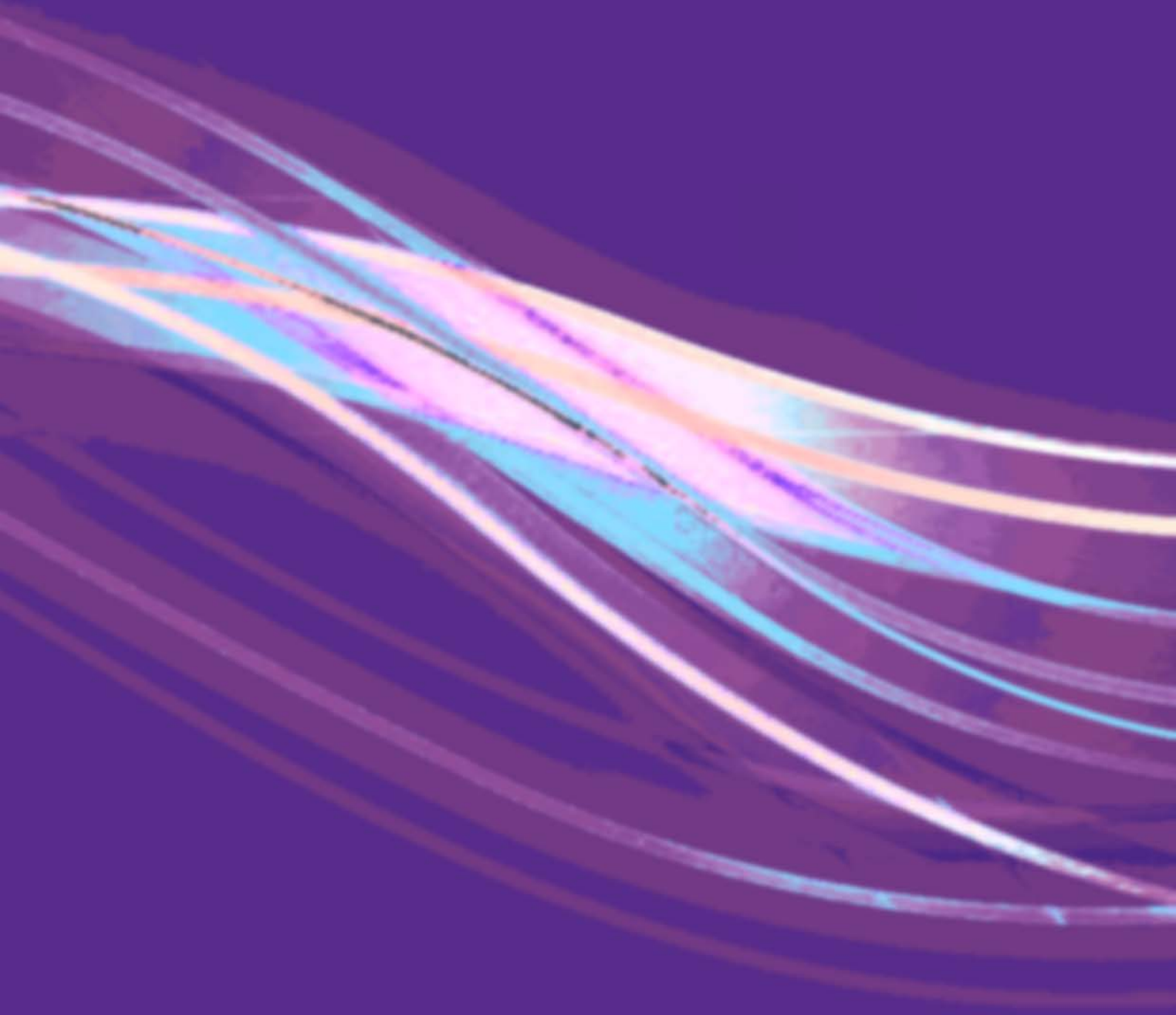


Synthesis of results and general discussion



The first objective of this thesis was to improve the information about the genetic structure of microeukaryotes groups, mainly at class-rank level, with the final goal of building a database of well-curated and reliable sequences. Once defined this reference sequence dataset, employing only molecular surveys done by Sanger sequencing, it was used as a framework for studying the global diversity of microeukaryotes in the deep ocean using high-throughput sequencing. The diversity of deep microeukaryotes plus the data collected on their abundance and biomass, allowed drawing a refined picture of the bathypelagic environment.

