We further aim to unveil microbial novelty and diversity (**Chapters 2, 3 and 4**) and how this diversity interacts in the Pyrenean lakes (**Chapter 5**), highlighting the potential of combining network analysis with tag-sequencing to unveil the microbial ecology of freshwater lakes. The expected outcome is a pioneer step to understand local community variability of the three domains of life (Bacteria, Archaea and Eukarya) within a regional scale and hopes to inspire further biogeographic studies in other alpine regions along the globe. A closer description of objectives per chapter follows:

## CHAPTER 1

We hypothesized predictable changes in microbial community succession in terms of taxonomic diversity, community composition and functional traits equivalent to plants and animals. In other words, to detect if communities exhibit specific characteristics associated to their maturity. For that purpose we gathered data from 121 16S rRNA gene libraries studies covering seven distinct habitats (gut-associated, plant-associated, soils, river biofilm, microbial mats, and saline lakes). We aimed to compare communities within each study across ‘early’ versus ‘late’ stages of succession, regardless of the different system timespans. We also compared how specific community-weighted microbial traits varied between early and late succes-

sional stages, under the assumption that some functional attributes should consistently become more or less important during different stages of succession, regardless of the habitat in question.

## CHAPTER 2

We hypothesized a role of dispersal and environmental filtering on the assembly of bacterial communities. Extrapolating environmental sources from their closest relatives, we intend to understand the different bacterial sources within the lake, therefore the dispersal process. Then, we assessed if sources of diversity were constrained by any environmental factors to find the main drivers of community assembly in the regional scale, therefore environmental filtering. Finally, we analyzed how communities arranged in the geographic region using alpha and beta-diversity metrics as community descriptors of historical processes.

## CHAPTER 3

According to previous studies at the local level, we hypothesized a high occurrence and diversity of archaeal assemblages at the regional scale in alpine lakes and we aimed to gain additional ecological understanding of their specific distributions. To elucidate their ecological role, we characterized their genetic diversity, occurrence, environmental preferences and habitat breadth. Specific groups of archaeal populations have been previously the target of studies in the Pyrenees through molecular techniques but specifically, we focused on groups with fewer references and more unknown roles, such as those from the DPANN group. Ultimately our study attempts to shed light on two previously undetected groups: Woesearchaeota and Pacearchaeota, and discuss their ecology within the context of the latest genomic surveys and their potential genomic traits.

## CHAPTER 4

According to the trends observed in microbial ecology studies, we hypothesized a larger diversity of microbial eukaryotic diversity in lake surfaces, than previously reported by traditional methods during many years of limnological studies in the Pyrenees. We evaluated their genetic novelty

through the V9 region of the 18S rRNA gene grouping them into taxonomic levels that respond to similar life forms and trophic roles. Furthermore, eukaryal groups were characterized by its genetic diversity and its geographic diversification (i.e. Beta-diversity). Altogether, we focused on the heterotrophic and parasitic part of the diversity, a fraction harder to access through microscopic assessments, to raise general conclusions on their diversity patterns, ecological structure and food-web roles.

## CHAPTER 5

We hypothesized a role of biotic interactions in microbial community assembly. We aimed at describing how the whole biosphere of the Pyrenees is interacting and to seek for potential ecological rules behind those interactions. We first attempt to separate biotic interactions from environmental or geographic dispersal coincidence to retrieve positive and negative interactions between bacteria, interactions between eukaryotes and interactions between bacteria and eukaryotes. Through network analysis we tested if different sources of bacteria may show differences on how they interact with each other. Relationships between different eukaryotic groups were assessed, with a focus on the potential relevance of fungi in the lake trophic chain. We implemented an approach that estimates network properties to differentiate and compare networks of co-occurrences and co-exclusions. Through node properties we evaluated microbial characteristics and their ecological significance, and their interaction potential in the Pyrenees context: a mixture of environmental sources of bacteria and different trophic roles of eukaryotes.





# General methods and sampling area

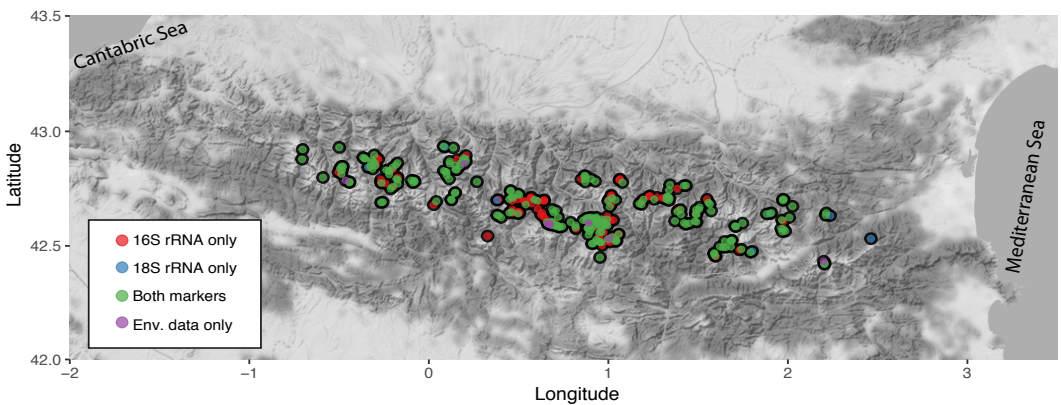


## General methods

This section provides a general description of the study system and an overlook on the sequencing results to highlight two fundamental approaches of this work, briefly introducing the environmental idiosyncrasies of the system and the contextual framework of the Pyrenees system.

### 1. Sampling and system details

Pyrenean alpine and montane lakes were sampled in July 2011 along a longitudinal gradient that ranged from longitudes  $0^{\circ}42'21.6''$  W to  $2^{\circ}27'45.9''$  E and latitudes  $42^{\circ}25'10.0''$  N to  $42^{\circ}56'06.4''$  N (Figure 1). The sampling design (a total of 320 lakes sampled) covered lakes of different sizes (range 0.1-56.9 ha) located along an altitudinal gradient (range 1459-2990 meters above sea level). The Pyrenean alpine catchments were composed of bare rock, scree, and alpine meadows, with 40% soil coverage on average. The predominant basin bedrock in the sampled area was granodiorite, with the presence of limestone and metamorphic rocks (schist and slate) being typically very poor in nutrients (Camarero and Catalan, 2012).



**Figure 1.** Location of sampled lakes across the Pyrenees mountain range. Lakes are colored by the analysis performed and results retrieved.

Regarding the geographic characteristics of the region, the Pyrenean range is mostly displayed in a longitudinal axis, although it is worth noticing that the longitudinal Pyrenean axis has a north-west to a south-east position. This relationship implies that there is a cross-effect: more western locations are also more northern ones (Figure 2). Therefore, when attribu-

ting an effect to longitude, it is possible that latitude also has a weight (and vice-versa). Also, the lake position in the range can be in the northern slope or the southern slope. This factor has a potential cross-effect with altitude: lakes in the northern slope are located at a lower altitude than lakes in the southern slope (del Castillo Jurado, 2003). Because in the west and east sides there are a higher proportion of lakes in the northern slope, altitude tends to be lower when the distance to the center (c. 0.6°E) is higher. Furthermore, the difference in altitude between the southern and northern slopes is not captured by latitude because of the oblique position of the Pyrenean axis. This information is relevant when discussing results related to geographic variables, since these cross-effects may be relevant.

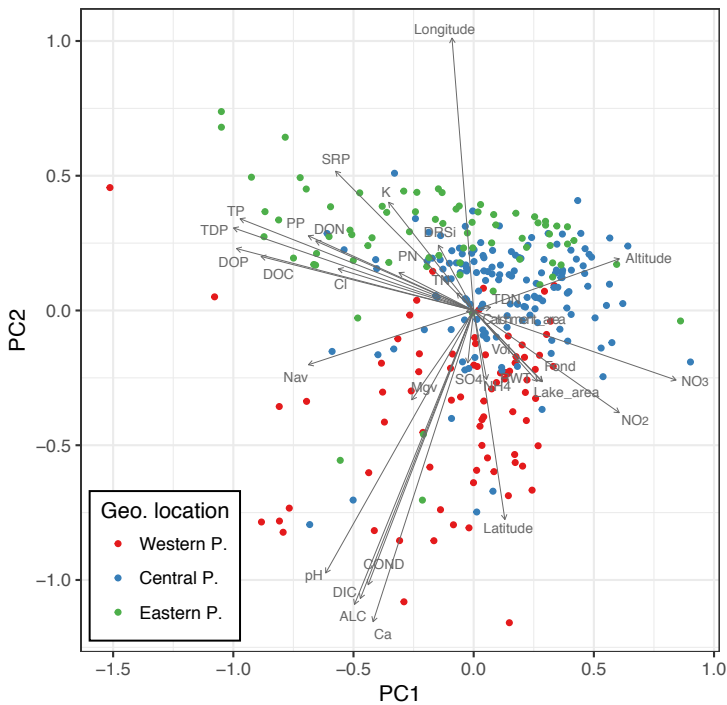
## 2. Environmental gradients and chemical variability

This thesis comprises the analysis of 30 environmental descriptors, agglomerating chemistry, ionic content, nutrients, and morphometric characteristics (Table 1).

Variable	Mean	SD	Min	Max
Lake area (ha)	4.85	6.3	0.1	56.9
Altitude (m)	2306.49	235.78	1459	2990
FOND	15.83	11.78	7	113
Vol	0.7	2.61	0	30.02
RWT	4.48	8.74	0	69.8
Total catchment (ha)	185.95	361.47	4	5437.9
pH	7.18	0.72	4.45	10.1
ALC	245.74	267.27	-35	1530
COND	33.06	30.47	3.9	168.8
DIC	2.97	3.1	0.1	17.5
DOC	0.92	1.02	0	8.7
SRP	14.75	9.88	3	74
TDP	69.43	53.4	3	484
TP	122.09	98.59	14	794
DOP	54.72	48.34	0	431
PP	52.64	59.84	0	568
NH4	0.42	0.4	0	4.2
NO2	0.08	0.05	0	0.22
NO3	5.95	5.48	0	23
TDN	12.51	6.54	3	77
TNv	14.62	7.06	4	77
DON	6.05	5.92	0	76
PN	2.21	2.72	0	25
DRSi	34.14	25	2	235
Clv	5.11	2.86	0	35
SO4	56.63	137.64	5	1339
Na	25.64	14.09	2	97
K	3.01	3.68	0	27
Ca	257.35	258.24	15	1388
Mg	29.48	71.47	2	595

**Table 1.** Environmental descriptors of the Pyrenean dataset. Mean values, standard deviation (SD), maximum and minimum values for 30 chemical and morphometric variables.

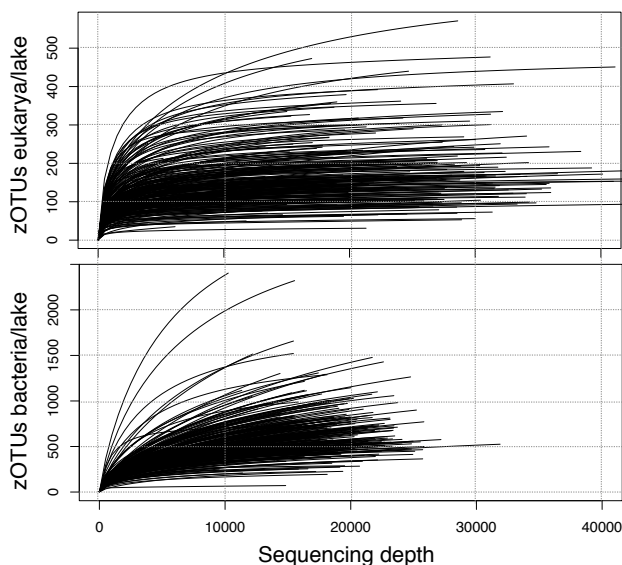
These have been estimated with standard procedures as described (Camarero et al., 2009). These variables are representative of the chemical and morphometric variability of the high mountain lakes of the Pyrenees mountain region. Through a principal components analysis (Figure 2) the main environmental gradients of the data are identified: the strongest gradient is related to pH and similar variables (COND, DIC, ALC, Ca), which was also associated to geography (more alkaline lakes with increased latitudes). The second strongest gradient was associated to a trophic status (variables representing carbon and phosphorus content opposite to nitrates and nitrites). The third gradient (third component, not shown), was associated to morphometric variables (lake area, fondaria, RWT). For further details, see a similar analysis performed in chapter 2.



**Figure 2.** Principal components analysis (PCA) performed with the 30 environmental variables (log-transformed and standardized) and the geographic coordinates of 320 lakes. The diagram describes the lineal combination of variables showing the main environmental gradients (components PC1 and PC2 represented 37% of the total environmental variance. With the third component, PC3, the analysis accumulates 49% of the environmental variance). The different colors indicate the geographic position.

## 2. Sequencing and on quantifying microbial diversity

Surface water samples were directly collected close to the outflow of the lakes, sampling c. 200 mL in 0.22-micron Sterivex pressure driven sterile filter units (Millipore R), pre-filtered through a pore size of 50  $\mu$ M, avoiding disturbance of the littoral zone. All lakes with a valid water sample were attempted for sequencing (320 lakes) although, sequencing failed in a different number of lakes for Bacteria (304 lakes successful) and Eukarya (227 lakes successful). Geographic distribution of available raw data is shown in Figure 1.



**Figure 3.** Rarefaction curves of lake samples for eukarya (upper panel) and bacteria (lower panel).

Microbial communities were sequenced through the Illumina MiSeq platform targeting the 16S rRNA and 18S rRNA partial gene sequences. These genomic regions have been the main tool to study microbial diversity, therefore these provide accurate classifications and descriptives. Specificities regarding the sequencing of bacterial, archaeal and eukaryal communities are extensively described in chapters 2, 3 and 4 respectively. Through a visualization of bioinformatically-filtered sequences, we observed that local communities were well sequenced (Figure 3), especially in the case of Eukarya. However, based on rarefaction curves of local communities, slopes

do not completely show a 'plateau' in the bacterial communities, suggesting that there is even more potential diversity per lake that was not detected (Figure 3). However, when combining the lakes into the regional pool of species, rarefaction curves indicate that there is not much bacterial diversity to further detect in the region under the given sampling protocol.

A variety of statistical methods and visualization tools are detailed in the methodology sections for each article / chapter. In brief, we included estimations of alpha diversity (richness (R), Shannon-Weaver (H') and phylogenetic diversity (PD) as metrics of microbial diversity), beta-diversity (i.e., the degree of heterogeneity of a given group of communities), multivariate ordinations and statistics, and network analysis. Most analysis and visualizations were conducted under the R-project language for statistical computing (R Core Team, 2017), through functions implemented in a wide variety of packages, or stated otherwise.





# Chapter 1

Consistent changes in the taxonomic structure and functional attributes of bacterial communities during primary succession

Ortiz-Álvarez, R., Fierer, N., de los Ríos, A., Casamayor, E. O. & Barberán, A. Consistent changes in the taxonomic structure and functional attributes of bacterial communities during primary succession. *ISME J.* (2018). doi:10.1038/s41396-018-0076-2 (IF: 9.664, ranking Q1 in Microbiology - 7/125- and in Ecology-3/153-)



Ortiz-Álvarez R, Fierer N, De los Ríos A, Casamayor EO, Barberán A. Consistent changes in the taxonomy and functional attributes of bacterial communities during primary succession. *ISME-J*. Vol. 12 (2018) : 1658–1667.

<http://dx.doi.org/10.1038/s41396-018-0076-2>

<https://www.nature.com/articles/s41396-018-0076-2>

Received: 18 September 2017 / Revised: 11 January 2018 / Accepted: 13 January 2018

© International Society for Microbial Ecology 2018

## **Abstract**

Ecologists have long studied primary succession, the changes that occur in biological communities after initial colonization of an environment. Most of this work has focused on succession in plant communities, laying the conceptual foundation for much of what we currently know about community assembly patterns over time. Because of their prevalence and importance in ecosystems, an increasing number of studies have focused on microbial community dynamics during succession. Here, we conducted a meta-analysis of bacterial primary succession patterns across a range of distinct habitats, including the infant gut, plant surfaces, soil chronosequences, and aquatic environments, to determine whether consistent changes in bacterial diversity, community composition, and functional traits are evident over the course of succession. Although these distinct habitats harbor unique bacterial communities, we were able to identify patterns in community assembly that were shared across habitat types. We found an increase in taxonomic and functional diversity with time while the taxonomic composition and functional profiles of communities became less variable (lower beta diversity) in late successional stages. In addition, we found consistent decreases in the rRNA operon copy number and in the high-efficient phosphate assimilation process (Pst system) suggesting that reductions in resource availability during succession select for taxa adapted to low-resource conditions. Together, these results highlight that, like many plant communities, microbial communities also exhibit predictable patterns during primary succession.

# Chapter 2

A complex integration of immigration and local pH shapes regional surface lake bacterial community assembly in the Pyrenean lacustrine district



Submitted paper. Embargoed until publication date





# Chapter 3

High occurrence of Pacearchaeota and Woesearchaeota (Archaea superphylum DPANN) in the surface waters of oligotrophic high-altitude lakes

Ortiz-Alvarez, R. & Casamayor, E. O. High occurrence of Pacearchaeota and Woesearchaeota (Archaea superphylum DPANN) in the surface waters of oligotrophic high-altitude lakes. *Environ. Microbiol. Rep.* 8, 210–217 (2016). (IF: 3.363, ranking Q2 in Microbiology)



Ortiz-Alvarez R, Casamayor, E. High occurrence of Pacearchaeota and Woesearchaeota (Archaea superphylum DPANN) in the surface waters of oligotrophic high-altitude lakes. *Environmental Microbiology Reports*. Vol. 8, Issue 2 (April 2016) : 210-217.

<http://dx.doi.org/10.1111/1758-2229.12370>

<https://onlinelibrary.wiley.com/doi/abs/10.1111/1758-2229.12370>

First published: 28 December 2015

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

## Abstract

We carried out a regional survey on the archaea composition from surface waters of > 300 high-altitude Pyrenean lakes (average altitude 2300 m, pH range 4.4–10.1) by 16S rRNA gene tag sequencing. Relative Archaea abundances ranged between 0% and 6.3% of total prokaryotes amplicons in the polymerase chain reaction (PCR) mixture, and we detected 769 operational taxonomic units (OTUs; grouped at 97% identity) that split into 13 different lineages, with altitude and pH having a significant effect on the community composition. Woesearchaeota and Pacearchaeota (formerly Euryarchaeota DHVEG-6 cluster) dominated the data set (83% of total OTUS), showed a high occurrence (presence in c. 75% of the lakes) and had relative abundances significantly and positively correlated with the phylogenetic diversity of bacterial communities. Micrarchaeota–Diapherotrites (formerly Euryarchaeota MEG cluster), Methanomicrobia, Thermoplasmata and ammonia-oxidizing thaumarchaeota (AOA) showed relative abundances between 1% and 3% and occurrences between 14% and 26%. Minor lineages were SM1K20, Aenigmarchaeota (formerly Euryarchaeota DSEG cluster), Methanobacteria, Bathyarchaeota and SCG. Environmental preferences substantially differed among lineages, with Aenigmarchaeota and Methanomicrobia having the largest habitat breadth, and Thermoplasmata, AOA and Micrarchaeota having the smallest. Pacearchaeota and Woesearchaeota had been mostly reported from saline habitats and sediments, but surface waters of oligotrophic alpine lakes are suitable environments for such ecologically spread and genetically diverse archaeal lineages.