The molecular taxonomy challenge: tools to define a group

The V4 region of 18S rDNA: the best compromise

The management of a large number of sequences that can be retrieved in molecular surveys needs a great initial effort in order to establish good working criteria. Sequences need to be organized in categories (OTUs) defined by given levels of similarity in order to get quantitative values of diversity. The first step of our work was defining the target region within the 18S rDNA, since no sequencing technology is currently ready to analyze the complete gene. In particular the choice was between V4 and V9 regions. This debate was born inside “BioMarKs”, an European research project that studies protistan diversity with molecular and microscopical techniques. Seminal studies using 454 pyrosequencing focused on the V9 region (Amaral-Zettler *et al.* 2009; Cheung *et al.* 2009, Stoeck *et al.* 2009), which is a very short region (up to 150 bp) and was optimal for the sequencing technology at that time. With technical advances yielding longer reads (currently up to 400 bp but longer sequences are expected in the near future) it was logical to analyze longer hypervariable regions, such as the V4. We contributed to this discussion by looking for the region that better represented the entire gene. Our results suggested that the variability detected in the V4 region (around 500 bp) was a good indicator of the variability that would be seen analyzing the entire gene. The respective plots (Figure 1, Chapter 2) showed that the slopes of the regression lines between the distances calculated with the entire gene and the V4 region were around 1.4 for three different Supergroups, so distances calculated with the V4 region could be translated to distances with the entire gene by dividing by this value.

How to define a taxonomic class: from phylogeny to clustering

Sequencing technology improves fast, maybe faster than our ability to manage it. Seminal molecular cloning methods typically yielded between 100 and 500 sequences per sample. Despite this number is very low compared with the output of advanced high-throughput sequencing methods,

the overall process is more controlled and the sequences obtained are longer. Thus, the output of clone libraries is very useful to build a reliable alignment that is the base for a correct phylogeny, as we have done in the first chapter (Figure 1). During the “age of clone libraries” building a tree was the best way to define a taxonomic group but nowadays the superproduction of sequences makes this operation more difficult, particularly aligning a great number of genetically distant sequences. Certainly, short sequences can not resolve the taxonomic relationships among distant lineages. In the “age of pyrosequencing” there is a progressive skip from phylogenetic trees to clustering approaches by which sequences are grouped in taxonomic units based on sequence similarity. Grouping of sequences in a tree is based on patristic distances between sequences (branch lengths), which depend on the number and variability of sequences considered and compares each sequence and all the rest. In contrast, the OTU clustering is based on similarity (or corrected genetic distances such as Jukes Cantor) between all pairs of sequences. The absolute value of similarity or distance is independent from the number of sequences considered, since only pairwise comparisons are performed. The scientific community is adopting as routine the clustering of sequences based on similarity, the problem is that in this process the shape of the tree is lost, so now is more difficult to identify outlier and fast-evolving lineages. In the first chapter we tested if the number of OTUs from the two approaches was different, and we found that at distances up to 0.10, the grouping using JC distances (almost equivalent to similarity) or patristic distances (from the phylogenetic tree) showed good correspondence (Figure 2, Chapter 1). Considering that often the clustering is done at 95 to 97% similarity, this is an acceptable result. Despite patristic distances would result in a more accurate and evolutionary robust clustering, similarity clustering is used with 454 datasets, since the alignment step is skipped and it provides a simple and intuitive way to analyze a high number of sequences.

Setting the limits of taxonomic groups

Working with clustering methods it is essential to answer two questions: what is the cut-off that defines an OTU with useful biological meaning (i.e. a species) and what is the maximal distance that can be found within a given group. About the first question, different authors have proposed different cut-off levels (Worden 2006, Jeon *et al.* 2006, Caron *et al.* 2009) but there is little support to justify each hypothetical level, and it would be important that the scientific community arrives at some conclusion into this direction. Due to intragenomic polymorphisms (Introduction, Table 1) and low-frequency sequencing errors we think it is not advisable to use 100% similarity as the OTU definition. A value between 97-99% similarity appears to be a more reasonable criterion because it is high enough to be rather strict, but not so high as to separate sequences only due to intragenomic polymorphism or sequencing errors. In the chapter two we addressed the second

question for groups roughly corresponding to taxonomic classes in classical systematics, pushed by the lack of the graphical output of the tree that made necessary an alternative way to rapidly identify outlier sequences. We performed trees to be sure about the affiliation of the sequences, which were later analyzed clustering group by group. We found that 75% of the class-rank groups had a corrected maximum pairwise genetic distance below 0.25. This is now our general reference to the maximum distance allowed within a class, and it is useful value to interpret the taxonomic equivalence of environmental ribogroups.