# Chapter 4

High planktonic diversity in mountain lakes contains similar richness of autotrophic, heterotrophic and parasitic eukaryotic lifeforms

Ortiz-Álvarez, R., Triadó-Margarit, X., Camarero, Ll., Casamayor, E.O. & Catalán, J. 2018. High planktonic diversity in mountain lakes contains similar richness of autotrophic, heterotrophic and parasitic eukaryotic lifeforms. *Scientific reports*. Accepted. (IF: 4.259, ranking Q1)



# SCIENTIFIC REPORTS

## OPEN High planktonic diversity in mountain lakes contains similar contributions of autotrophic, heterotrophic and parasitic eukaryotic life forms

Received: 18 November 2017

Accepted: 1 March 2018

Published online: 13 March 2018

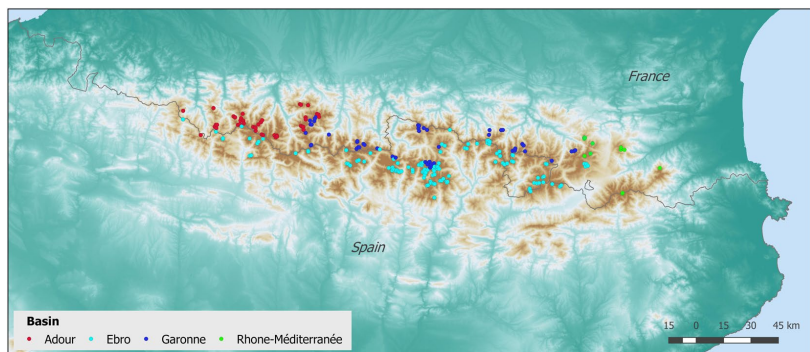
Rüdiger Ortiz-Álvarez<sup>1</sup>, Xavier Triadó-Margarit<sup>1</sup>, Lluís Camarero<sup>1</sup>, Emilio O. Casamayor<sup>1</sup> & Jordi Catalan<sup>2</sup>

A rich eukaryotic planktonic community exists in high-mountain lakes despite the diluted, oligotrophic and cold, harsh prevailing conditions. Attempts of an overarching appraisal have been traditionally hampered by observational limitations of small, colorless, and soft eukaryotes. We aimed to uncover the regional eukaryotic biodiversity of a mountain lakes district to obtain general conclusions on diversity patterns, dominance, geographic diversification, and food-web players common to oligotrophic worldwide distributed freshwater systems. An unprecedented survey of 227 high-altitude lakes comprising large environmental gradients was carried out using Illumina massive tag sequencing of the 18S rRNA gene. We observed a large Chrysophyceae dominance in richness, abundance and novelty, and unveiled an unexpected richness in heterotrophic phagotrophs and parasites. In particular, Cercozoa and Chytridiomycota showed diversity features similar to the dominant autotrophic groups. The prominent beta-dispersion shown by parasites suggests highly specific interactions and a relevant role in food webs. Interestingly, the freshwater Pyrenean metacommunity contained more diverse specific populations than its closest marine oligotrophic equivalent, with consistently higher beta-diversity. The relevance of unseen groups opens new perspectives for the better understanding of planktonic food webs. Mountain lakes, with remarkable environmental idiosyncrasies, may be suitable environments for the genetic diversification of microscopic eukaryotic life forms.

Hidden species richness in ultraoligotrophic freshwaters ecosystems has long been suspected as early microscopic assessments indicated abundant soft and rather small eukaryotic organisms<sup>1</sup>. In ecological studies, many species of recognized relevant groups (e.g. Chrysophyceae) were only tentatively determined or, quite often, lumped into size classes, due to the lack of a consistent taxonomy<sup>2–5</sup>. Recently, the introduction of molecular techniques is progressively unveiling a great eukaryotic diversity in mountain lakes<sup>6–8</sup>. The use of molecular tools for identification of sequences originating from environmental DNA by reference to sequence databases can overcome many limitations of traditional microscopic approaches<sup>9</sup>, although they may lack specificity to link morphology to molecular data<sup>10</sup>. So far, the molecular diversity of protists in inland waters appears higher than that of the morphospecies and cultivated species catalogued in public databases<sup>11–13</sup>.

Despite remarkable traditional attempts to develop a morphological taxonomy (e.g. studies by H. Skuja, A. Pascher, K. Starmach, P. Bourrelly among others), the realistic appraisal of the species richness across oligotrophic lakes was hampered by technical limitations. The relevance of scarcely visible organisms, usually heterotrophs, has been a matter of speculation or they have been just simply ignored. Heterotrophic flagellates (HNF) were usually amalgamated in a single functional guild, and amoeboid forms were rarely considered and, if so, with limited taxonomic resolution<sup>14,15</sup>. Recent findings in marine waters have revealed large unseen diversity of heterotrophic

<sup>1</sup>Integrative Freshwater Ecology Group, Center for Advanced Studies of Blanes-CSIC. Acc. Cala St Francesc 14, E-17300, Blanes, Catalonia, Spain. <sup>2</sup>CREAF - CSIC, Campus UAB, Edifici C, 08193, Cerdanyola del Vallès, Catalonia, Spain. Correspondence and requests for materials should be addressed to E.O.C. (email: [casamayor@ceab.csic.es](mailto:casamayor@ceab.csic.es))



**Figure 1.** Distribution of 227 lakes surveyed and sequenced across the Pyrenees. The main river basins are indicated. Dots of nearby lakes may overlap. Map created with ArcGIS 10.1 (©ESRI).

protistan groups<sup>16</sup>, and many studies are highlighting the role of fungi in aquatic food webs as a promising research topic<sup>9,11,17</sup>.

Oligotrophic lakes are widely heterogeneous regarding climatic, chemical and morphological conditions. This heterogeneity sustains the proposal that freshwater microbial diversity might be higher than in marine environments<sup>18–20</sup>. An appraisal of the general eukaryotic diversity in oligotrophic planktonic systems requires an extensive survey of sites comprising several environmental gradients. Pyrenean mountain lakes provide such an environmental variation because of their distribution across altitude, and bedrock diversity<sup>21</sup>. Their abundance in a relatively restricted territory also permits considering patterns related to the spatial distribution and connectivity between lakes. In the present investigation, we have characterised the eukaryotic regional metacommunity of the mountain lakes of the Pyrenees and evaluated its novelty using the V9 region of the 18S rRNA gene as a taxonomic indicator. Diversity, dominance, and geographic diversification of the main groups were examined. We grouped operational taxonomic units (OTUs) in taxonomic levels that responded to similar life forms and trophic roles. Thus, some general conclusions on diversity patterns, ecological structure and food-web players were unveiled.

## Materials and Methods

**Lakes survey.** The lakes sampling design covered the whole geographical distribution of the mountain lakes in the Pyrenees, from 0°42′21.6″W to 2°27′45.9″E longitude and 42°25′10.0″N to 42°56′06.4″N latitude (Fig. 1). The 227 sampled lakes were distributed across altitudinal gradient (1459–2990 m above sea level; average 2300 m) and different bedrock types to encompass the main environmental gradients influencing the species distribution in mountain lakes<sup>21</sup>. Lake area ranged between 0.1 and 57 ha (average 4.6 ha). We intended to sample all lakes within a selected sub-basin to accurately mimic the lake size, and hydrological connectivity distributions. Most of the lakes were located in the alpine belt where bare rock and scree predominated, with 40% soil coverage by meadows and shrubs on average. The predominant bedrock was granodiorite, which shaped highly diluted waters (average conductivity 33  $\mu\text{S}/\text{cm}$ ). Additional areas of limestone and metamorphic rocks (schist and slate) provided a broad range of chemical conditions (e.g., pH range 4.4–10.1, average 7.2). Typically, the lakes were poor in phosphorus (average soluble reactive phosphorous 73 nM), and mostly ranked as oligotrophic<sup>22</sup>. Plankton samples were collected in the littoral zone along the summer period to keep the temporal interval the shortest as possible and to preserve comparability among samples. About 200 mL <50  $\mu\text{m}$  size water lake samples were filtered *in situ* using 0.22-micron Sterivex pressure driven sterile filter units (Millipore R), and preserved in lysis buffer (40 mM EDTA, 50 mM Tris, pH 8.3, 0.75 M sucrose) as recently reported<sup>23</sup>. For the genomic DNA extraction, the membranes were enzymatically digested with lysozyme, proteinase K and sodium dodecylsulfate incubation, followed by phenol extraction, and DNA purification and concentration with Amicon® Ultra 4 Centrifugal Filter Units – 100000 NMWL (Millipore).

**Sequencing and data filtering.** High-speed multiplexed 18S rRNA gene sequencing with the Illumina MiSeq System ( $2 \times 150$  bp) was carried out with the “universal” 1391 f and EukBr eukaryotic primers (V9 region)<sup>24</sup>. These primers are known to recover the known eukaryotic diversity without major qualitative or quantitative biases<sup>16,25</sup>. Guidelines from the Earth microbiome project (EMP) protocol, and the genomics core facilities and methods of the RTSF-MSU (Michigan State University, USA) were followed. Raw sequences were analysed with UPARSE<sup>26</sup>. Overall quality of sequences was high, with 77.6% reaching Q30. The total number of sequences before quality filtering was 12,288,923. After merging of read pairs and filtering by read length (above or equal to 161 pb), quality-score distribution (ASCII ‘B’) and an expected error of 0.25, we kept 6,384,407 sequences. Out of them 921,749 were unique and 205,694 not singletons. The latter were clustered at 97% identity, after chimera removal, resulting in 6398 OTUs. Sequences were mapped back into OTUs and classified with the SILVA 119 database<sup>27</sup>. Overall, we retrieved a mean of 24,428 sequences per lake (range 10677–45940).

Group	Regional parameters				Local parameters		
	Closest match (%)	Mean OTU occurrence	$\alpha$ -diversity (OTU Nr)	$\gamma$ -diversity ( $H'$ )	$\alpha$ -diversity (OTU Nr)	$\alpha$ -diversity ( $H'$ )	BP dominance (%)
Diatomea	97.48	4.29	93	1.98	4.01	0.62	0.58
Cryptomonadales	96.95	6.47	93	1.36	6.04	0.6	0.54
Ciliophora	96.89	2.96	105	2.64	3.12	0.7	0.33
Kathablepharidae	96.89	23.46	24	0.86	5.65	0.47	0.79
Chlorophyta	96.87	3.51	378	3.26	11.86	1.17	0.12
Dinoflagellata	95.60	6.25	422	2.79	26.53	1.4	0.21
Chytridiomycota	93.71	1.92	710	4.93	13.73	1.7	0.1
Chrysophyceae	93.63	9.19	818	3.88	75.53	2.29	0.12
Choanomonada	92.48	3.83	103	2.41	3.96	0.58	0.45
Cercozoa	91.98	2.49	542	3.81	13.56	1.51	0.16

**Table 1.** Regional (lacustrine district) and local (per lake) diversity indexes. The mean closest match identity and OTU occurrence, Alpha and Gamma diversities estimated by the total number of OTUs (Richness, OTU Nr), Shannon diversity index ( $H'$ ), and Berger-Parker dominance index (BP dominance) are shown.

**Ecological classification and diversity indexes.** Sequences were classified using the RDP classifier by default as implemented in Qiime<sup>28</sup> and the Silva-NGS pipeline with the Silva 119 database<sup>29</sup>. Relevant discrepancies were double-checked through local BLAST (search September 2015), and if not resolved left as “Unclassified”. Around 7% of the OTUs could not be classified beyond “domain”. Mean number of OTUs per lake (Richness) and average occurrence of OTUs were determined. Shannon diversity<sup>30</sup> per lake (alpha diversity), Shannon diversity in the whole region (gamma diversity) and Berger-Parker dominance<sup>31</sup> indexes were calculated.

For classification *ad hoc* taxonomical levels reflecting the main ecological life-styles were used. Furthermore, preferences for carbon source, nutrition type and habitat were labelled. Thus, groups corresponded to distinct taxonomic hierarchical levels but shared the common characteristic that most of the species within the group showed a similar life form and functional role within the planktonic food web. For the first ten most relevant groups, we evaluated in detail more different diversity patterns (Table 1).

**Genetic novelty analysis.** The novelty in the dataset was explored by 18S rRNA gene BLAST identity searches against GenBank sequences (search September 2015) to both the closest environmental match (CEM) and the closest cultured match (CCM) available in GenBank (e.g., Massana *et al.*<sup>32</sup>). Only OTUs sharing sequence identity values and alignment coverage values above 80% to nt database were considered for downstream analyses. The closest match identity values (either to environmental or cultured) were used to explore novelty-abundance, and novelty-occurrence relationships. High novelty was defined for those OTUs with <97% identity for both CCM and CEM (see more details in Supplementary information S1).

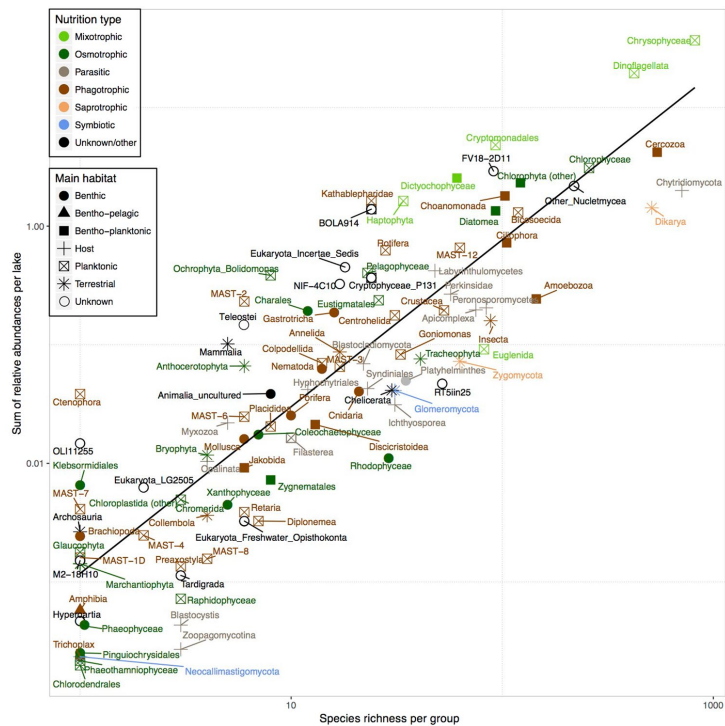
**Geographic patterns and beta-dispersion.** The relationship between geographic Euclidean distance and community dissimilarity (Bray-Curtis) was evaluated using Mantel tests (R package vegan) with the Spearman method both for the whole eukaryotic assemblage and for taxonomic groups of interest. We assigned lakes to basins and sub-basins using the ‘Spatial Join’ routine from ArcGIS 10.1 (©ESRI). We estimated average beta-dispersion (i.e., average distance to group centroids) (R package vegan) per catchment or sub-catchment, and tested it with ANOVA (999 permutations) using the taxonomic groups as levels of a factor ‘taxonomy’. If significant, *post hoc* pairwise comparisons or HSD Tukey tests were carried out to find differences between groups (R package agricolae). The coherence between pairs of taxonomic groups regarding how beta-dispersion changed in sub-basins was assessed using Spearman correlations (R package base).

## Results

The total number of properly identified OTUs reached 6086 (mean 303 per lake, range 82–715), comprising 105 differential groups (Fig. 2). Most of the metacommunity richness corresponded to planktonic (37%) and benthic-planktonic (20%) free organisms, followed by host-dependent organisms (18%). The remaining component were putative zoospores, diaspores and traces of predominantly terrestrial (11%) and benthic (3%) organisms. The most abundant were autotrophic organisms (73%), but most of them (63% of the total relative abundance) were mixotrophs –i.e., autotrophs that also ingest prokaryotes– and only a minority (10%) exclusively osmotrophs. Phagotrophic heterotrophs accounted for 12% of the reads, while other heterotrophic life-forms were less abundant (c. 3–0.1%). However, most of the OTUs richness belonged to heterotrophs: phagotrophs (26%), parasites (18%), saprotrophs (9%), and symbiotic (0.5%). Autotrophic organisms accounted for 36% of the richness (mixotrophs, 25%; exclusively osmotrophs, 11%) and trophic mode was unknown for the remaining OTUs. Overall, there was a general positive relationship between abundance and richness of the groups (Fig. 2).

Taxonomically, the regional pool was dominated by phytoplanktonic mixotrophic flagellates of Chrysophyceae (37% of the sequences and 818 OTUs), followed by Dinoflagellata (20%, 422 OTUs), Cryptomonadales (5%, 93 OTUs), and Haptophyta (2%, 34 OTUs). Among exclusively osmotrophic phytoplankton, Chlorophyta were the most abundant and diverse (5% and 378 OTUs), particularly Chlorophyceae, including also many flagellated forms. Diatomea showed relative abundance slightly above 1% of the sequences with 93 OTUs.

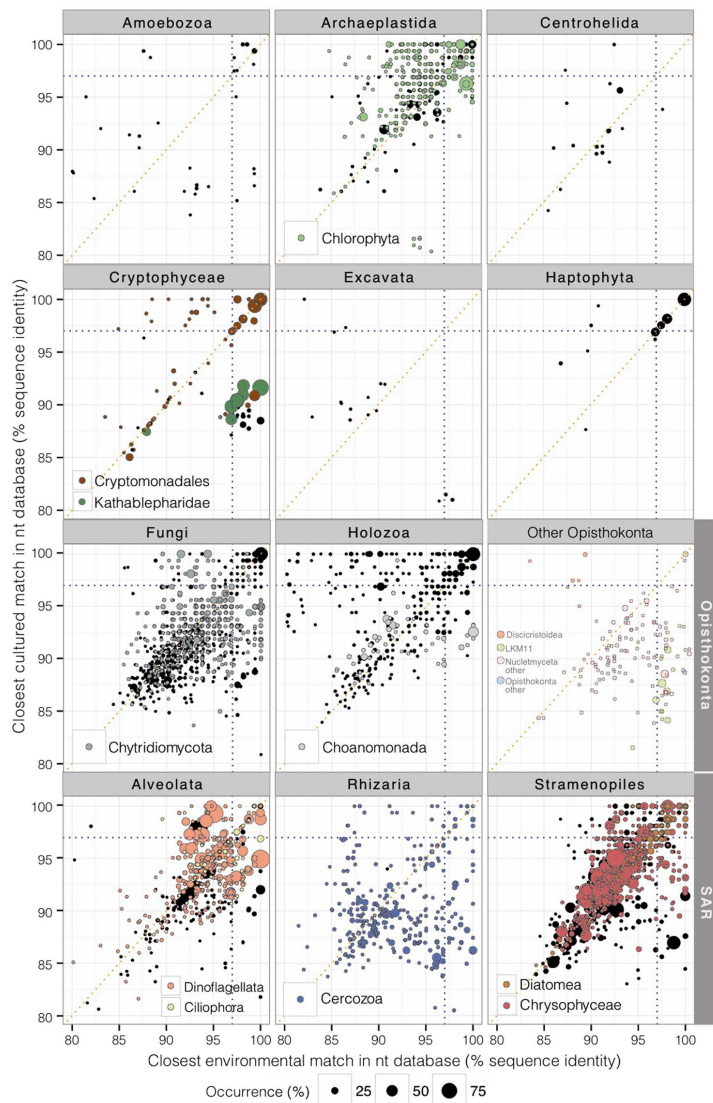




**Figure 2.** OTU richness and cumulative relative abundance of the main groups found in the plankton of Pyrenean lakes in logarithmic scale. Colours indicate dominant nutrition type while shapes indicate their main habitats. Autotrophs are split into two groups: those purely osmotrophs and those also able of phagotrophic ingestion (mixotrophs). A fitted linear model is drawn.