The importance of a curated reference database

A good reference database is an essential tool for any molecular study based on short sequences, but current databases target only prokaryotes (GreenGenes) or give a less accurate treatment to eukaryotes (SILVA). The main problem of SILVA is that often lacks of good taxonomic assignments for eukaryotic sequences, especially for the newly discovered ribogroups. A new tool for the identification of eukaryotic 18S rDNA has been published very recently, the PR2 database (Protist Ribosomal Reference Database, Guillou *et al.* 2013). This tool did not exist at the beginning of this thesis. In this frame, I want to highlight the importance of the pool of well-curated sequences from chapter two (8291) that constitutes the core, improved with PR2, of an in-house reference database (MAS9013), which has been used for taxonomic assignation and chimera detection in the second part of the thesis and in other publications in preparation.

Typical composition of epipelagic microeukaryotes

The high-rank diversity observed in chapter 1 in terms of relative abundance of specific lineages (Figure 1), is the typical found in other molecular surveys of marine picoeukaryotes (Massana and Pedrós-Alió 2008, Vaulot *et al.* 2008) and resembles the abundance of groups in chapter 2 (Table 1). Alveolates, mainly MALV-I and MALV-II, dominated the community and represented 47% of the clones, followed by Stramenopiles (19%) and Rhizaria (13%). Fungi were not considered in the first two chapters, since generally they are little represented in the epipelagic environment, globally they are less than 1% of the sequences in clone libraries (Massana and Pedrós-Alió 2008). Differences in the taxonomic composition of epipelagic and deep microeukaryotes, even at high rank clustering levels, are evident (Figure 6, Chapter 4) and will be analyzed in the second part of this discussion. Interestingly, at low distances (minimal 1300 Km), samples strongly differed when analyzed by clone libraries (Figure 6, Chapter 1), and at that time this was explained by undersampling. However, at comparable distances and with a major sequencing effort, we see that differences among deep samples are still present (Figure 5, Chapter 4). This is a clear signal of the strong effect of the environment on community selection.

The deep ocean

Counting microeukaryotes: the need of flow-cytometry

As far as we know this is the first study that applies flow-cytometry, together with microscopy, on a large scale investigation of the abundance of microeukaryotes (Chapter 3). Epifluorescence microscopy is time consuming and is prone to errors of the operator, while flow-cytometry presents other types of problems. Whereas it is possible to identify several populations of photosynthetic microeukaryotes thanks to their pigments (Olson *et al.* 1985; Li *et al.* 1994; Marie *et al.* 2001), to detect heterotrophic microeukaryotes a general stain (in this case SYBR Green) is required. However, this does not discriminate between prokaryotes and eukaryotes, and a continuum between large bacteria and small eukaryotes exists. To solve this problem, we used epifluorescence microscopy on selected samples to set the counting gate in the cytometry software (then applied to the complete dataset) and to check stations with unrealistic abundances. An alternative method, to save time and recover the size and shape of the cells, would be automatic epifluorescence microscopy, which will be implemented in the future in our lab. The comparison of flow-cytometry and microscopic counts (Figure 2, Chapter 3) was very good ($R^2=0.82$, $p<0.0001$). Therefore, a large number of samples were then processed by flow-cytometry. The abundance of microeukaryotes thus determined was one of the two parameters in the description of the deep ocean global community.

General features of microeukaryotes in the bathypelagic ocean

Considering the bathypelagic region (1000 to 4000 m), also the focus of the diversity study, the abundance of microeukaryotes averages 14 cells mL⁻¹. This concentration is not constant, and this is particularly evident in the South Pacific, where there is a peak of 58 cells mL⁻¹ in the deepest sample of station 98. Regarding cell size structure, the percentage of very small cells (equivalent diameter <3 μm) decreases with depth (Figure 6, Chapter 3) and from the pictures taken for biomass measurements, we know that some of these cells are clearly flagellated. The average biomass of microeukaryotes is 50 pg C mL⁻¹ in the 1400 to 4000 m layer. For the diversity study of the deepest samples (~4000 m), we filtered 120 L, meaning that we collected about 1,320,000 cells per sample. Most of the sequences retrieved belong to Rhizaria, followed by Alveolata and Stramenopiles (Figure 6, Chapter 4). At a local level and at lower taxonomic rank, communities are dominated basically by three classes (Collodaria, Chrysophytes, Basidiomycota) and one ribogroup (MALV-II). Formally Collodaria is an *order*, but considering the value of its maximal distance retrieved in the first chapter (Table 1) and that Polycystinea is not a monophyletic group

(Figure 3b, Chapter 2) it was regarded as a *class*. Differences in microeukaryotic abundance and diversity among deep samples are clearly related both to abiotic (oxygen, temperature) and biotic (prokaryotic and viral abundance) parameters.