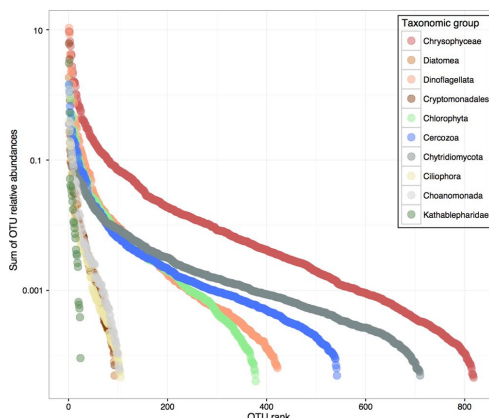
Regarding the motile heterotrophic components of the samples, the pseudopodia-feeding group Cercozoa (Rhizaria) with 4.2% of the sequences and 542 OTUs was the most abundant and diverse. Other groups above 1% of the sequences were Choanomonada, Kathablepharidae, and Bicosoecida. Ciliophora were slightly less abundant than these groups but with substantial richness. In contrast, Kathablepharidae and Cryptophyceae showed poor OTUs richness. Among mycoplankton, the mostly parasitic fungi Chytridiomycota (2% relative abundance) showed high richness (710 OTUs). Mainly terrestrial Dikarya, including Ascomycota and Basidiomycota (511 OTUs), and Zygomycota (63 OTUs) were probably present as dispersal stages.

**Eukaryotic genetic novelty in the Pyrenean lacustrine district.** As a common trait, large proportions of highly novel representatives were observed for the different taxa. The main major groups SAR (Stramenopiles-Alveolata-Rhizaria), Opisthokonta, and Archaeplastida, showed 81%, 74% and 54% of the OTUs within the highest novelty, respectively (Fig. 3). This value raised to >90% in the minor groups Centrohelida and Excavata. SAR had the largest diversity of both autotrophic (Chrysophyceae) and heterotrophic (Cercozoa) groups and also the greatest novelty. Chrysophyceae were mainly recovered within the highest novelty plot area, equally distant from environmental and cultured references even for those OTUs showing the highest occurrence. Numerous Rhizaria, which essentially corresponded to Cercozoa, were far below the 97% threshold. Within Opisthokonta, it was remarkable the extremely high novelty in Fungi. Indeed, most of the Chytridiomycota OTUs, true aquatic fungi, were in the highest novelty area of the plot and the mean value of the closest match was substantially low (Table 1). A smaller set of OTUs related to potential basal fungi (LKM11 and LKM15) previously found in environmental studies, showed low relatedness to cultured counterparts. Choanomonada also showed a great degree of novelty, although the most common OTU had been previously reported in environmental surveys. Photosynthetic Chlorophyta and Cryptophyceae showed close cultured references. Conversely, the heterotrophic group of Cryptophyceae (Kathablepharidae) lacked close cultured counterparts (<93% identity to CCM). Finally, Dinoflagellata and Haptophyta showed low novelty. In general, the relationship between the degree of genetic novelty and both abundance ( $r = 0.21$ ,  $p < 0.01$ ) and occurrence of the OTUs ( $r = 0.23$ ,  $p < 0.001$ ) was low. Thus, the most abundant and frequent OTUs were not necessarily previously known.



**Figure 3.** Novelty of OTUs as indicated by their closest cultured and environmental matches identity values. The highest novelty area is indicated, and falls below the 97% thresholds. Color displays the main groups of study, leaving other OTUs of the high taxonomic ranks in black. Dot size indicates OTU occurrence in the dataset.

**Planktonic eukaryotic richness and ecological diversity in oligotrophic waters.** Chrysophytes, Chytridiomycota, Dinoflagellata, Cercozoa, and Chlorophyta showed the highest richness, Shannon diversity  $>1$  per lake, and Berger-Parker dominance around 15% (Table 1). Lower richness and Shannon diversity (around 0.5) was observed in Kathablepharidae, Cryptomonadales, Diatomea, Ciliophora, and Choanomonada with Berger-Parker dominance  $>30\%$ . The metacommunity patterns of gamma-diversity (rank-abundance curves) slightly diverge from the previous alpha-diversity picture. The gentler the slope of the curves, the greater the



**Figure 4.** Rank abundance plot (log scale) of OTUs by taxonomic groups of interest. Cumulative sum of relative abundances determines OTU rank per group.

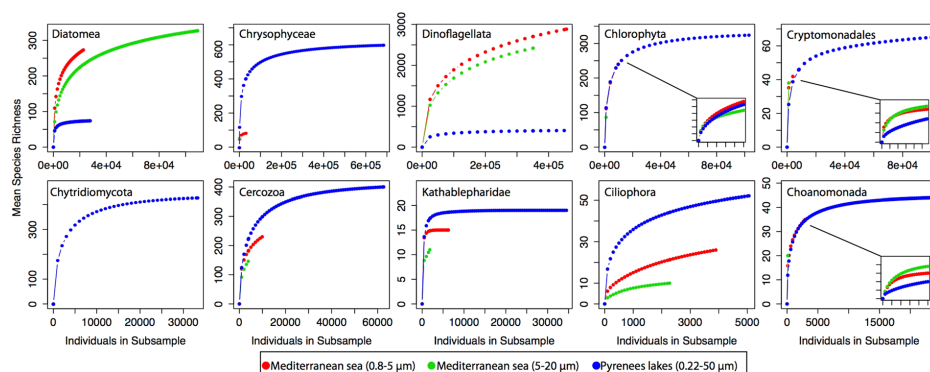
gamma-diversity (Fig. 4). Chytridiomycota, Chrysophyta, and Cercozoa were again the most diverse groups but the two heterotrophic groups showed more spatial beta-diversity than Chrysophyta. Chlorophytes and Dinoflagellata showed intermediate values and the diversity for the remaining groups was low, particularly, in Kathablepharidae.

Interestingly, the freshwater Pyrenean metacommunity contained more diverse specific populations than the closest marine oligotrophic equivalent, surface waters of the Mediterranean Sea (Tara Oceans Project<sup>16</sup>) after normalizing to equivalent sequencing efforts (Fig. 5). This result was more apparent for Chrysophyceae, Cercozoa, Katablepharidae, and Ciliophora. Conversely, Diatomea and Dinoflagellata, appear to hold richer plankton populations in the Mediterranean Sea. Similar values were found for Choanomonada, and very few chytrids were reported in the Mediterranean dataset. In addition, higher beta-diversity was found for the regional freshwater study than in the marine surface samples across the Mediterranean (see more details in Supplementary information S2).

In general, the similarity between lake communities was significantly related to the geographic distance (Mantel test, Spearman method,  $r = 0.21$ ,  $p = 0.001$ ). In particular, this relationship was significant for Chrysophyceae ( $r = 0.16$ ), Dinoflagellata, Ciliophora, and Cercozoa although the variance explained was low (Table 2). Therefore, we further analysed the diversity partition across the region considering catchments and sub-catchments. The mean beta-dispersion of basin and sub-basin communities was 0.53 ( $sd = 0.04$ ) and 0.40 ( $sd = 0.07$ ), respectively. The ANOVA of distances to group centroids was not significant for catchments ( $F = 1.89$ ,  $p = 0.12$ ). Conversely, the same analysis for sub-catchments was significant ( $F = 2.28$ ,  $p < 0.001$ ), and the *post hoc* Tukey HSD test showed that the taxonomic groups could be categorised into different levels of sub-basin beta-diversity (Table 2). The analysis highlighted the main heterotrophs with high beta-dispersion (i.e., Chytridiomycota and Cercozoa, 'a') from the main primary producers with low beta-dispersion (Chrysophyceae and Dinoflagellata 'c'), and Kathablepharidae ('d') as the group with the lowest beta-dispersion. There was a positive correlation among the beta-dispersion of most of the groups per sub-basins (Fig. S1), being especially remarkable the high correlation and degree of significance ( $P < 0.001$ ) of the Chytridiomycota with many of the other groups.

## Discussion

Historically, knowledge on protist assemblages from high mountain lakes was mainly based on ecological studies (e.g. Capblancq, 1972; Tilzer, 1972)<sup>33,34</sup>. They suspected a high diversity in some groups such as Chrysophyceae and Dinoflagellata, similarly as found in other oligotrophic systems (Nordic and Central European sites) by early taxonomical studies. Our study broadens this suspected high diversity and indicates a high degree of novelty, even in the case of Dinoflagellata with apparently closer culture counterparts especially for the most abundant OTUs. The comparison with the marine samples primarily highlights the Chrysophyceae richness in mountain lakes, whereas dinoflagellates are richer in the marine environment. Since chrysophytes lack carbon concentration mechanisms<sup>35</sup>, these are particularly suitable for these soft water environments, and we can interpret richness differences as an evolutionary divergence of the ecology between the two groups. This hidden high diversity in Pyrenean chrysophytes was already suspected based on 210 stomatocyst morphotypes found along 105 lakes<sup>36</sup>. The molecular survey carried out in the present study, suggests that the chrysophycean richness is at least four-fold larger. Interestingly, in classical treatises of Chrysophyceae morphospecies descriptions, the number raised to about one thousand (e.g., Starmach, 1985)<sup>37</sup>, though only a few of them have been cultivated. Chrysophyceae also includes heterotrophic organisms. Many of them are usually handled as *Spumella*-like



**Figure 5.** Rarefaction curves per taxonomic group of the Pyrenean pool (227 samples rarefied as a single pool) versus the Mediterranean Sea pool from the Tara Oceans project (11 samples per size fraction). Despite the variable number of samples, we equalized both pools by rarefying at equivalent sequencing depths per pool (see Supplementary information).

Group	Mantel-r	Basin beta-dispersion (mean $\pm$ SD)	Sub-basin beta-dispersion (mean $\pm$ SD)	Tukey HSD test groups
Chytridiomycota	-0.02	0.66 $\pm$ 0.03	0.52 $\pm$ 0.07	a
Cercozoa	0.07*	0.63 $\pm$ 0.04	0.49 $\pm$ 0.08	ab
Chlorophyta	0.05	0.61 $\pm$ 0.07	0.49 $\pm$ 0.08	ab
Ciliophora	0.07**	0.64 $\pm$ 0.05	0.48 $\pm$ 0.11	ab
Diatomea	-0.02	0.64 $\pm$ 0.03	0.47 $\pm$ 0.11	abc
Choanomonada	0.01	0.60 $\pm$ 0.06	0.46 $\pm$ 0.12	bc
Cryptomonadales	-0.07	0.57 $\pm$ 0.07	0.45 $\pm$ 0.11	bc
Chrysophyceae	0.16**	0.56 $\pm$ 0.05	0.41 $\pm$ 0.10	c
Dinoflagellata	0.09**	0.56 $\pm$ 0.03	0.42 $\pm$ 0.09	c
Kathablepharidae	0.04	0.47 $\pm$ 0.05	0.32 $\pm$ 0.14	d

**Table 2.** Geographic relationships statistics among group communities. Mantel test results between group communities and geographic distances (\* $p < 0.01$ , \*\* $p < 0.001$ ). Beta dispersion comparison per basin and sub-basin. Beta-dispersion was calculated per sub-basin when there were at least two lakes to compare. Statistical significance was assessed through *post-hoc* Tukey HSD for sub-basins: groups a–d are groups of analogous beta-dispersion patterns.

organisms, and c. 50 Pyrenean OTUs could be assigned to this genus. A recent study indicates that *Spumella*-like flagellates are polyphyletic and that their lack of morphological differences can be seen as a convergence to a successful live form under certain ecological circumstances<sup>38</sup>. Mountain lakes, with a dark long ice-covered period<sup>39</sup>, may be a suitable environment for the genetic diversification of this life-form.

Dinoflagellates are abundant and diverse in mountain lakes, although less than in marine samples<sup>16</sup>. The high richness is particularly concentrated in small and unarmored forms (e.g. *Gymnodinium*). As is the case for Chrysophyceae, the molecular techniques provide a tool to investigate solidly the ecology of mountain dinoflagellates, which with the few exceptions of some *Peridinium* have remained elusive to traditional taxonomy. The result obtained for diatoms conformed to the established knowledge. Most of the diversity largely corresponds to benthic forms and, therefore, the modest contribution to total planktonic diversity does not come as a surprise. Moreover, as expected, the bulk of the OTUs corresponded to centric diatoms. The match with the cultured *Fragilaria nanana* agrees with the observation of this species in mixing epilimnia<sup>40</sup>. The group Cryptophyceae (both Cryptomonadales and Kathablepharidae) showed low diversity as compared with its high abundance and widespread occurrence. This feature is not something particular to mountain lakes but general for the group. The low morphological differentiation does not hide a high molecular diversity, although molecular taxonomy revealed additional variation<sup>41</sup>. Finally, we confirmed the presence of Haptophyta (Pavlovophyceae and Prymnesiophyceae) most of them closely related to cultured counterparts (e.g. *Chrysochromulina parva*), a group with little penetration in freshwaters and that had been only sceptically reported in mountain lakes<sup>42</sup>.

Molecular approaches provide taxonomic accuracy to the ecological studies, which up-to-now was only available to a few taxonomic specialists, as in the large variety of biflagellate Archaeplastida previously reported<sup>13,42</sup> or ciliate species. The case of Ciliophora stands out since their morphological diversity is well known in at least

one lake<sup>2,42</sup>. Combining metabarcoding and morphology of Ciliophora would allow straightforward understanding of their ecology and behaviour. Cultured references in general offered poor taxonomic quality, as in the case of non-flagellated Archaeplastida (e.g. Chlorococcales). But in some cases OTUs had close counterparts, as in Haptophyta or flagellated Archaeplastida. Molecular approaches can help discern between organisms with homologous morphologies. However the V9 region of the 18 rRNA gene lacks resolution to discern between species<sup>10</sup>. Furthermore, the degree of differentiation of this highly conserved region between organisms can vary between lineages, and its delimitation lies beyond the scope of this article. The fact that the same cultured strain was the CCM to several OTUs (e.g. *Gymnodinium*) or to marine genera, further confirms that the resolution of V9 18S rRNA region is low, and especially critical in the Alveolata group (see more details in Supplementary information). An effort to connect metabarcoding with morphological observations appears as a straightforward and fruitful way to follow. The wide range of monographs on freshwater Archaeplastida or Diatoms<sup>21,43–45</sup>; based on microscopical morphological observations could benefit of single-cell genomics<sup>46</sup> or multi marker gene analysis to help fully disentangle a complex overlooked eukaryotic biosphere and their ecological dynamics. For example, if the correspondence between cyst types and Chrysophyceae barcodes is confirmed, the molecular techniques can provide more comparable ecological studies across sites. The cysts sedimentary record would enlighten a view of the population dynamics over millennia in paleoecological studies, so far mostly addressed with the diatom records.

**Unveiling unexpected heterotrophic diversity in oligotrophic freshwaters.** The variety of heterotrophic flagellates in planktonic ecosystems has been traditionally underestimated<sup>47</sup>, particularly in mountain lakes<sup>48</sup>. Our results unveil an unexpected high heterotrophic diversity in several main groups. Further than the aforementioned *Spumella*-like organisms from Chrysophyceae, less richer phagotrophic flagellated groups, such as Choanomonada, Kathablepharidae or Bicosoecida<sup>49</sup> also presented a higher diversity than expected. Chytridiomycota was the second richest group. The close relationship with cultured counterparts indicated that some of them may be pollen saprophytes. However, a vast richness remains with unknown roles. Chytridiomycota show the highest beta-dispersion, whose patterns per basin were significantly different to the main primary producers (Chrysophyceae and Dinoflagellata). This agrees with specialised saprophytic and parasitic roles<sup>9</sup>. The marked difference between the total species richness (gamma diversity) and the average richness per lake (alpha diversity) indicates a high specificity and punctual incidence in the communities as previously suggested<sup>50</sup>. However, only time series studies in a selected number of lakes can further provide the right answer to this hypothesis. The richness of Chytridiomycota in different aquatic systems is cropping out as more habitats are investigated with molecular techniques<sup>51</sup>. The low average richness per lake compared to the high metacommunity diversity suggests highly specific host interactions. The outstanding heterotrophic diversity is present also in Cercozoa (Rhizaria), whose role has been largely ignored, yet specimens were often observed in microscopic assessments. Parts of this diversity are benthic organisms such as the ameboid forms with organic or siliceous theca. However, flagellated forms (Cercamonadidae and Glissomonadida) spherical ameboid forms with filopodia, such as the Vampyrellidae, which are likely truly planktonic, also contributed with a substantial richness. Overall, and according to very recent findings in aquatic environments (see below), this trait is not exclusive of mountain lakes but probably an unseen key piece gearing planktonic food webs.

**Reconsidering planktonic food webs.** Recent studies have been highlighting the relevance of the heterotrophic component in the marine protists diversity<sup>16</sup> and biological interactions for the community dynamics<sup>52</sup>. The large diversity observed in the present investigation provides the building blocks for reconstructing mountain lake food webs. In these small lakes, the fuelling of the planktonic food web does not depend exclusively on its own primary production but also on organic materials supplied at the interfaces with the atmosphere, the benthic lake habitats, peri-lacustrine environments and terrestrial soils through direct and diffuse runoff<sup>7,43,53</sup>. The external loading of organic material enhances the heterotrophic character of the metacommunity by facilitating growth of saprophytic organisms releasing nutrients into the food webs (e.g. Chytridiomycota growing on pollen grains) and phagotrophic heterotrophs grazing on associated microbial communities. Indeed, the gamma diversity comparison with marine samples, from sites of similar primary productivity remarks the exacerbated saprophytic character of the lake communities. We may further wonder about the role of clearly terrestrial organisms, such as Dikarya or Zygomycota. Their high abundance and richness probably correspond to dispersal stages that constitute a source of food for protists and aquatic invertebrates<sup>54</sup>. In addition, the high richness of truly aquatic fungi or mycoplankton (e.g. Chytridiomycota) indicates a key role of parasitism on the lake plankton dynamics. Detailed studies relating host and Chytridiomycetes infections have focused on some key species (e.g. *Asterionella*<sup>55</sup> or *Daphnia*<sup>56</sup>). Others have studied their implication on the proposed 'mycoloop'<sup>57</sup>, which would sustain zooplankton populations under algal bloom conditions by increasing palatability<sup>58</sup>. How this loop would function under highly oligotrophic circumstances has high research potential. The increasing awareness of the role that Chytridiomycetes play in controlling host populations<sup>59–61</sup> highlights parasitism in lake ecosystem dynamics as a key missing pathway as some time ago occurred with viral infections<sup>62</sup>. The latest molecular approaches provide solid foundations for reviewing the food web paradigm in oligotrophic freshwater lakes<sup>9</sup> where parasitic groups are suspected to have high relevance<sup>63</sup>.

A large variety of parasites from all the main evolutionary branches were also abundant like in Stramenopiles (Labyrinthulomycetes, Peronosporomycetes, Hyphochytriales, Opalinata and *Blastocystis*), Alveolata (Apicomplexa, Perkinsidae and Syndiniales), and Opisthokonta (Blastocladiomycota, Ictyosporae, Myxozoa, Zoopagomycotina). Many of them may be present only as dispersal stages (e.g. zoospores and sporozoites). In any case, the large number of OTUs found demand more consideration of all these organisms in the lake ecosystem. Up to recently, some of the groups were mostly known from marine systems (e.g. Labyrinthulomycetes, Perkinsidae, Syndiniales). Perkinsidae have been found in freshwater only a few years ago<sup>64</sup>, and molecular surveys are increasingly reporting them<sup>65</sup>. These organisms can be parasites of algae, fish, bivalves and amphibians.

Parasitism on benthic invertebrates and macrophytes may have a potential impact on the planktonic community by top-down interactions since littoral can be considered extending to the whole lake in many small mountain lakes. Parasites infecting protist, rotifers and planktonic crustaceans may have a direct influence on the planktonic food web (e.g. Chytridiomycota, Peronosporomycetes, Perkinsidae, Syndiniales, Zoopagomycotina). The stocking of Fish in mountain lakes throughout the world by humans may be highly relevant at this respect<sup>66</sup>. Thus parasites related to them (e.g. Peronosporomycetes, Perkinsidae, Syndiniales, Ichtyosporea, Myxozoa) may merit more insight in future studies, particularly those more virulent. As Chytridiomycota, some groups include a range of life forms; for instance, Blastocladiomycota, are zoospore fungi that some are saprophytic on refractory materials (e.g., pollen, keratin, cellulose, chitin), grow in the submerged parts of marsh plants and may be specific parasites of nematodes, midges and crustaceans. Like LKM11, the ecological role of potentially basal fungal groups remains to be assessed in these oligotrophic lakes. In-depth studies are required but, whatever the specific roles that will come out, the conclusion that saprotrophic and parasitic interactions are central to the dynamics of the lake planktonic communities will probably remain.

**Data accessibility.** Raw files for the genetic datasets are available in the NCBI SRA database under project id PRJNA413654.

## References

- Rodhe, W. Phototrophy and heterotrophy in high mountain lakes. *Verh. Internat. Verein. Limnol.* **16**, 302–313 (1966).
- Felip, M., Camarero, L. & Catalan, J. Temporal changes of microbial assemblages in the ice and snow cover of a high mountain lake. *Limnol. Oceanogr.* **44**, 973–987 (1999).
- Fott, J., Blažo, M., Stuchlik, E. & Štruncový, O. Phytoplankton in three Tatra Mountain lakes of different acidification status. *1999* **58**, 10 (1999).
- Pugnetti, A. & Bettinetti, R. Biomass and species structure of the phytoplankton of an high mountain lake (Lake Paione Superiore, Central Alps, Italy). *1999* **58**, 4 (1999).
- Felip, M. & Catalan, J. The relationship between phytoplankton biovolume and chlorophyll in a deep oligotrophic lake: decoupling in their spatial and temporal maxima. *J. Plankton Res.* **22**, 91–105 (2000).
- Triadó-Margarit, X. & Casamayor, E. O. Genetic diversity of planktonic eukaryotes in high mountain lakes (Central Pyrenees, Spain). *Environ. Microbiol.* **14**, 2445–56 (2012).
- Kammerlander, B. *et al.* High diversity of protistan plankton communities in remote high mountain lakes in the European Alps and the Himalayan mountains. *FEMS Microbiol. Ecol.* **91**, 1–10 (2015).
- Filker, S., Sommaruga, R., Vila, I. & Stoeck, T. Microbial eukaryote plankton communities of high-mountain lakes from three continents exhibit strong biogeographic patterns. *Mol. Ecol.* **25**, 2286–2301 (2016).
- Frenken, T. *et al.* Integrating chytrid fungal parasites into plankton ecology: research gaps and needs. *Environ. Microbiol.* **19**, 3802–3822 (2017).
- Stoeck, T. *et al.* A morphogenetic survey on ciliate plankton from a mountain lake pinpoints the necessity of lineage-specific barcode markers in microbial ecology. *Environ. Microbiol.* **16**, 430–444 (2014).
- Debroas, D. *et al.* Overview of freshwater microbial eukaryotes diversity: a first analysis of publicly available metabarcoding data. *FEMS Microbiol. Ecol.* **93**, 1–14 (2017).
- Triadó-Margarit, X. & Casamayor, E. O. High genetic diversity and novelty in planktonic protists inhabiting inland and coastal high salinity water bodies. *FEMS Microbiol. Ecol.* **85**, 27–36 (2013).
- Triadó-Margarit, X. & Casamayor, E. O. High protists diversity in the plankton of sulfurous lakes and lagoons examined by 18S rRNA gene sequence analyses. *Environ. Microbiol. Rep.* **7**, 908–917 (2015).
- Štráňkrabová, V. *et al.* Investigations on pelagic food webs in mountain lakes - aims and methods. *1999* **58**, 11 (1999).
- Wille, A., Sonntag, B., Sattler, B. & Psenner, R. Abundance, biomass and size structure of the microbial assemblage in the high mountain lake Gossenköllesee (Tyrol, Austria) during the ice-free period. *1999* **58**, 10 (1999).
- de Vargas, C. *et al.* Eukaryotic plankton diversity in the sunlit ocean. *Science (80-)*. **348** (2015).
- Wurzbacher, C. M., Bärlocher, F. & Grossart, H. P. Fungi in lake ecosystems. *Aquat. Microb. Ecol.* **59**, 125–149 (2010).
- Logares, R. *et al.* Infrequent marine-freshwater transitions in the microbial world. *Trends in Microbiology* **17**, 414–422 (2009).
- Auguet, J.-C., Barberán, A. & Casamayor, E. O. Global ecological patterns in uncultured Archaea. *ISME J.* **4**, 182–90 (2010).
- Barberán, A. & Casamayor, E. O. A phylogenetic perspective on species diversity,  $\beta$ -diversity and biogeography for the microbial world. *Mol. Ecol.* **23**, 5868–76 (2014).
- Catalan, J. *et al.* Ecological thresholds in European alpine lakes. *Freshw. Biol.* **54**, 2494–2517 (2009).
- Camarero, L. & Catalan, J. Atmospheric phosphorus deposition may cause lakes to revert from phosphorus limitation back to nitrogen limitation. *Nat. Commun.* **3**, 1118 (2012).
- Ortiz-Alvarez, R. & Casamayor, E. O. High occurrence of Pacearchaeota and Woesearchaeota (Archaea superphylum DPANN) in the surface waters of oligotrophic high-altitude lakes. *Environ. Microbiol. Rep.* **8**, 210–217 (2016).
- Amaral-Zettler, La, McCliment, Ea, Ducklow, H. W. & Huse, S. M. A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA genes. *PLoS One* **4**, e6372 (2009).
- Stoeck, T. *et al.* Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol. Ecol.* **19**(Suppl 1), 21–31 (2010).
- Edgar, R. C. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* **10**, 996–8 (2013).
- Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–6 (2013).
- Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7**, 335–336 (2010).
- Klindworth, A. *et al.* Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* **41**, e1 (2013).
- Margalef, R. Information theory in ecology. *Mem R Acad Cienc Y Artes Barcelona* **32**, 373–449 (1957).
- Berger, W. H. & Parker, F. L. Diversity of Planktonic Foraminifera in Deep-Sea Sediments. *Science (80-)*. **168**, 1345 LP–1347 (1970).
- Massana, R., Pernice, M., Bunge, J. A. & Campo, Jdel Sequence diversity and novelty of natural assemblages of picoeukaryotes from the Indian Ocean. *ISME J.* **5**, 184–195 (2011).
- Capblanq, J. Phytoplankton et productivité primaire de quelques lacs d'altitude dans les Pyrénées. *Annl. Limnol.* **8**, 231–321 (1972).
- Tilzer, M. Dynamik und produktivität von phytoplankton und pelagischen Bakterien in einen Hochgebirgsee (Vorderer Finstertaler See, Österreich). *Arch. Hydrobiol. Suppl.* **42**, 201–273 (1972).
- Maberly, S. C., Ball, L. A., Raven, J. A. & Sültemeyer, D. Inorganic carbon acquisition by chrysophytes. *J. Phycol.* **45**, 1052–1061 (2009).
- Pla, S., Camarero, L. & Catalan, J. Chrysophyte cyst relationships to water chemistry in Pyrenean lakes (NE Spain) and their potential for environmental reconstruction. *J. Paleolimnol.* **30**, 21–34 (2003).

37. Starmach, K. *Chrysophyceae und Haptophyceae*. 1, (VEB Gustav Fisher Verlag, 1985).
38. Grossmann, L., Bock, C., Schweikert, M. & Boenigk, J. Small but Manifold - Hidden Diversity in 'Spumella-like Flagellates'. *J. Eukaryot. Microbiol.* **63**, 419–439 (2016).
39. Catalan, J. Evolution of dissolved and particulate matter during the ice-covered period in a deep, high-mountain lake. *Can. J. Fish. Aquat. Sci.* **49**, 945–955 (1992).
40. Catalan, J. *et al.* Lake Redo ecosystem response to an increasing warming in the Pyrenees during the twentieth century. *J. Paleolimnol.* **28**, 129–145 (2002).
41. Shalchian-Tabrizi, K. *et al.* Diversification of unicellular eukaryotes: cryptomonad colonizations of marine and fresh waters inferred from revised 18S rRNA phylogeny. *Environ. Microbiol.* **10**, 2635–2644 (2008).
42. Felip, M., Bartumeus, F., Halac, S. & Catalan, J. Microbial plankton assemblages, composition and biomass, during two ice-free periods in a deep high mountain lake (Estany Redó, Pyrenees). *J. Limnol.* **58**, 193–202 (1999).
43. Cameron, N. G. *et al.* Surface-sediment and epilithic diatom pH calibration sets for remote European mountain lakes (AL-PE Project) and their comparison with the Surface Waters Acidification Programme (SWAP) calibration set. *J. Paleolimnol.* **22**, 291–317 (1999).
44. Rivera-Rondón, C. A. & Catalan, J. Diatom diversity in the lakes of the Pyrenees: an iconographic reference. *Limnetica* **36**, 127–396 (2017).
45. Wunsam, S., Schmidt, R. & Klee, R. Cyclotella-taxa (Bacillariophyceae) in lakes of the Alpine region and their relationship to environmental variables. *Aquat. Sci.* **57**, 360–386 (1995).
46. Hamilton, P. B., Lefebvre, K. & Bull, R. Single cell PCR amplification of diatoms using fresh and preserved samples. *Front. Microbiol.* **6** (2015).
47. Boenigk, J. & Arndt, H. Bacterivory by heterotrophic flagellates: community structure and feeding strategies. *Antonie Van Leeuwenhoek Int. J. Gen. Mol. Microbiol.* **81**, 465–480 (2002).
48. Straskrabova, V. *et al.* In *Patterns and factors of biota distribution in remote European mountain lakes* (eds Catalan, J., Curtis, C. J. & Kernan, M.) **62**, 19–53 (Schweizerbart science publishers, 2009).
49. del Campo, J. & Massana, R. Emerging diversity within chrysophytes, choanoflagellates and bicosoecids based on molecular surveys. *Protist* **162**, 435–448 (2011).
50. Kagami, M., de Bruin, A., Ibelings, B. W. & Van Donk, E. Parasitic chytrids: their effects on phytoplankton communities and food-web dynamics. *Hydrobiologia* **578**, 113–129 (2007).
51. Comeau, A. M., Vincent, W. F., Bernier, L. & Lovejoy, C. Novel chytrid lineages dominate fungal sequences in diverse marine and freshwater habitats. *Sci. Rep.* **6**, 30120 (2016).
52. Lima-Mendez, G. *et al.* Determinants of community structure in the global plankton interactome. *Science (80-)*. **348** (2015).
53. Ruiz-González, C., Niño-García, J. P. & del Giorgio, P. A. Terrestrial origin of bacterial communities in complex boreal freshwater networks. *Ecol. Lett.* **18**, 1198–1206 (2015).
54. Kagami, M., von Elert, E., Ibelings, B. W., de Bruin, A. & Van Donk, E. The parasitic chytrid, *Zygorhizidium*, facilitates the growth of the cladoceran zooplankton, *Daphnia*, in cultures of the inedible alga, *Asterionella*. *Proc. R. Soc. B-Biological Sci.* **274**, 1561–1566 (2007).
55. Ibelings, B. W. *et al.* Host parasite interactions between freshwater phytoplankton and chytrid fungi (Chytridiomycota). *J. Phycol.* **40**, 437–453 (2004).
56. Johnson, P. T. J., Ives, A. R., Lathrop, R. C. & Carpenter, S. R. Long-term disease dynamics in lakes: causes and consequences of chytrid infections in *Daphnia* populations. *Ecology* **90**, 132–144 (2009).
57. Kagami, M., Miki, T. & Takimoto, G. Mycology: chytrids in aquatic food webs. *Front. Microbiol.* **5**, 166 (2014).
58. Rasconi, S., Grami, B., Niquil, N., Jobard, M. & Sime-Ngando, T. Parasitic chytrids sustain zooplankton growth during inedible algal bloom. *Front. Microbiol.* **5**, 1–19 (2014).
59. Chambouvet, A., Morin, P., Marie, D. & Guillou, L. Control of Toxic Marine Dinoflagellate Blooms by Serial Parasitic Killers. *Science (80-)*. **322**, 1254–1257 (2008).
60. Lefèvre, E., Roussel, B., Amblard, C. & Sime-Ngando, T. The Molecular Diversity of Freshwater Picoeukaryotes Reveals High Occurrence of Putative Parasitoids in the Plankton. *PLoS One* **3**, e2324 (2008).
61. Lepère, C., Domaizon, I. & Debroas, D. Unexpected importance of potential parasites in the composition of the freshwater small-eukaryote community. *Appl. Environ. Microbiol.* **74**, 2940–2949 (2008).
62. Thingstad, T. F. Are viruses important partners in pelagic food webs? *TREE* **8**, 209–213 (1993).
63. Jephcott, T. G. *et al.* Ecological impacts of parasitic chytrids, synidiales and perkinsids on populations of marine photosynthetic dinoflagellates. *Fungal Ecol.* **19**, 47–58 (2016).
64. Lefranc, M., Thenot, A., Lepere, U. & Debroas, D. Genetic diversity of small eukaryotes in lakes differing by their trophic status. *Appl. Environ. Microbiol.* **71**, 5935–5942 (2005).
65. Brate, J. *et al.* Freshwater Perkinsea and marine-freshwater colonizations revealed by pyrosequencing and phylogeny of environmental rDNA. *Isme J.* **4**, 1144–1153 (2010).
66. Ventura, M. *et al.* In *Challenges for high mountain conservation in a changing world* (eds Catalan, J., Ninot, J. M. & Aniz, M. M.) 181–205 (Springer, 2017).

## Acknowledgements

This research was funded by Grants BRIDGES CGL2015–69043-P from the Spanish Office of Science (MINECO) and European funding (ERDF) and DISPERSAL (Ref. 829/2013) from the Spanish National Parks research program (OAPN-MAGRAMA), to EOC, and LACUS CGL2013–45348-P (MINECO-ERDF) to J.C. We thank technicians and students who participated in the field sampling expeditions, and specially M Sala and C Palacin for their excellent performance, and X Torras for ArcGIS support.

## Author Contributions

L.C. and J.C. carried out the field survey and provided the environmental and biological data. R.O.A. and X.T.M. performed the data analyses, with the contributions from J.C. and E.O.C., R.O.A. and J.C. led the writing, with the contributions from all co-authors. E.O.C. lead and supervised the whole study. All authors approved the final version of the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-22835-3>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018





# Chapter 5

Microbial interactions based on co-occurrence and co-exclusion probabilities lie behind the ecology of Pyrenean freshwater lakes



Submitted paper. Embargoed until publication date



# General discussion



## 1. The Pyrenean lake district microbiota

### 1.1. Core and transient organisms in the community ecology of lake surfaces

High mountain lakes are very dynamic ecosystems. Lake water flows getting eventually mixed with different hydrographic compartment such as streams and rivers (Niño-García *et al.*, 2016) or sub-surface reservoirs (Winter, 1999). In this regard, lake surface acts as a mirror of the dispersal mechanisms connecting the catchment (Catalan *et al.*, 2006), and a gateway to further colonize the water column. The high water retention time characteristic of lakes translates into microbial communities capable of forming stable populations with long persistence in the lake. Network analysis of bacterial source-tracking allowed us to visualize for the first time that aquatic-related bacteria in surface Pyrenean lakes waters were strongly gathered into a module of a high number of interactions, confirming the expectation of a core (or autochthonous) group of ecologically relevant aquatic bacteria, which also had the highest abundances (Chapter 5). In addition, the modularity analysis unveiled that some soil / sediment-related bacteria with lower abundances may be transient (or allochthonous), or even dead organisms following catchment processes, whereas other populations of higher abundance may actually have successfully colonized the water column and also have active ecological roles in situ (Lynch and Neufeld, 2015; Ruiz-González *et al.*, 2015). This community fingerprint is certainly an expected outcome, since a high proportion of freshwater surface Bacteria may be transferred from complex soil environments (Crump *et al.*, 2012) , which also harbor specific habitat populations and interactions (Chaffron *et al.*, 2010). We interpret therefore that aquatic core organisms and some seeded from terrestrial environments may establish more permanent populations than stochastically dispersed transient organisms, for which the aquatic environment acts just as a dispersal vector (i.e., soil particles and attached microbiota non viable in the plankton) (Zwart *et al.*, 2003).

The lake surface is also an active part of the planktonic food web. In fact, the continuous supply of organic materials from the interface with the atmosphere, benthos, peri-lacustrine environments and soils (Felip *et al.*, 1999; Ruiz-González *et al.*, 2015). This was also demonstrated through the eukaryotic metacommunity composition, which out of the total richness 11% were terrestrial organisms, such as lichen or vegetal remains. In the other hand, lake and soil bacteria were equal contributors of richness (based on the environmental annotation and database search from chapter 2). This implies that



water flows mobilize more easily terrestrial bacteria, probably due to their smaller sizes, highlighting the role of the lake surface as a dispersal vector for transient bacteria. Alternatively, bacteria may have higher resistance to cells losses. Probably, despite the lower richness, allochthonous eukaryotes may constitute the main food source for the unexpectedly rich Chytridiomycota, that support the heterotrophic character of the surface meta community releasing nutrients into the planktonic food web. These observations highlight that bacterial and eukaryotic components have different roles in the surface meta-community. Therefore, unveiling the hidden diversity and understanding the patterns behind their distributions leads to the understanding of different processes happening at different spatial scales.