Culturing bias and deep protists

Surprisingly the analysis of the first twenty OTUs that represent 50% of total reads show several examples of high similarity with cultured organisms (Table 3, Chapter 4). This fact was previously reported for Fungi (Bass *et al.* 2007, Richards *et al.* 2012) but was unexpected for the other groups. For example three OTUs that explain 79% of the chrysophyte tags are very similar to cultured species. We know that in the epipelagic environment there is normally little agreement between the diversity retrieved in molecular surveys and in culture-based surveys (Massana *et al.* 2004a), and considering the environmental characteristics of the dark ocean and its relative isolation we expected to find even less agreement. However, one third of the most abundant OTUs are more than 97% similar with a cultured species and represent totally 28% of the tags. Moreover two of these OTUs are 99% similar to cultivated species. Considering that most cultured species derive from epipelagic samples, we can deduce that at least one quarter of the tags are shared between surface and deep communities. This proves the great capacity of adaptation of microeukaryotes to different environments. In addition, the different impact of the culturing bias in surface and deep communities remains as an intriguing observation worth to be further addressed.

Ubiquity: Everything is everywhere?

Finlay *et al.* (2004) stated the difference between “ubiquity” and “ubiquitous dispersal”. They claimed that most protists have ubiquitous dispersal, implying that they are not necessarily found everywhere but should be present in suitable habitats around the world (Caron *et al.* 2009). Another important aspect of this debate is that the presence of the same sequence in separate oceans can give us information about the ubiquity of the respective species whereas the absence can be simply due to undersampling or to temporal successions. In the Malaspina dataset, 42 OTUs are present in the total of 27 stations, and these represent 80% of the pyrosequences. This qualitative view of communities is pretty close to the idea that “everything is everywhere”. However, following “the environment select” concept, the distribution of these OTUs is not uniform, and an OTU that belongs to the rare biosphere in one sample often is the dominant OTUs in another, as observed in several examples, for instance with Fungi. As commented in the introduction, the environmental homogeneity of the deep ocean is an old and misleading idea and, despite putative high dispersal ability, there is a strong environmental effect on community composition. Probably the intrinsic properties of deep water masses plus the presence of preys or alternative food allows

the existence of different trophic niches.

Phagotrophy: the relation with prokaryotes

Phagotrophy was expected to be the principal trophic pathway for microeukaryotes in the deep ocean. To test this hypothesis, we first analyze the relation between the abundance of prokaryotes and microeukaryotes, and second the relation between their ratio and the diversity. Considering the entire dark water column, the abundance of microeukaryotes correlates well with that of prokaryotes ($R^2=0.50$, $p=0.0001$), except in South Pacific stations ($R^2=0.08$, $p>0.05$, Figure 5, Chapter 3). Subsequent multiple regression analyses proved that this result is independent from depth (Chapter 3). However, despite this significant relationship, prokaryotes explain only a part of the variance of the microeukaryotes abundance. We use prokaryotes abundance also to explain the variability in diversity but the result is not significant ($p>0.058$), although 34% of the variability is significantly explained by the ratio Prokaryotes to microeukaryotes ($p=0.001$). Low values of this ratio, similar to the ones in epipelagic waters (ca. 2000), correspond to communities dominated by Collodaria and Chrysophyceae, suggesting a putative grazing role for these two classes. Considering that less than half of the variability of abundance and diversity is explained by prokaryotes and that the abundance ratio has an average value higher than at surface, we conclude that phagotrophy in the deep ocean seems to give also a significant space for other trophic pathways such as osmotrophy and parasitism.

Osmotrophy: the role of Fungi

Comparing the prokaryotes to microeukaryotes abundance ratio with the relative abundance of several taxa, only Fungi presents a significant relationship (Figure 7, Chapter 3). Considering that Fungi are proved osmotrophs and probably unable to perform phagotrophy (Richards *et al.* 2012), this suggests that where the community is dominated by Fungi there is a lower grazing pressure on prokaryotes, and this could favour the high abundance ratio.

As seen in the introduction (Figure 7b) the distribution of DOC is not constant in the bathypelagic region. In fact, DOC decreases along the deep conveyor belt resulting in an overall higher concentration in Atlantic than in Pacific waters. It is possible to associate part of the decrease of DOC from Southern Ocean to North Pacific to the presence of Fungi in these waters. However, considering that DOC is more concentrated in the Atlantic Ocean, especially in the north, is difficult to understand why Fungi generally do not thrive in these waters. I present next several possible explanations for the absence of Fungi in the Atlantic Ocean:

- *Fungi versus prokaryotes*. A first simple hypothesis is that prokaryotes are the main con-

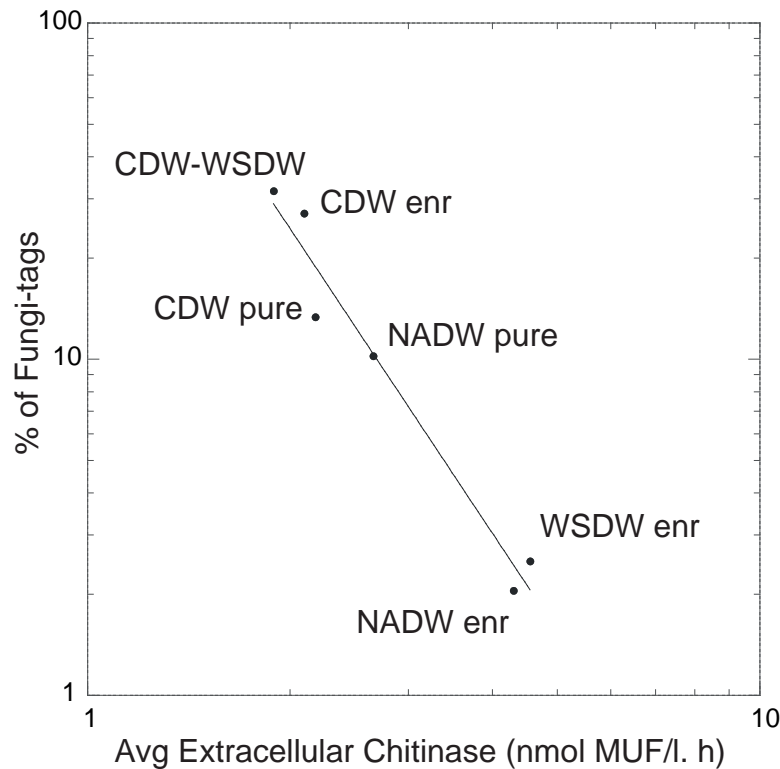


Figure 1. Relationship between the average extracellular chitinase concentration and the average fungal tag abundance in the different water masses defined.

sumers of DOC in Atlantic waters and outcompete Fungi. An antagonistic relation of Fungi and bacteria has been seen in previous experiments, as for example in Moller *et al.* (1999) where Fungi and Prokaryotes clearly compete for the same DOC.

- *Fungi versus Chrysophytes.* Chrysophytes can be phagotrophs and osmotrophs (Sandgren *et al.* 1995, Sanders *et al.* 2001), and they may survive thanks to a combination of these two strategies. Except in station 32 they never share dominance with Fungi. Våge *et al.* (2013) built a model to test the importance of mixotrophy compared with pure osmotrophy. They demonstrated that at low size ratio between prey (prokaryotes) and predators (Chrysophytes), as occurs in the deep ocean, for a mixotroph is very convenient the “eating the competitor” strategy (Thingstad *et al.* 1996). So in this case mixotrophs (Chrysophytes) could suppress pure osmotrophs (Fungi) by foraging on them, and this could happen in Atlantic and North Pacific waters.
- *Fungi and recalcitrant carbon.* The presence of Fungi in older waters poorer in DOC could be explained by their specialization to assimilate recalcitrant carbon. A possible mechanism is the secretion of superoxide molecules, in particular oxidized forms of Mn, which oxidizes recalcitrant carbon to more bioavailable forms (Hansel *et al.* 2012). This mechanism is probably shared with bacteria. So, another explanation of higher ratio of

Prokaryotes to microeukaryotes is that in a Fungi dominated community prokaryotes could also thrive on recalcitrant DOC.