### 1.2. Unveiling new diversity of interacting microbiota and their ecological roles

A primary goal in most diversity studies, regardless of further ecological descriptors, is to quantify how much biological diversity is present in a given system. The discovery of new biological diversity is and has been subject to research driven at first by human curiosity, but further explored in detail because of their implications in ecology. Capturing the whole or most of the diversity in the study system provides the foundations to scientifically explore ecosystem-level and evolutionary questions. For instance, lately there is a growing interest to explore diversity radiations within Archaea, that only the latest sequencing techniques have been able to disentangle. Along the Pyrenean survey, 82% of archaeal sequences belonged to a group of unknown function (DHVEG-6), comprised by two clearly differentiated lineages, further coined as Pacearchaeota and Woesearchaeota (Castelle *et al.*, 2015) which occur in 75% of the lakes. Resolving the ecological role of such widely distributed Pacearchaeota and Woesearchaeota groups requires of the evaluation of environmental breadths available from extensive surveys like the one carried out in the Pyrenees. Intermediate amplitude, may indicate that chemical variables are not fundamental. Instead, the abundance relationship with bacterial diversity, linked to genomic traits characteristic of symbiotic organisms, may suggest an unseen to date archaea-bacteria direct biotic interaction (Wrede *et al.*, 2012). A worth to mention case is also present in the heterotrophic eukaryotic component mostly represented by highly diverse aquatic fungi. Again, we detected an overlooked crucial component regarding biotic interactions. Thus, recent studies and reviews are highlighting the role of Chytridiomycota in the 'Mycoloop' as controllers of phytoplankton enhancing zooplankton. Furthermore, there is also a really high diversity and abundance of the LKM11 group (a.k.a. Cryptomycota or

Rozellomycota) with an ecological role still unclear although we observed close links of some Cryptomycota representatives with phytoplankton through co-occurrence networks that deserve further studies. The diversity captured and described by our spatial sampling represents faithfully the regional surface water microbiome of the Pyrenees in summer, although additional insights will require seasonal studies covering heterogeneous aerial inputs (Hervàs, 2009; Peter *et al.*, 2014), and explore variability in the vertical axes sampling deeper lake strata (Auguet *et al.*, 2012; Özgül *et al.*, 2015).

## 2. Environmental filtering in bacterial and eukaryotic communities

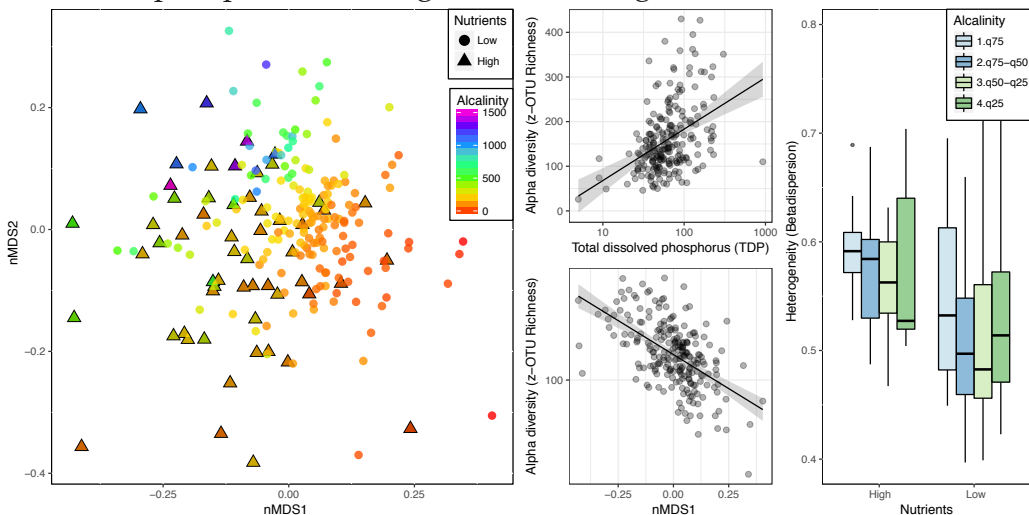
### 2.1 *The process of environmental filtering on environmental preferences*

From the available regional pool of microbes with the potential to reach the lake waters (Chase, 2003), only some of them will successfully establish as core populations locally (Nee *et al.*, 1991; Ruiz-González *et al.*, 2017). Classic niche theories establish that microbes assemble on the basis of their preferences and affinities towards multiple environments of concrete characteristics (Hutchinson, 1961). This process is known as environmental filtering or species sorting (Vellend, 2010). As we have demonstrated empirically (Chapter 1), the deterministic environmental filtering per habitat becomes stronger with time, and more mature communities tend to comprise more similar communities under equivalent environmental filters. In microbes, filters act on the microbial functional characteristics. Indeed, our observations in bacterial assemblages from multiple habitats indicate that pioneer or non-mature communities are comprised by fast growers (cells with a high number of copies of ribosomal RNA) and that are capable of assimilating rapidly inorganic phosphorus from the environment (through the Pst operon system), a metabolic feature that may be specially relevant in oligotrophic systems with P acting as a limiting nutrient. Despite these observations are shared across habitats, there are other traits likely to reflect the specificities of each given system, where the distribution range of a population is related to the genomic size and associated metabolic versatility (Barberán *et al.*, 2014). However the opposite strategy also exists, with different microbes spreading into different niches, so not all of them compete in the same environmental ranges and have specific functional traits. This process can act at the individual level (a single organism with suitable functional traits to live in the given conditions), at the group level (taxa that share a common

functional trait) or at the interaction level (at least two interacting organisms which interaction is only possible under the given conditions in situ). In lakes, aerial colonizers (Peter *et al.*, 2014) and soil colonizers (Niño-García *et al.*, 2017), are environmentally selected by the conditions of each lake to assemble the resulting communities (Niño-García *et al.*, 2016)(Niño-García *et al.*, 2016).

## 2.2. Spatial trends of environmental filtering in eukaryotes

Since some environmental factors have a regional structure, those that have the most significant filtering effect will mimic the resulting community structure along the region. As mentioned, eukaryotes have a lower inter-habitat mobility and cell viability than bacteria. So it is likely that their community patterns may be more related to the specific environmental filtering rather than more stochastic communities driven by random dispersal. The dissimilarity arrangement of communities confirms this hypothesis. Eukaryotes are strongly affected by alkalinity (~pH, with a slightly stronger correlation value) and total phosphorus (although C and N variables co-varied with P in these system, as recurrently observed (Taylor and Townsend, 2010; Camarero and Catalan, 2012). Aiming to expand and lead further studies, community heterogeneity was contextualized under low or high nutrient conditions (phosphate or nitrogen limited) (Figure 1).



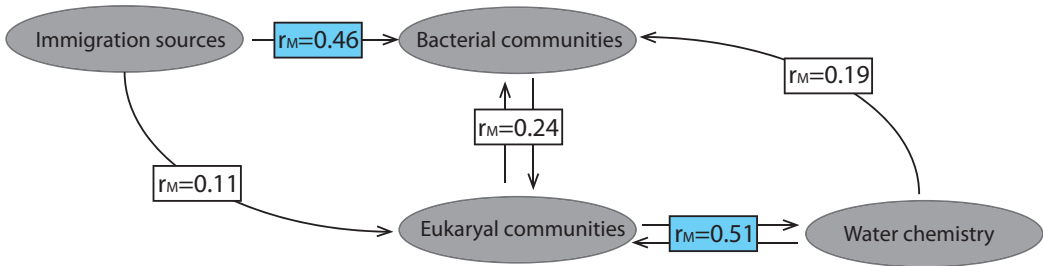
**Figure 1.** Eukaryotic pyrenees metacommunity displayed with the main environmental drivers. Also, relationships with alpha diversity, and beta-diversity based on differential phosphate/nitrogen limitation. Betadispersion is based on diversity categories and in alkalinity (strongest env. correlation) quartiles.

Previous studies in eukaryotic communities with fewer lakes, had already shown that pH (associated to lithology), and phosphorous (trophic status) were the main environmental drivers (Triadó-Margarit and Casamayor, 2012), and key drivers for phytoplankton composition (Buchaca and Catalan, 2008; Grossmann *et al.*, 2016b). The fact that pH (or similar variables) appeared as the most important factor shaping both bacteria and eukarya community composition along the whole regional scale was not surprising as this fact has recurrently been reported in freshwaters (Ren *et al.*, 2015), and agrees with the idea of the Pyrenean lake district being a whole region subject to the same pool of potential colonizers, but separated by the alkalinity substrate gradient. Given the strong relationship observed between alpha diversity of eukaryotes and total phosphorus (Figure 1), we infer that the more phosphorus available, the more diversity is established. But furthermore, local variability under the high P scenario was substantially higher than under low P and available nitrate. It is plausible that the high diversity of primary producers based on Chrysophyceae, Chlorophyta and Cryptophyta may assemble in alternative clusters depending on the trophic situation. And that there are more potential genetic types available at the higher phosphorus situation, and therefore with more possibilities to assemble differently. These alternatives related to a trophic axis of TP have been previously observed in early 2000's, during a spatial study (n=82 lakes), and related it to introduced fish presence (Buchaca and Catalan, 2007). Fish introductions have a strong impact to zooplankton and therefore to the whole trophic chain (Ventura *et al.*, 2017), increasing eutrophia (Buchaca *et al.*, 2011). Because of the fact that both the phosphorus and nitrogen cycles have been affected by global change (Rockström *et al.*, 2009; Steffen *et al.*, 2015), and given the sensitivity of the Pyrenean lakes to varying atmospheric loads of nitrogen and phosphorus (Camarero and Catalan, 2012) we can expect substantial community switches in the near future. Furthermore, the fertilization effects observed in the Pyrenees are likely occurring in other lake districts of alpine biomes. Preserving lakes as close to their oligotrophic natural conditions may help prevent global change effects, and on this matter ongoing projects are restoring lakes to their pristine fishless conditions showing evident changes in the eukaryotic trophic structure along the span of a few years (Ventura *et al.*, 2017).

### 2.3 Environmental filtering and dispersal in bacteria vs. eukaryotes

The aforementioned community patterns in eukaryotes are strongly dri-

ven by deterministic environmental filtering. However, as shown in chapter 2, phosphorus was not a strong factor shaping dissimilarity patterns in bacteria and the predictive strength of pH was limited. The network analysis of bacteria and eukaryotic species interactions showed that network properties differed greatly both separately and in the specific network of bacteria-eukarya interactions. In fact, at least partially, bacterial modules seem to be characterized by intermobility of habitat-specialized populations. Which means, that the strength of stochastic processes and deterministic processes appeared different for bacteria and eukaryotes (Figure 2). Again, this agrees with the previously stated idea of bacterial populations showing higher mobility in the hydrographic network than larger organisms such as protists, algae or fungi. In fact, it doesn't look likely that most freshwater eukaryotes detected can thrive in soils, while we have confirmed that some soil bacteria can seed populations into the water system. Water movements and seeding along the seasonal cycle seem more important for bacteria than the chemical status of lakes, which in turn is crucial determining the eukaryotic communities (Figure 2). Unfortunately the current dataset did not allow to evaluate these phenomena in the archaeal communities, although, because of their sizes it is likely that these are affected by dispersal as bacteria. A major conclusion here is that the regional metacommunity of bacteria is essentially structured by dispersal, while the eukaryotic metacommunity has a stronger effect of environmental filtering.



**Figure 2.** Relationships of distance metrics of communities, immigration sources and water chemistry. Mantel tests with the Spearman method were carried out to assess the strength of links between communities and community descriptors (dispersal and environmental filtering).

### 3. Microbial community assembly and biogeography on the Pyrenees

#### 3.1. Links of environmental and biotic preferences in community assembly

Integrating the concepts of environmental filtering and biotic interactions is a research topic of special interest to enlighten the ecology of the new unveiled diversity, such as novel archaea or aquatic fungal branches. Taking a closer look on diversity of fungal groups, we observed a direct link between P and Richness of Chytridiomycota and Cryptomycota. Given that these groups are heterotrophic and rely on other biotic sources, environment filtering and biotic interactions should be strongly dependent. Another example is the Aenigmarchaeota group that within archaea has the largest habitat breadth, whereas ammonia-oxidizing archaea (AOA) showed the smallest. This fact suggests that, opposite to AOA, the ecological role of Aenigmarchaeota does not rely on specific environmental characteristics, but may be linked to the metabolic activity of other groups such as AOA as previously suggested (Restrepo-Ortiz and Casamayor, 2013). Aenigmarchaeota, and also Pacearchaeota and Woesearchaeota are groups with a reduced genomic repertoire and lack major metabolic pathways. Given the positive relationship between bacterial phylogenetic richness and Pacearchaeota and Woesearchaeota relative abundances, our results support the idea of these groups having a symbiotic or host-dependent lifestyle (Spang *et al.*, 2017). These insights confirm that the evaluation of environmental preferences together with biotic preferences can help enlighten the ecology of groups with unclear lifestyles and trophic roles. As shown by network analysis, species thrive and interact together, and that must modulate growth of others. Especially in the case of facilitator species, interacting microbes may be impossible to grow in axenic cultures without the facilitator. Network analysis suggests the potential groups of microbes that may grow together under *in situ* conditions as co-cultures of microbial consortia (Golyshina *et al.*, 2017). This approach would also be useful to disentangle host-parasite interactions as in the case of Chytridiomycota or cryptomycota (Frenken *et al.*, 2017) and efficiently guide further microbiological studies.

General patterns on the combination of environmental and biotic preferences can give additional clues into the differences between bacteria and eukaryotic behaviour. Co-occurring species pairs are greatly explained by environmental overlap (i.e., equivalent environmental filtering). However, the proportion of eukaryotic pairs explained by the environment was close to 100%. This indicates that the eukaryotic food web is strongly related to

the environmental filtering, modulating it or the other way round. Still, the significant correlation between bacterial and eukaryal communities cannot be solely the result of niche overlap, and as observed in chapter 5 there is a higher percentage of B-E pairs that cannot be explained by the environment, and should be interpreted as commensalism (Christensen *et al.*, 2002) or symbiosis (Lima-Mendez *et al.*, 2015).

### 3.2. *Implication of biotic interactions on community resilience and resistance*

Ecological theory posits that a single stable state or equilibrium is likely to happen in systems with small regional species pools, high rates of connectivity, and low productivity (Chase, 2003; Fukami, 2015), such as primary successional systems. Equilibrium is a concept hardly applicable in lake surfaces, such a dynamic environment with a higher richness than most productive environments (Chase, 2003; Fukami, 2015)(Chase, 2003; Fukami, 2015). However, communities can be more or less stable, and more or less vulnerable to environmental changes or perturbations. In this regard, the strength of the community network of interactions is key on the responsiveness (Lozupone *et al.*, 2012; García-Callejas *et al.*, 2017), since high diversity can sustain the ecosystem functioning under perturbation scenarios (Girvan *et al.*, 2005; Grossmann *et al.*, 2016a). In fact, more mature communities have been traditionally considered to be more stable (Margalef, 1963; Odum, 1969) and predictable than earlier communities (Chapter 1).

Many examples in the different chapters highlight biotic interactions with a potential effect on ecosystem dynamics. For instance, just checking the observed percentage of host-dependent eukaryotic organisms (18%), we can infer that at least these do not thrive alone. The high amount of parasites observed, with hosts such as benthic invertebrates or macrophytes may have a potential impact on the planktonic community by top-down interactions since littoral can be considered extending to the whole lake in many small mountain lakes. Resource sharing, or niche sharing, through apparently negative relationships (that are observed as positive co-occurrences), was in fact a consequence observed in the meta-community. Given the observed patterns in successional trajectories (Chapter 1), more mature communities can have higher diversity of microorganisms, but this diversity likely cooperate with many other microbes and therefore need to share the available resources while limiting their abundances (Chapter 5). This is not a trivial fact, since the carrying capacity of the system (i.e., the total amount of resources that sustain the community) may be a predictor of community vulnerabili-

ty against a perturbation. Interestingly, the network properties observed do not completely resemble those of a network capable of quickly reacting to perturbations (Zhou *et al.*, 2010), since we observed a mix of properties between random networks and small-world networks. What we can say at this point is that topological properties of the bacterial network (slightly higher clustering coefficient and low average path length) resemble those linked to fast reactions to perturbations more than the eukaryotic counterpart. In other words, it seems that once the effect of interactions associated with the environment along with niche overlapping diversity is removed, the eukaryotic pool is more vulnerable than the bacterial pool.

### 3.3 Trends in space of temporal succession

Seeking a temporal succession in a spatial survey requires a keen understanding of the process by which communities change through time, regardless of the timescale. Only understanding this process, spatial patterns can be explained accordingly. For that reason, understanding the overall trend of decreasing beta diversity and increasing alpha diversity reported in chapter 1 is a necessary step before attempting to explain spatial patterns in the Pyrenean lake district. The regional relationship between the geographic gradient and bacterial diversity, from central Pyrenees to both ends (western and eastern Pyrenees), which is also observed when quantifying species interactions, can be interpreted under the concepts associated to successional patterns. The regional coincidence in the increase of alpha diversity and their composition similarity in the Pyrenees can imply that the most specialized and complex communities (eastern and western Pyrenees) are those that have had more time for suitable bacterial accumulation from different sources. Dispersal mechanisms, remoteness, and isolation of high mountain lakes contrast with more hydrographically connected regions such as boreal networks. In fact, in these contrasting systems, highly diverse communities from headwaters and rivers become less diverse towards lakes due to an increased environmental filtering and species sorting (Ruiz-González *et al.*, 2015). However, in these systems, water residence time along the hydrological continuum is a strong predictor of community composition (Niño-García *et al.*, 2016), while in the Pyrenean metacommunity water residence time did not explain the microbial diversity patterns at the regional scale. This difference implies that, despite dispersal has a strong role in the Pyrenees there is not a link between colonization and persistence of taxa related to lake characteristics such as water residence or lake area. When we rule this out, the hypothesis of communities being assembled according to historical



processes becomes stronger. And if we consider the strong seeding of soil-related taxa into the lakes, patterns of lake diversity patterns may be a mirror of historical processes existent in the terrestrial environment. However only future studies could confirm these ideas. Also, the catchment scale was not directly assessed in the case of bacteria, but the eukaryotic populations considered in chapter 4 do show differences regarding catchments. It is possible that seasonal processes acting at the catchment scale may be more linked to hydrology although at the regional scale this potential effect is diluted. Also, the case of eukaryotes is relevant because we did not observe the same diversity pattern that in bacteria (not shown). It is likely that trophic gradients override other sub-regional diversity patterns under the selected tag-sequencing marker V9, which has a low phylogenetic resolution. Equivalent relationships indicating historical processes have been observed in some crustacean populations, but only towards eastern Pyrenees using specific markers (Ventura et al., 2014), that are similar to our findings in the bacterial communities. It is also likely that functional traits related to the lake trophic status in eukaryotic orders or genera are masking biogeographic patterns at the species level. Probably the V4 loop in the 16S rRNA gene has enough information to unveil this pattern in bacteria, whereas the V9 loop in the 18S rRNA gene lacks enough taxonomic resolution. Marker combinations of V9 with more variable regions such as ITS (Internal transcribed spacer), could enhance population studies with a higher resolution and more detailed phylogeographic approaches for specific groups along the spatial area.

Furthermore, if we interpret that most positive interactions are mutualistic interactions, we can link the diversity pattern to community stability (Thebault and Fontaine, 2010; Lever et al., 2014; Rohr et al., 2014). So, if bacterial communities in the two ends of the Pyrenees are more stable than those of the Central Pyrenees, it may be worth to conjecture that colonization by soil taxa may differ if community stability varies, and so the environmental filtering of immigrants in the lakes may be weaker or stronger. Potentially, our results do indicate that alpha and beta diversity indexes contextualized within longitude and altitude on a regional scale are proxies of long-term ecological succession since the de-glaciation of the Pyrenees. In other words, spatial patterns along more than 500 km could be linked to temporal processes that shaped more than 300 microbial communities. This opens a question for future studies that may answer whether the patterns observed are explained solely by recent seasonality or if diversity patterns are long-term proxies associated with the geological scale.

# General conclusions



- By combining studies that have examined primary succession patterns across a wide variety of habitats, we reasonably identified consistent and predictable trends in community composition, diversity, and functional attributes across successional gradients. These trends are in agreement with current concepts about how plant communities shift during primary succession. A deterministic trend towards community assembly is observed, enhancing slow growers (lower rRNA operon copy number), which may show adaptations for oligotrophic conditions and resource sharing (Pst operon) (**Chapter 1**).
- Processes driving community assembly in high mountain lakes are directly related to the dispersal and environmental filtering of concrete environmental sources such as epilimnion or sediments. We found trends towards specific epilimnetic bacteria and a significant impact of Sediment, River and Wetland populations on local community structures. We observed that pH filtered the epilimnion and sediment sources differently (**Chapter 2**).
- Alpha and beta diversity of bacterial communities indicate that the spatial structure in the Pyrenean lacustrine district has trends of a temporal gradient along the mountain range geography, as a result of a complex interaction of environmental source dispersal and environmental filtering. This hints to lake community maturity as a compendium of habitat sources being the aquatic taxa more deterministically selected rather than the more stochastic populations from soil, biofilms or sediments (**Chapter 2**).
- Within the DPANN archaeal group, Woesearchaeota, Pacearchaeota, Diapherotrites, Aenigmarchaeota, and Micrarchaeota are important groups in the summer Pyrenean lacustrine district but require of additional genetic and microbiological studies to improve the current knowledge on their ecological roles (**Chapter 3**).
- Pacearchaeota and Woesearchaeota were highly occurring taxa, and showed intermediate environmental amplitude, suggesting that chemical variables are not fundamental drivers. Instead, the abundance relationship with bacterial diversity, linked to genomic traits characteristic of symbiotic organisms, may suggest an archaea-bacteria direct biotic interaction (**Chapter 3**).

- Unexpectedly, the high eukaryotic diversity in Pyrenean lacustrine district equally split among autotrophic, heterotrophic and parasitic components. The large phytoplanktonic diversity comprised mostly Chrysophyceae, Dinoflagellata and Chlorophytes (**Chapter 4**).
- Planktonic food web on the basis of the high diversity of Chytriomycota, Cercozoa and other heterotrophs needs to be reconsidered. The external loading of terrestrial organic material enhances the heterotrophic character of communities by facilitating growth of saprophytic organisms releasing nutrients into the food webs. Also, the increasing awareness of the role that Chytridiomycetes play in controlling host populations highlights parasitism in lake ecosystem dynamics as a key missing pathway (**Chapter 4**).
- We inferred species interactions through removing environmental niche effects or geographic dispersal limitation as drivers of co-occurrences, with particular network topologies. According to that, responsiveness of communities towards environmental changes would be different for eukaryotic than for bacterial communities (**Chapter 5**).
- We further observed a role of environmental sources regarding freshwater bacterial interactivity. Supply of sediment and soil bacteria may have a transient role or be strong community competitors and facilitators, which we were only capable of identify through separating the analysis of positive and negative interactions. In this regard, competition seems linked to a sharing strategy, through limiting maximum abundances in the community (**Chapter 5**).
- If competition through resource sharing limits maximum species abundances (**Chapter 5**), which is likely to happen more in more diverse and complex late successional communities (**Chapter 1**), it is possible to infer that organisms capable of strong competition and sharing of resources need to function under oligotrophic conditions. Microbes may do so through decreasing their rRNA copy number and their ability to mobilize inorganic phosphorus (**Chapter 1**).

# Bibliography

from the general sections



- Alfreider A, Pernthaler J, Amann R, Sattler B, Glockner FO, Wille a, *et al.* (1996). Community analysis of the bacterial assemblages in the winter cover and pelagic layers of a high mountain lake by in situ hybridization. *Appl Environ Microbiol* **62**: 2138–2144.
- Allen TR, Walsh SJ. (1996). Spatial and compositional pattern of alpine treeline, Glacier National Park, Montana. *Photogramm Eng Remote Sensing* **62**: 1261–1268.
- Auguet J-C, Casamayor EO. (2013). Partitioning of Thaumarchaeota populations along environmental gradients in high mountain lakes. *FEMS Microbiol Ecol* **84**: 154–164.
- Auguet J-C, Triadó-Margarit X, Nomokonova N, Camarero L, Casamayor EO. (2012). Vertical segregation and phylogenetic characterization of ammonia-oxidizing Archaea in a deep oligotrophic lake. *ISME J* **6**: 1786–1797.
- Barberán A, Casamayor EO. (2014). A phylogenetic perspective on species diversity,  $\beta$ -diversity and biogeography for the microbial world. *Mol Ecol* **23**: 5868–76.
- Barberán A, Casamayor EO, Fierer N. (2014a). The microbial contribution to macroecology. *Front Microbiol* **5**: 1–8.
- Barberán A, Ramirez KS, Leff JW, Bradford MA, Wall DH, Fierer N. (2014b). Why are some microbes more ubiquitous than others? Predicting the habitat breadth of soil bacteria. *Ecol Lett* **17**: 794–802.
- Bartrons M, Catalan J, Casamayor EO. (2012). High Bacterial Diversity in Epilithic Biofilms of Oligotrophic Mountain Lakes. *Microb Ecol* **64**: 860–869.
- Blois JL, Zarnetske PL, Fitzpatrick MC, Finnegan S. (2013). Climate change and the past, present, and future of biotic interactions. *Science* **341**: 499–504.
- Buchaca T, Catalan J. (2007). Factors influencing the variability of pigments in the surface sediments of mountain lakes. *Freshw Biol* **52**: 1365–1379.
- Buchaca T, Catalan J. (2008). On the contribution of phytoplankton and benthic biofilms to the sediment record of marker pigments in high mountain lakes. *J Paleolimnol* **40**: 369–383.
- Buchaca T, Skov T, Amsinck SL, Gonçalves V, Azevedo JMN, Andersen TJ, *et al.* (2011). Rapid Ecological Shift Following Piscivorous Fish Introduction to Increasingly Eutrophic and Warmer Lake Furnas (Azores Archipelago, Portugal): A Paleoecological Approach. *Ecosystems* **14**: 458–477.
- Caliz J, Casamayor EO. (2014). Environmental controls and composition of anoxygenic photoheterotrophs in ultraoligotrophic high-altitude lakes (Central Pyrenees). *Environ Microbiol Rep* **6**: 145–151.
- Camarero L, Catalan J. (2012). Atmospheric phosphorus deposition may cause lakes to revert from phosphorus limitation back to nitrogen limitation. *Nat Commun* **3**: 1118.
- Camarero L, Rogora M, Mosello R, Anderson NJ, Barbieri A, Botev I, *et al.* (2009). Regionalisation of chemical variability in European mountain lakes. *Freshw Biol* **54**: 2452–2469.
- Cardinale BJ, Duffy JE, Gonzalez A, Hooper DU, Perrings C, Venail P, *et al.* (2012). Biodiversity loss and its impact on humanity. *Nature* **486**: 59–67.
- Caro-Quintero A, Deng J, Auchtung J, Brettar I, Höfle MG, Klappenbach J, *et al.*



- (2011). Unprecedented levels of horizontal gene transfer among spatially co-occurring *Shewanella* bacteria from the Baltic Sea. *ISME J* **5**: 131–40.
- Castelle CJ, Wrighton KC, Thomas BC, Hug L a, Brown CT, Wilkins MJ, *et al.* (2015). Genomic expansion of domain archaea highlights roles for organisms from new phyla in anaerobic carbon cycling. *Curr Biol* **25**: 690–701.
- del Castillo Jurado M. (2003). Morfometría de lagos. Una aplicación a los lagos del Pirineo. Universitat de Barcelona: Barcelona.
- Catalan J, Barbieri MG, Bartumeus F, Bitusk P, Botev I, Brancelj A, *et al.* (2009). Ecological thresholds in European alpine lakes. *Freshw Biol* **54**: 2494–2517.
- Catalan J, Camarero L, Felip M, Pla S, Ventura M, Buchaca T, *et al.* (2006). High mountain lakes: extreme habitats and witnesses of environmental changes. *Limnetica* **64**: 123–145.
- Catalan J, Pla S, Rieradevall M, Felip M, Ventura M, Buchaca T, *et al.* (2002). Lake Redo ecosystem response to an increasing warming in the Pyrenees during the twentieth century. *J Paleolimnol* **28**: 129–145.
- Chaffron S, Rehrauer H, Pernthaler J, Von Mering C. (2010). A global network of coexisting microbes from environmental and whole-genome sequence data. *Genome Res* **20**: 947–959.
- Chase JM. (2003). Community assembly: When should history matter? *Oecologia* **136**: 489–498.
- Chesson PL, Warner RR. (1981). Environmental Variability Promotes Coexistence in Lottery Competitive Systems. *Am Nat* **117**: 923–943.
- Christensen BB, Haagenen JAJ, Heydorn A, Molin S. (2002). Metabolic commensalism and competition in a two-species microbial consortium. *Appl Environ Microbiol* **68**: 2495–2502.
- Cicczazzo S, Esposito A, Borruso L, Brusetti L. (2016). Microbial communities and primary succession in high altitude mountain environments. *Ann Microbiol* **66**: 43–60.
- Crump BC, Amaral-Zettler LA, Kling GW. (2012). Microbial diversity in arctic freshwaters is structured by inoculation of microbes from soils. *ISME J* **6**: 1629–1639.
- Delmas M. (2015). The last maximum ice extent and subsequent deglaciation of the Pyrenees: an overview of recent research. *Cuad Investig Geográfica* **41**: 359–287.
- Delmas M, Gunnell Y, Braucher R, Calvet M, Bourlès D. (2008). Exposure age chronology of the last glaciation in the eastern Pyrenees. *Quat Res* **69**: 231–241.
- Downing JA, Prairie YT, Cole JJ, Duarte CM, Tranvik LJ, Striegl RG, *et al.* (2006). Abundance and Size Distribution of Lakes, Ponds and Impoundments. *Ref Modul Earth Syst Environ Sci* **51**: 2388–2397.
- Faust K, Raes J. (2012). Microbial interactions: from networks to models. *Nat Rev Microbiol* **10**: 538–550.
- Felip M, Bartumeus F, Halac S, Catalan J. (1999a). Microbial plankton assemblages, composition and biomass, during two ice-free periods in a deep high mountain lake (Estany Redó, Pyrenees). *J Limnol* **58**: 193–202.
- Felip M, Camarero L, Catalan J. (1999b). Temporal changes of microbial assemblages in the ice and snow cover of a high mountain lake. *Limnol Oceanogr* **44**:

- 973–987.
- Frenken T, Alacid E, Berger SA, Bourne EC, Gerphagnon M, Grossart H-P, *et al.* (2017). Integrating chytrid fungal parasites into plankton ecology: research gaps and needs. *Environ Microbiol* **19**: 3802–3822.
- Fukami T. (2015). Historical contingency in community assembly : integrating niches, species pools, and priority effects. *Annu Rev Ecol Evol Syst* **46**: 1–23.
- García-Callejas D, Molowny-Horas R, Araújo MB. (2017). Multiple interactions networks: towards more realistic descriptions of the web of life. *Oikos* **127**: 5–22.
- Gilbert B, Levine JM. (2017). Ecological drift and the distribution of species diversity. *Proc R Soc B Biol Sci* **284**: 20170507.
- Girvan MS, Campbell CD, Killham K, Prosser JI, Glover LA. (2005). Bacterial diversity promotes community stability and functional resilience after perturbation. *7*: 301–313.
- Golyshina O V., Toshchakov S V., Makarova KS, Gavrillov SN, Korzhenkov AA, La Cono V, *et al.* (2017). 'ARMAN' archaea depend on association with euryarchaeal host in culture and in situ. *Nat Commun* **8**: 1–11.
- Grossmann L, Beisser D, Bock C, Chatzinotas A, Jensen M, Preisfeld A, *et al.* (2016a). Trade-off between taxon diversity and functional diversity in European lake ecosystems. *Mol Ecol* **25**: 5876–5888.
- Grossmann L, Jensen M, Pandey RV, Jost S, Bass D, Psenner R, *et al.* (2016b). Molecular investigation of protistan diversity along an elevation transect of alpine lakes. *Aquat Microb Ecol* **78**: 25–37.
- Van der Gucht K, Cottenie K, Muylaert K, Vloemans N, Cousin S, Declerck S, *et al.* (2007). The power of species sorting: Local factors drive bacterial community composition over a wide range of spatial scales. *Proc Natl Acad Sci* **104**: 20404–20409.
- Hanski I, Gilpin ME. (1991). *Metapopulation Dynamics: Empirical and Theoretical Investigations*.
- Hervàs A. (2009). Bacterial dispersal on airborne saharan dust particles: survival and colonization in alpine lakes of the limnological observatory of the Pyrenees. Universitat de Barcelona: Blanes.
- Hervas A, Casamayor EO. (2009). High similarity between bacterioneuston and airborne bacterial community compositions in a high mountain lake area. *FEMS Microbiol Ecol* **67**: 219–228.
- Hutchinson GE. (1961). The paradox of the plankton. *Am Nat* **95**: 137–145.
- Lauga B, Malaval S, Largier G, Regnault-Rogert C. (2009). Two lineages of *Trifolium alpinum* (Fabaceae) in the Pyrenees: Evidence from random amplified polymorphic DNA (RAPD) markers. *Acta Bot Gall* **156**: 317–330.
- Leavitt PR, Vinebrooke RD, Donald DB, Smol JP, Schindler DW. (1997). Past ultraviolet radiation environments in lakes derived from fossil pigments. *Nature* **388**: 457–459.
- Leibold MA, Holyoak M, Mouquet N, Amarasekare P, Chase JM, Hoopes MF, *et al.* (2004). The metacommunity concept: A framework for multi-scale community ecology. *Ecol Lett* **7**: 601–613.

- Lever JJ, van Nes EH, Scheffer M, Bascompte J. (2014). The sudden collapse of pollinator communities. *Ecol Lett* **17**: 350–359.
- Li H, Zeng J, Ren L, Wang J, Xing P, Wu QL. (2017). Contrasting patterns of diversity of abundant and rare bacterioplankton in freshwater lakes along an elevation gradient. *Limnol Oceanogr*. e-pub ahead of print, doi: 10.1002/lno.10518.
- Lima-Mendez G, Faust K, Henry N, Decelle J, Colin S, Carcillo F, *et al.* (2015). Determinants of community structure in the global plankton interactome. *Science* (80- ) **348**. e-pub ahead of print, doi: 10.1126/science.1262073.
- Llorens–Marès T, Liu Z, Allen LZ, Rusch DB, Craig MT, Dupont CL, *et al.* (2017). Speciation and ecological success in dimly lit waters: horizontal gene transfer in a green sulfur bacteria bloom unveiled by metagenomic assembly. *ISME J* **11**: 201–211.
- Lotter AF. (1989). Subfossil and modern diatom plankton and the paleolimnology of Rotsee (Switzerland) since 1850. *Aquat Sci* **51**: 338–350.
- Lotter AF, Bigler C. (2000). Do diatoms in the Swiss Alps reflect the length of ice-cover? *Aquat Sci* **62**: 125.
- Lozupone C a, Stombaugh JI, Gordon JI, Jansson JK, Knight R. (2012). Diversity, stability and resilience of the human gut microbiota. *Nature* **489**: 220–30.
- Lynch MDJ, Neufeld JD. (2015). Ecology and exploration of the rare biosphere. *Nat Rev Microbiol* **13**: 217–229.
- Malanson GP, Butler DR, Fagre DB, Walsh SJ, Tomback DF, Daniels LD, *et al.* (2007). Alpine Treeline of Western North America: Linking Organism-To-Landscape Dynamics. *Phys Geogr* **28**: 378–396.
- Manzano S, Carrión JS, López-Merino L, González-Sampériz P, Munuera M, Fernández S, *et al.* (2017). Mountain strongholds for woody angiosperms during the Late Pleistocene in SE Iberia. *Catena* **149**: 701–712.
- Marchetto A, Rogora M, Boggero A, Musazzi S, Lami A, Lotter AF, *et al.* (2009). Response of Alpine lakes to major environmental gradients, as detected through planktonic, benthic and sedimentary assemblages. *Adv Limnol* **62**: 419–440.
- Margalef R. (1963). On Certain Unifying Principles in Ecology. *Am Nat* **97**: 357.
- Margalef R. (1968). Perspectives in Ecological Theory. University of Chicago Press: Chicago, Illinois, USA.
- McPeck MA, Gomulkiewicz R. (2005). Assembling and depleting species richness in metacommunities: insights from ecology, population genetics and macroevolution. In: Holyoak M, M.A. L, Holt RD (eds). *Metacommunities: spatial dynamics and ecological communities*. University of Chicago Press: Chicago, Illinois, USA, pp 355–373.
- Müller B, Lotter AF, Sturm M, Ammann A. (1998). Influence of catchment quality and altitude on the water and sediment composition of 68 small lakes in Central Europe. *Aquat Sci* **60**: 316–337.
- Nee S, Gregory RD, May RM. (1991). Core and Satellite Species: Theory and Artefacts. *Oikos* **62**: 83–87.
- Niño-García JP, Ruiz-González C, del Giorgio PA. (2017). Exploring the ecological coherence between the spatial and temporal patterns of bacterioplankton in bo-

- real lakes. *Front Microbiol* **8**: 1–12.
- Niño-García JP, Ruiz-González C, del Giorgio PA. (2016). Interactions between hydrology and water chemistry shape bacterioplankton biogeography across boreal freshwater networks. *ISME J* 1–12.
- Odum EP. (1969). The strategy of ecosystem development. *Science* **164**: 262–270.
- Ortiz-Álvarez R, De Los Ríos A, Fernández-Mendoza F, Torralba-Burrial A, Pérez-Ortega S. (2015). Ecological specialization of two photobiont-specific maritime cyanolichen species of the genus *Lichina*. *PLoS One* **10**. e-pub ahead of print, doi: 10.1371/journal.pone.0132718.
- Özgül İ, Llíros M, Crowe SA, García-armisen T, Morana C, Darchambeau F, *et al.* (2015). Vertical Distribution of Functional Potential and Active Microbial Communities in Meromictic Lake Kivu. e-pub ahead of print, doi: 10.1007/s00248-015-0612-9.
- Pace NR, Stahl DA, Lane DJ, Olsen GJ. (1986). The Analysis of Natural Microbial Populations by Ribosomal RNA Sequences. In: *Advances in Microbial Ecology*. pp 1–55.
- Palmer MW. (1994). Variation in species richness: Towards a unification of hypotheses. *Folia Geobot Phytotaxon* **29**: 511–530.
- Pernthaler J, Glöckner F. (1998). Seasonal community and population dynamics of pelagic bacteria and archaea in a high mountain lake. *Appl Environ Microbiol* **64**: 4299–4306.
- Pernthaler J, Glöckner F-O, Unterholzner S, Alfreider A, Psenner R, Amann R. (1998). Seasonal community and population dynamics of pelagic bacteria and Archaea in a high mountain lake. *Appl Environ Microbiol* **64**: 4299–4306.
- Pernthaler J, Šimek K, Sattler B, Schwarzenbacher A, Bobkova J, Psenner R. (1996). Short-term changes of protozoan control on autotrophic picoplankton in an oligo-mesotrophic lake. *J Plankton Res* **18**: 443–462.
- Peter H, Hörtnagl P, Reche I, Sommaruga R. (2014). Bacterial diversity and composition during rain events with and without Saharan dust influence reaching a high mountain lake in the Alps. *Environ Microbiol Rep* **6**: 618–624.
- Peter H, Sommaruga R. (2016). Shifts in diversity and function of lake bacterial communities upon glacier retreat. *ISME J* **10**: 1–10.
- Pla S, Camarero L, Catalan J. (2003). Chrysophyte cyst relationships to water chemistry in Pyrenean lakes (NE Spain) and their potential for environmental reconstruction. *J Paleolimnol* **30**: 21–34.
- R Core Team. (2017). R: A Language and Environment for Statistical Computing. R Found Stat Comput Vienna, Austria **0**: {ISBN} 3-900051-07-0.
- Ren L, Jeppesen E, He D, Wang J, Liboriussen L, Xing P, *et al.* (2015). pH Influences the Importance of Niche-Related and Neutral Processes in Lacustrine Bacterioplankton Assembly. *Appl Environ Microbiol* **81**: 3104–14.
- Restrepo-Ortiz CX, Casamayor EO. (2013). Environmental distribution of two widespread uncultured freshwater Euryarchaeota clades unveiled by specific primers and quantitative PCR. *Environ Microbiol Rep* **5**: 861–7.
- Rockström J, Steffen W, Noone K, Persson Å, Chapin FS, Lambin E, *et al.* (2009).

- Planetary boundaries: Exploring the safe operating space for humanity. *Ecol Soc* **14**. e-pub ahead of print, doi: 10.1038/461472a.
- Rodriguez-Brito B, Li L, Wegley L, Furlan M, Angly F, Breitbart M, *et al.* (2010). Viral and microbial community dynamics in four aquatic environments. *ISME J* **4**: 739–751.
- Rohr RP, Saavedra S, Bascompte J. (2014). On the structural stability of mutualistic systems. *Science (80- )* **345**: 1253497–1253497.
- Ruiz-González C, Niño-García JP, del Giorgio PA. (2015). Terrestrial origin of bacterial communities in complex boreal freshwater networks. *Ecol Lett* **18**: 1198–1206.
- Ruiz-González C, Niño-García JP, Kembel SW, del Giorgio PA. (2017). Identifying the core seed bank of a complex boreal bacterial metacommunity. *ISME J* **11**: 2012–2021.
- Schmitt T. (2017). Molecular Biogeography of the High Mountain Systems of Europe: An Overview. In: Catalan J (ed). *High Mountain Conservation in a Changing World*. Advances in Global Change Research, pp 63–74.
- Shurin JB, Allen EG. (2001). Effects of competition, predation, and dispersal on species richness at local and regional scales. *Am Nat* **158**: 624–637.
- Simek K, Bobkova J, Macek M, Nedoma J, Psenner R. (1995). Ciliate grazing on picoplankton in a eutrophic reservoir during the summer phytoplankton maximum: A study at the species and community level. *Limnol Oceanogr* **40**: 1077–1090.
- Šimek K, Hartman P, Nedoma J, Pernthaler J, Springmann D, Vrba J, *et al.* (1997). Community structure, picoplankton grazing and Zooplankton control of heterotrophic nanoflagellates in a eutrophic reservoir during the summer phytoplankton maximum. *Aquat Microb Ecol* **12**: 49–63.
- Sommaruga R, Krossbacher M, Salvenmoser W, Catalan J, Psenner R. (1995). Presence of large virus-like particles in a eutrophic reservoir. *Aquat Microb Ecol* **9**: 305–308.
- Sommaruga R, Oberleiter A, Psenner R. (1996). Effect of UV radiation on the bacterivory of a heterotrophic nanoflagellate. *Appl Environ Microbiol* **62**: 4395–4400.
- Spang A, Caceres EF, Ettema TJG. (2017). Genomic exploration of the diversity, ecology, and evolution of the archaeal domain of life. *Science (80- )* **357**. e-pub ahead of print, doi: 10.1126/science.aaf3883.
- Spang A, Saw JH, Jørgensen SL, Zaremba-Niedzwiedzka K, Martijn J, Lind AE, *et al.* (2015). Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature* **521**: 173–179.
- Steffen W, Richardson K, Rockström J, Cornell SE, Fetzer I, Bennett EM, *et al.* (2015). Sustainability. Planetary boundaries: guiding human development on a changing planet. *Science* **347**: 1259855.
- Taylor PG, Townsend AR. (2010). Stoichiometric control of organic carbon-nitrate relationships from soils to the sea. *Nature* **464**: 1178–1181.
- Thebault E, Fontaine C. (2010). Stability of Ecological Communities and the Architecture of Mutualistic and Trophic Networks. *Science (80- )* **329**: 853–856.
- Tolotti M, Thies H, Nickus U, Psenner R. (2012). Temperature modulated effects

- of nutrients on phytoplankton changes in a mountain lake. *Hydrobiologia* **698**: 61–75.
- Triadó-Margarit X, Casamayor EO. (2012). Genetic diversity of planktonic eukaryotes in high mountain lakes (Central Pyrenees, Spain). *Environ Microbiol* **14**: 2445–56.
- Triadó-Margarit X, Casamayor EO. (2013). High genetic diversity and novelty in planktonic protists inhabiting inland and coastal high salinity water bodies. *FEMS Microbiol Ecol* **85**: 27–36.
- Triadó-Margarit X, Casamayor EO. (2015). High protists diversity in the plankton of sulfurous lakes and lagoons examined by 18s rRNA gene sequence analyses. *Environ Microbiol Rep* **7**: 908–917.
- Vellend M. (2010). Conceptual synthesis in community ecology. *Q Rev Biol* **85**: 183–206.
- Ventura M, Petrusek A, Miró A, Hamrová E, Buñay D, De Meester L, *et al.* (2014). Local and regional founder effects in lake zooplankton persist after thousands of years despite high dispersal potential. *Mol Ecol* **23**: 1014–1027.
- Ventura M, Tiberti R, Buchaca T, Buñay D, Sabás I, Miró A. (2017). Why should we preserve fishless high-mountain lakes? In: Catalan J, Ninot JM, Aniz MM (eds). *Challenges for high mountain conservation in a changing world*. Springer, pp 181–205.
- Walker LR, del Moral R. (2003). *Primary Succession and Ecosystem Rehabilitation*. Cambridge University Press.
- Walker M, Johnsen S, Rasmussen SO, Popp T, Steffensen J-P, Gibbard P, *et al.* (2009). Formal definition and dating of the GSSP (Global Stratotype Section and Point) for the base of the Holocene using the Greenland NGRIP ice core, and selected auxiliary records. *J Quat Sci* **24**: 3–17.
- Wang J, Meier S, Sojininen J, Casamayor EO, Pan F, Tang X, *et al.* (2016). Regional and global elevational patterns of microbial species richness and evenness. *Ecography* **40**: 393–402.
- Wille A, Sonntag B, Sattler B, Psenner R. (1999). Abundance, biomass and size structure of the microbial assemblage in the high mountain lake Gossenköllesee (Tyrol, Austria) during the ice-free period. *J limnol* **58**: 117–126.
- Wilson DS. (1992). Complex interactions in metacommunities, with implications for biodiversity and higher levels of selection. *Ecology* **73**: 1984–2000.
- Winter TC. (1999). Relation of streams, lakes, and wetlands to groundwater flow systems. *Hydrogeol J* **7**: 28–45.
- Wrede C, Dreier A, Kokoschka S, Hoppert M. (2012). Archaea in symbioses. *Archaea* **2012**: 596846.
- Zaremba-Niedzwiedzka K, Caceres EF, Saw JH, Bäckström Di, Juzokaite L, Vancaester E, *et al.* (2017). Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature* **541**: 353–358.
- Zhou J, Deng Y, Luo F, He Z, Tu Q, Zhi X. (2010). Functional molecular ecological networks. *MBio* **1**: e00169-10.
- Zwart G, Van Hannen EJ, Kamst-van Agterveld MP, Van der Gucht K, Lindström

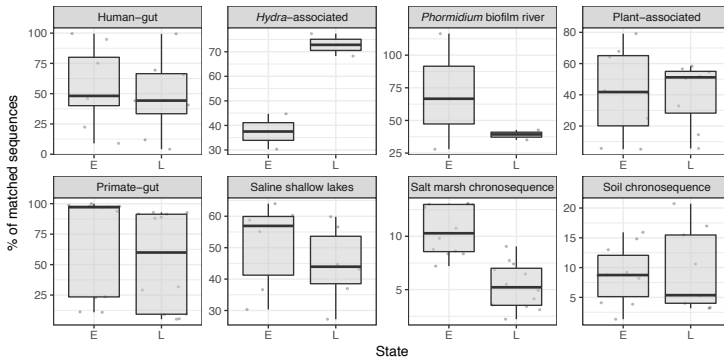
ES, Van Wichelen J, *et al.* (2003). Rapid Screening for Freshwater Bacterial Groups by Using Reverse Line Blot Hybridization? *Appl Environ Microbiol* **69**: 5875–5883.

ANNEX:  
Supplementary information

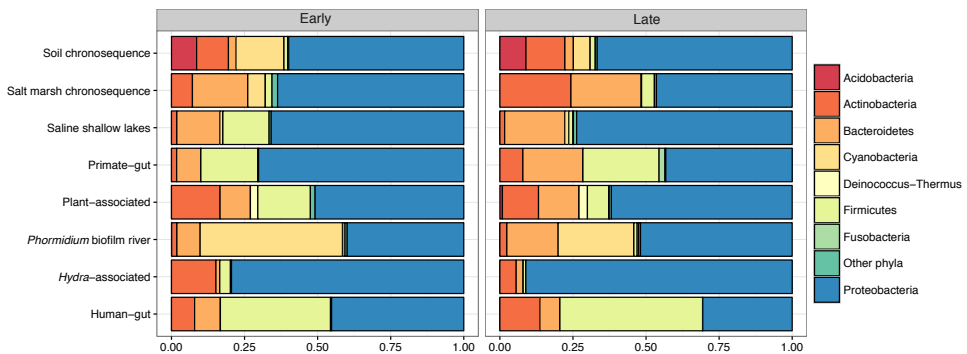




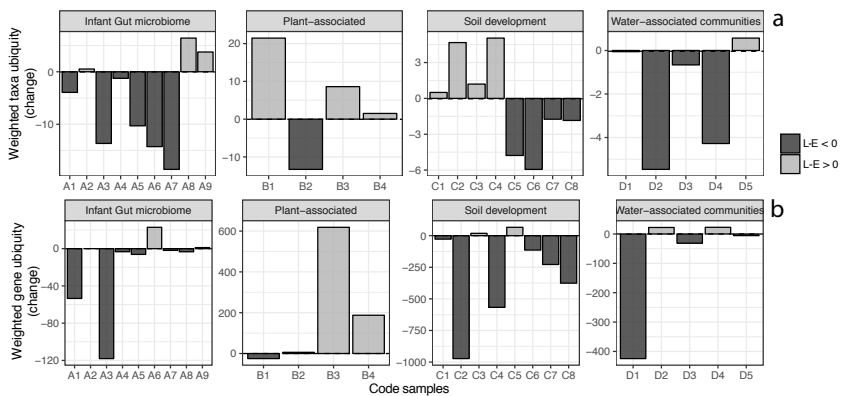
## SUPPLEMENTARY INFORMATION CHAPTER 1



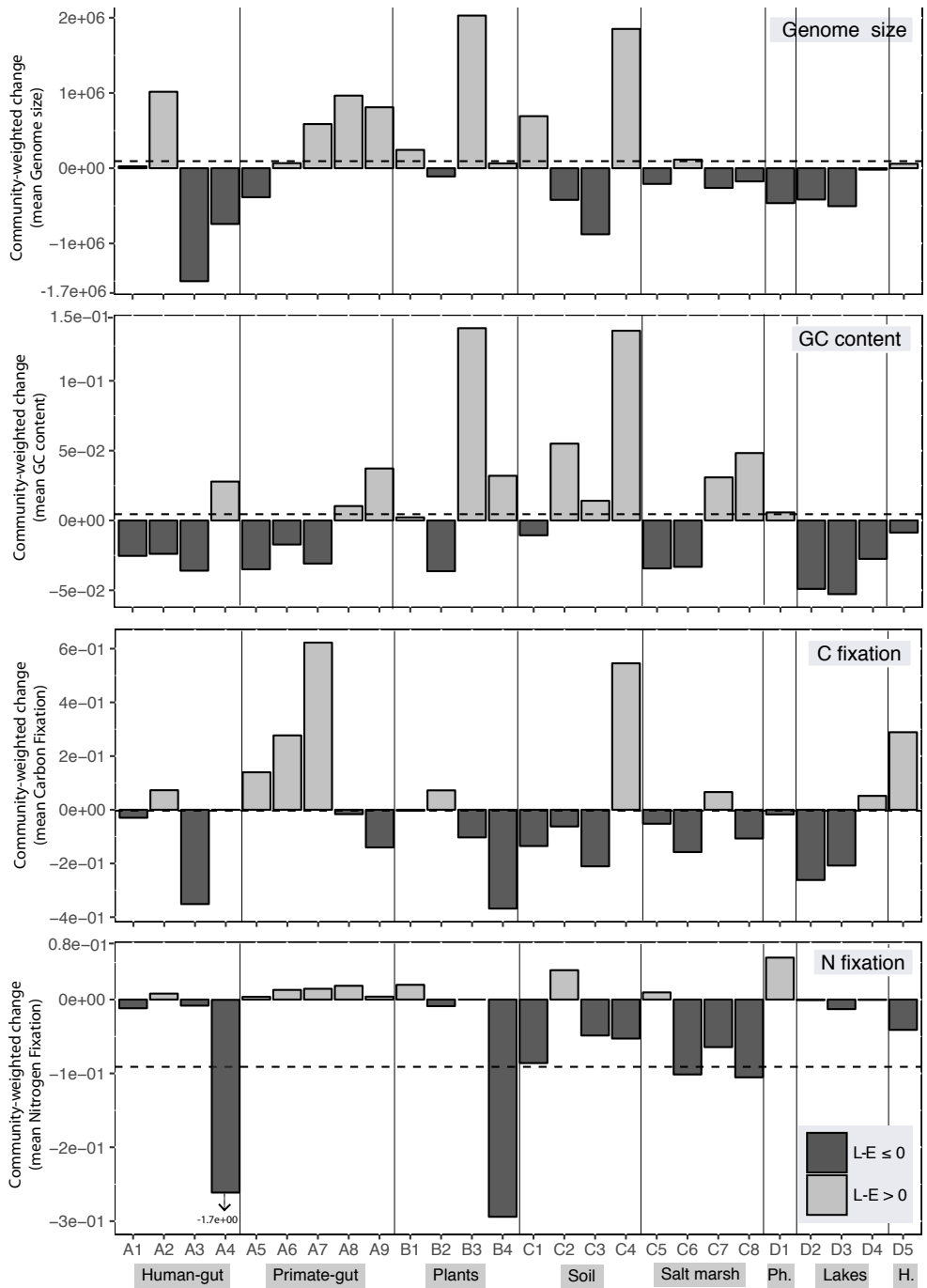
**Figure S1.** The proportion of matched sequences in early and late stages per habitat type.



**Figure S2.** Distribution of major microbial phyla across habitat types, for early and late stages.



**Figure S3.** Changes in the weighted occurrence of taxa (a) and functions (b) between early and late stages per sample.



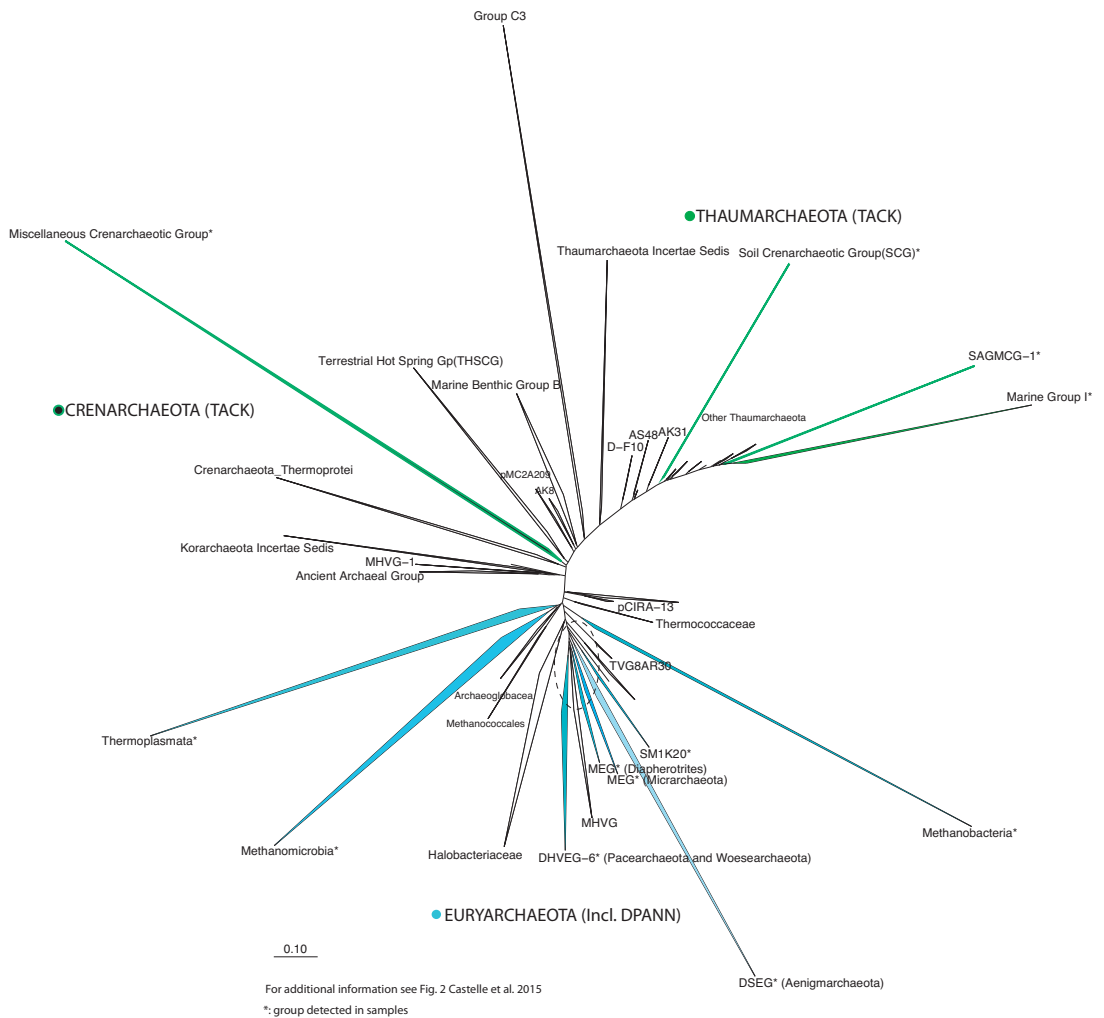
**Figure S4.** Changes in genome size, GC content, carbon fixation and nitrogen fixation community weighted functional traits.

Code	Authors	Broad habitat (A, B, C, D)	Habitat	Specific	Early	Late	Timespan	Replicates
A1	Costello et al 2013	Infant Gut microbiome	Human-gut	Infant-Gut	Day 8	Day 21	13 days	2
A2	Koenig et al 2011	Infant Gut microbiome	Human-gut	Infant-Gut	Day 5 and 10	Day 77 and 84	74 days	2
A3	Valles et al 2014	Infant Gut microbiome	Human-gut	Infant-Gut	Week 1	Week 9	63 days	2
A4	La Rosa et al 2014	Infant Gut microbiome	Human-gut	Infant-Gut	Day 8	Day 41	33 days	2
A5	McKenney et al 2015	Infant Gut microbiome	Primate-gut	Varecia	Birth	Nursing	11 days	2
A6	McKenney et al 2015	Infant Gut microbiome	Primate-gut	Varecia	Birth	Nursing	11 days	2
A7	McKenney et al 2015	Infant Gut microbiome	Primate-gut	Varecia	Birth	Nursing	11 days	2
A8	McKenney et al 2015	Infant Gut microbiome	Primate-gut	Lemur	Birth	Nursing	14 days	2
A9	McKenney et al 2015	Infant Gut microbiome	Primate-gut	Lemur	Birth	Nursing	14 days	2
A10	McKenney et al 2015	Infant Gut microbiome	Primate-gut	Lemur	Birth	Nursing	14 days	2
B1	Redford & Fierer 2009	Plant-associated	Plant-associated	Phyllosphere populus	May, Jun, Jul	Sept, Oct	150 days	1
B2	Malignien et al 2014	Plant-associated	Plant-associated	Phyllosphere Arabidopsis	Days 19 and 24	Days 67 and 73	54 days	3
B3	Bengtsson et al 2012	Plant-associated	Plant-associated	Kelp lamina	2 Months	11 Months	270 days	1
B4	Shade et al 2013	Plant-associated	Plant-associated	Apple flower	29-may	2-jun	5 days	2
C1	Brown & Jumpponen 2013	Soil development	Glacier chronosequence	Lyman Glacier (WA, USA)	150m	600m	?	2
C2	Knelman et al 2014	Soil development	Glacier chronosequence	Puca Glacier (Peru)	5yr	85yr	80 years	4 (E), 3 (L)
C3	Cutler et al 2014	Soil development	Lava chronosequence	Hekla Mt (Iceland)	164yr	852yr	688 years	2
C4	Fernandez-Martinez et al 2017	Soil development	Glacier chronosequence	Parry and Pia Glaciers (Chile)	1/5yr	19/66	66 years	2
C5	Dini-Andreote et al 2014	Soil development	Salt marsh chronosequence	Spring sampling	0years	105 years	105 years	3
C6	Dini-Andreote et al 2014	Soil development	Salt marsh chronosequence	Summer sampling	0years	105 years	105 years	3
C7	Dini-Andreote et al 2014	Soil development	Salt marsh chronosequence	Autumn sampling	0years	105 years	105 years	3
C8	Dini-Andreote et al 2014	Soil development	Salt marsh chronosequence	Winter sampling	0years	105 years	105 years	3
D1	Brasel et al 2015	Water-associated communities	Phormidium biofilm river	Phormidium (River)	Day 8	Day 29/32	24 days	3
D2	Casamayor unpublished	Water-associated communities	Saline shallow lakes	Camarón	New water November	January	50 days	2
D3	Casamayor unpublished	Water-associated communities	Saline shallow lakes	Guallar	New water November	February	50 days	2
D4	Casamayor unpublished	Water-associated communities	Saline shallow lakes	Pito	New water November	February	50 days	2
D5	Franzenburg et al 2013	Water-associated communities	Animal-associated	Hydra	Week 1	Week 15	105 days	2

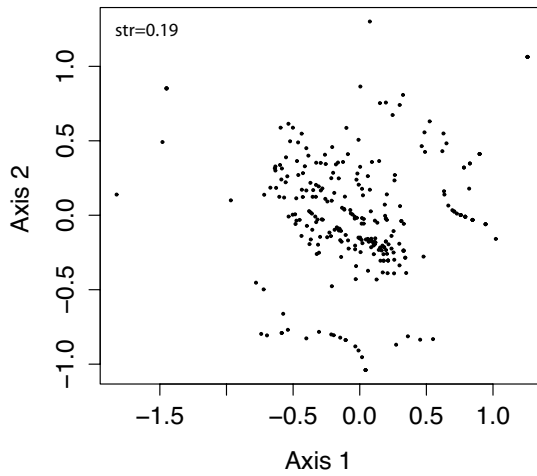
Supplementary Table S1 : Codes, references and brief information of the articles included in the meta-analysis.



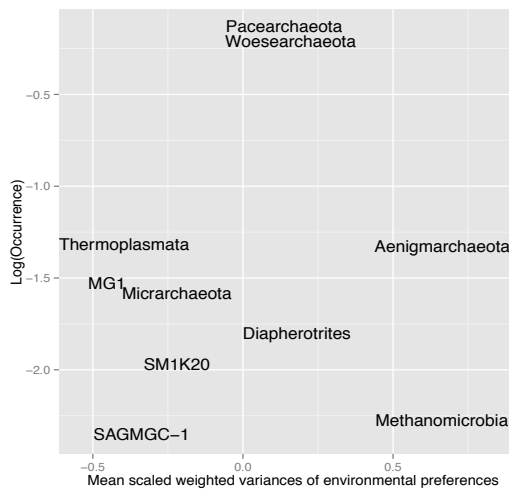
## SUPPLEMENTARY INFORMATION CHAPTER 3



**FigS1.** Archaeal lineages found in the Pyrenean lakes data set within the TACK, DPANN and Euryarchaeota superphyla.



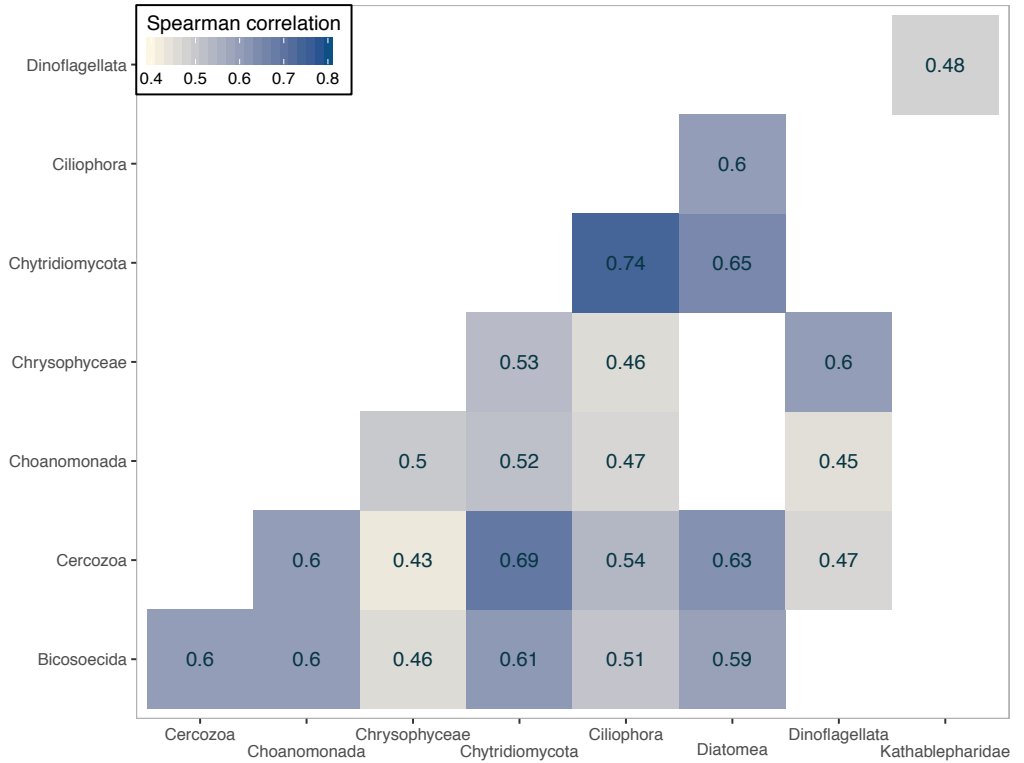
**Fig S2.** Ordination nMDS analysis of the lakes sampled based on the archaeal community composition.



**Fig S3.** Lack of correlation between environmental preferences weighted by relative abundance and occurrences.

## SUPPLEMENTARY INFORMATION CHAPTER 4

Figure S1



**Supplementary figure S1.** Correlations between taxonomic groups of interest based on average beta-dispersion per sub basin. Darker color indicates higher correlations. Only correlation values with strong significance ( $p < 0.001$ ) are displayed.



**Supplementary information S1.** *Cultured references that matched with sequences of previously detected organisms (through morphology and microscopy observations)*

The highly conserved nature of the V9 18S rRNA region may preclude species classification. However, some CCMs confirm some previous observations in high mountain lake planktonic communities or indicate knowledge gaps of particular interest. At this respect, the number of Chlorophyta OTUs with close cultured counterparts was remarkable (115). About one-third of them were present in more than ten lakes. The results confirm that many elongate flagellated forms are already known (e.g., some *Chlamydomonas* and *Monomastix* species). In contrast, the CCMs for chlorophyte coccal forms are poorly taxonomically determined (e.g. *Chlorococcum* spp., *Chlorella* spp., and *Coccomixa* spp.). Abundant OTUs classified as Cryptophyceae matched cultures of *Cryptomonas marssonii* and, *Plagioselmis nannoplantica*, among others. Some OTUs matched heterotrophic forms of the group such as *Goniomonas truncata* and several strains of *Chilomonas*, that were present in about 20% and 10% of the lakes, respectively. The few OTUs from Haptophyta matched cultures of Pavlovophyceae species, and a widespread one in the lake dataset matched *Chrysochromulina parva* (Prymnesiophyceae).

Among the aquatic fungi, less than 10% of the Chytridiomycota OTUs matched cultured strains. Most of them related to *Rhizophyidium*, *Rhizophlyctis* or *Rhizoclostridium*. Ten of the OTUs were widespread (present in 20-30% of the samples) but not particularly abundant. Some of the related cultures correspond to species growing on pine pollen (e.g. *Rhizophlyctis harderi* and *Kappamyces laurelensis*).

Cercozoa - the group with the highest diversity among phagotrophic heterotrophs in our samples – includes several life-forms with pseudopodia that correspond to the main phylogenetic branches (Cavalier-Smith, 2003)(Cavalier-Smith, 2003). The most common OTU (20% samples) matched cultures of the thecate *Rhogostoma schuessleri* (Cl. Thecofilosea). Some flagellated forms were also common; OTUs matching *Cercomonas rotunda* and *Paracercomonas crassicauda* cultures were present in about 5% of the samples among other species of the same genera. The species with silica plates (Imbricatea) were less frequent (< 1%) but showed matches with a large diversity of genera (*Allas*, *Assulina*, *Peregrinia*, *Euglypha*, *Spongomonas*).

The SAR groups largely differed regarding culture references. Ciliophora showed a low number of CCM and only an OTU matching a cultures of *Rimostrombidium lacustris* (20% of samples) was relevant. On the other hand, Dinoflagellata showed CCM for many of the most ubiquitous and abundant OTUs. Several OTUs matching the same culture of an undetermined *Gymnodinium* were ubiquitous and indicate a potentially large hidden variability in these abundant small dinoflagellates. Similarly, may occur with *Woloszynskia*-like species. The most abundant OTU in the complete survey belongs to this group. Only two OTUs closely matched cultured strains *W. leopoliensis* and *W. pascheri*, present in 10 and 2% of the samples,

respectively. Among armoured forms, *Peridinium inconspicuum* and *P. cinctum* cultures are close to OTUs present in 3 and 10% of the lakes. Within Chrysophyceae only 51 OTUs had a CCM reference, despite being the richer protist group in the lake samples. OTUs related to cultures of several *Dinobryon* species, conspicuous forms with large colonies of cells with theca, were regularly present in many lakes. Among the OTUs with high occurrence, several matched cultures of flagellated forms with siliceous scales (Synurales), which are currently split from the Chrysophyceae into a separate class (Synurophyceae). OTUs matching *Mallomonas* species were present in more than 50% of the lakes. Most Chrysophyceae are planktonic, therefore, the appearance in a few lakes of OTUs matching benthic forms, rarely reported in field surveys, are worth to mention (e.g., *Chrysochaete*, *Chrysonephele* and *Chrysocapsa*). In contrast, the naked flagellated autotrophic forms, which are extremely rich in these lakes only showed a few cultured matches. OTUs matching cultured heterotrophic Chrysophycean forms were abundant. Unfortunately, the taxonomy of the cultures was poor; most of them are simply labelled as “*Spumella*-like”. Some diatom OTUs showed a high matching with cultured referents; unfortunately the latter were poorly determined. They mostly belonged to Mediophyceae, centric diatoms that are truly planktonic. An exception to the general taxonomic uncertainty was *Fragilaria nanana*, with a 100% match with one of the OTUs and present in 25% of lakes.

## **Supplementary information S2:** *Comparison of pyrenean freshwater and mediterranean sea surface samples*

### *Context*

The comparison between freshwater and marine planktonic systems of similar primary production can enlighten the general ecological constraints and evolutionary trajectories that have been shaping planktonic communities. For example, Pyrenean lakes and Mediterranean open waters show roughly a similar primary production, about  $100 \pm 50 \text{ g C m}^{-2} \text{ yr}^{-1}$  (Catalan *et al.*, 2006; Siokou-Frangou *et al.*, 2010). Species of oligotrophic freshwater and marine planktonic assemblages have been evolving in ecosystems of similar flow of autochthonous carbon but contrasting environmental conditions, the size of the system and inputs of subsidiary organic matter from neighbouring ecosystems. Therefore, one may expect both common and contrasting features between oligotrophic marine and freshwater planktonic communities. We compared the diversity patterns found in our survey with those of the Tara Oceans expedition in the Mediterranean system (Vargas *et al.*, 2015). Commonalities and major differences provided elements for discussing on how the common oligotrophy but the different environmental setting may have been influencing the evolutionary history and the ecological configuration of these planktonic systems.

### *Specific methods*

We downloaded a set of 22 marine libraries of the V9 18S rRNA region, from the Tara Oceans expedition (Vargas *et al.*, 2015) to compare the richness of different groups. We chose subsurface (mixed layer) Mediterranean samples, because of the similar primary productivity status to the lakes studied (Siokou-Frangou *et al.*, 2010). We selected samples from 0.8-5  $\mu\text{m}$  and 5-20  $\mu\text{m}$  sizes, which despite dismissing organisms  $>20 \mu\text{m}$  was a close equivalent to the plankton size typically found in the Pyrenean lakes (0.22-50  $\mu\text{m}$ ) (Sarmiento *et al.*, 2015).

Data filtering and classification was carried out with the same parameters as for the Pyrenean samples, except filtering to a read length of 163. For simplicity, the comparison was made using a single classification method (RDP classifier). To minimise information loss and potential sampling strategy bias, we established a compromise between the number of samples and the sequencing depth for each sequence pool. A total of 227 Pyrenees samples were rarefacted to 10,000 sequences (2,270,000 sequences). And, in accordance, 11 Tara samples (for each filter size) were rarefacted to 200,000 sequences (2,200,000 sequences) in order to equalize sequencing effort between studies. Good's coverage was 0.77 and 0.74 for the marine pools and 0.94 for the Pyrenean one, indicating a better capture of the regional diversity in the latter pool, but still adequate for our goal. We calculated richness and Shannon diversity in the marine samples for alpha diversity comparison with the Pyrenean samples.

*Comparison results*

The diversity comparison between the Pyrenean lakes and the Mediterranean Sea samples from the Tara Oceans expedition cannot be straightforward because of the size partition in the marine survey and its much larger number of samples and sequences. Therefore, we concentrated in the main taxonomic groups of the Pyrenean lakes and compared them with the rarefied results of the two ocean size-filters (Suppl. Table 1, Figure 5). Comparing individual samples, the average richness and Shannon diversity ( $H'$ ) per group were higher in the marine samples than in the freshwater ones (Suppl. Table 1) with a few exceptions. Chrysophyceae showed more species per lake (average, 68) than per marine spot (27, adding the two filters), although  $H'$  was slightly higher in the marine samples of the small filter (0.8-5  $\mu\text{m}$ ). In general, this filter showed much higher  $H'$  values than the 5-20  $\mu\text{m}$  filter, except for diatoms. The richness of Ciliophora and Kathablepharidae was similar between the freshwater and marine samples, although  $H'$  was more than twice higher in the marine samples; indicating a more evenly distributed abundance of the species for these groups in the sea. Chytridiomycota richness and  $H'$  were similar to Chlorophyta and Cercozoa in the freshwater samples (about ten species per sample on average) whereas they were absent in the rarefied samples of the Tara Oceans expedition. Beyond these exceptions, the other groups were richer and more diverse in the marine samples than in the mountain lake samples. The difference was particularly outstanding for Dinoflagellata, with an order of magnitude more species in both marine filters (ca. 300 species per sample) than in the freshwater ones (ca. 30) although this group places second in the lake's richness rank.

**Table S1:** Average richness and Shannon diversity for each group for Mediterranean Sea samples (size fractions '0.8-5  $\mu\text{m}$ ' and '5-20  $\mu\text{m}$ ') and Pyrenees samples (a single size fraction Med (0.22-20  $\mu\text{m}$ ), after rarefaction to 10000 all samples and transform them to relative abundances. OTU classification into taxonomic groups was done using RDP classifier only.

Group	Richness (OTU nr)			$H'$		
	Pyrenees	Med (0.8-5 $\mu\text{m}$ )	Med (5-20 $\mu\text{m}$ )	Pyrenees	Med (0.8-5 $\mu\text{m}$ )	Med (5-20 $\mu\text{m}$ )
Chytridiomycota	10	0	0	1.42	0	0
Ciliophora	2	2	2	0.33	0.65	0.62
Kathablepharidae	6	7	2	0.45	1.66	0.59
Choanomonada	2	6	1	0.28	1.66	0.34
Cryptomonadales	6	10	2	0.55	1.89	0.5
Diatomea	4	23	31	0.55	2.37	2.31
Chlorophyta	10	17	6	1.15	2.46	1.05
Cercozoa	11	19	6	1.37	2.57	1.19
Chrysophyceae	68	25	2	2.17	2.8	0.62
Dinoflagellata	30	321	216	1.44	4.5	4.12

The high richness and diversity of Dinoflagellata and Diatomea per marine sample remain when the overall metacommunity diversity is compared amalgamating all the samples and rarefying per individuals (Figure S2). However, the situation changes for the other groups. The freshwater metacommunity results more diverse for Chrysophyceae, Chlorophyta, Cryptomonadales, Cercozoa, Katablepharidae and Ciliophora, and similar for Choanomonada. This difference between average features per sample (alpha diversity) and the patterns of the overall ensemble (gamma diversity) indicates more different assemblages between the mountain lake samples than between the marine ones. This indicates higher beta-diversity in the freshwater studied area than in the marine samples.

## References

- Catalan J, Camarero L, Felip M, Pla S, Ventura M, Buchaca T, *et al.* (2006). High mountain lakes: extreme habitats and witnesses of environmental changes. *Limnetica* **64**: 123–145.
- Cavalier-Smith T. (2003). Protist phylogeny and the high-level classification of Protozoa. *Eur J Protistol* **39**: 338–348.
- Sarmiento H, Casamayor EO, Auguet JC, Vila-Costa M, Felip M, Camarero L, *et al.* (2015). Microbial food web components, bulk metabolism, and single-cell physiology of piconeuston in surface microlayers of high-altitude lakes. *Front Microbiol* **6**: 1–12.
- Siokou-Frangou I, Christaki U, Mazzocchi MG, Montresor M, d'Alcala MR, Vaquer D, *et al.* (2010). Plankton in the open Mediterranean Sea: a review. *Biogeosciences* **7**: 1543–1586.
- Vargas C De, Audic S, Henry N, Decelle J, Mahé F, Logares R, *et al.* (2015). Eukaryotic plankton diversity in the sunlit ocean. **348**: 1–12.