- *Fungi versus chitinase*. One of the typical features of Fungi is the presence of chitin in the outer wall (Richard *et al.* 2012). During the Malaspina expedition we measured extracellular enzymatic activities, including chitinase, the enzyme that digests chitin. The chitinase activity turns out to be higher in the Atlantic than in Pacific and Indian Oceans. Considering averaged values for each defined watermass, the relative abundance of Fungi tags show an inverse relationship with the chitinase activity (Figure 1) in a highly significant correlation ($p=0.0005$; R^2 of 0.95). Chitinase enzymes could be produced by prokaryotes as well as microeukaryotes (Cottrell *et al.* 2000). Although it is unknown if Fungi are the principal targets of this enzyme, the extracellular chitinase probably produces an environment not favourable for their life.

Osmotrophy is also present in other taxonomic groups, such as the labyrinthulids (Raghukumar *et al.* 2001) or Excavata (Lara *et al.* 2009). Indeed, the extent of the osmotrophic process should be studied to understand the impact of heterotrophic protists in the global carbon balance.

Parasitism: the hidden relationships

The inference of parasitic interactions from sequencing data is quite hard because there is no clear correspondence between host and parasite abundance (Skovgard *et al.* 2014). And, more severely, a naked sequence sometimes tells very little about ecological performance. Several marine microeukaryotic clades are considered to be parasites, being MALV-I and MALV-II the most abundant. Several species within dinoflagellates and Fungi can also be parasites. The supposed hosts of these parasites are other microeukaryotes and also macrofauna. In our dataset, the relative abundance of MALV-II tags have a significant correlation with the relative abundance of Metazoan tags ($R^2=0.45$, $p=0.0005$) (Figure 2). This relation is better for MALV-I ($R^2=0.60$) and less strong for dinoflagellates ($R^2=0.41$), but undetected for fungi or the other classes. The three relations are particularly evident in Atlantic samples, where R^2 is respectively 0.85, 0.89 and 0.76. Generally parasitism is strongly present in Fungi and probably also in its marine clades, especially since the principal OTU is highly similar to parasitic species. Indeed, *Engyodontium album* is a parasite of the *Felis domesticus* (Dennis 1995). Nevertheless, assuming a random distribution of the hosts it is expected a random distribution of parasite tags. In this sense, the distribution of Alveolate tags better fits with this scenario compared with the non-random Fungi distribution.

The last decades of protist research have focused on the study of the diversity of microeukaryotes,

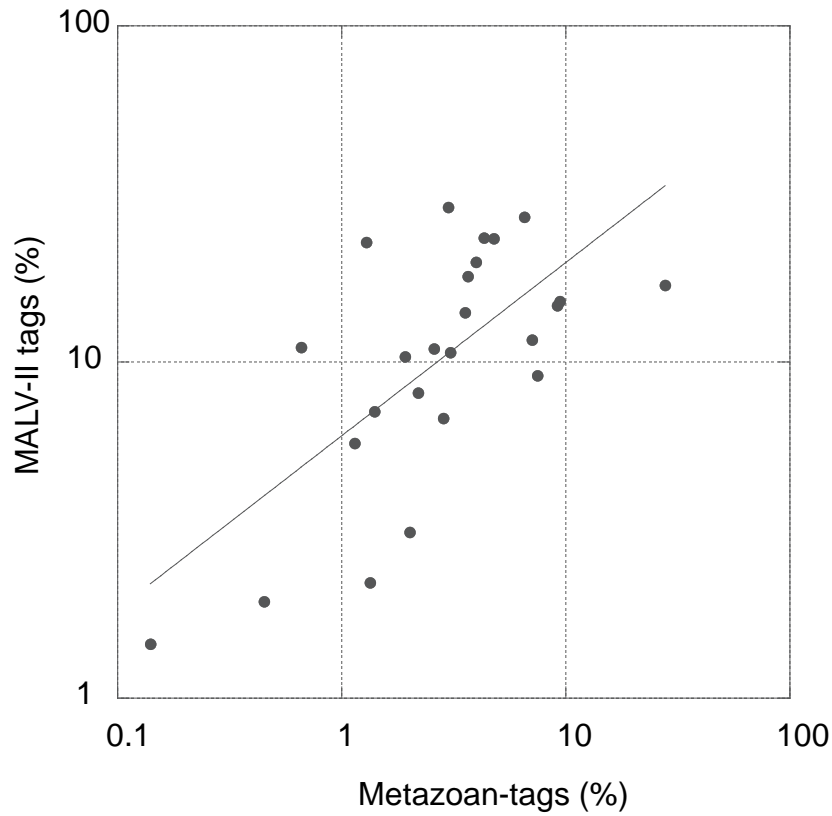


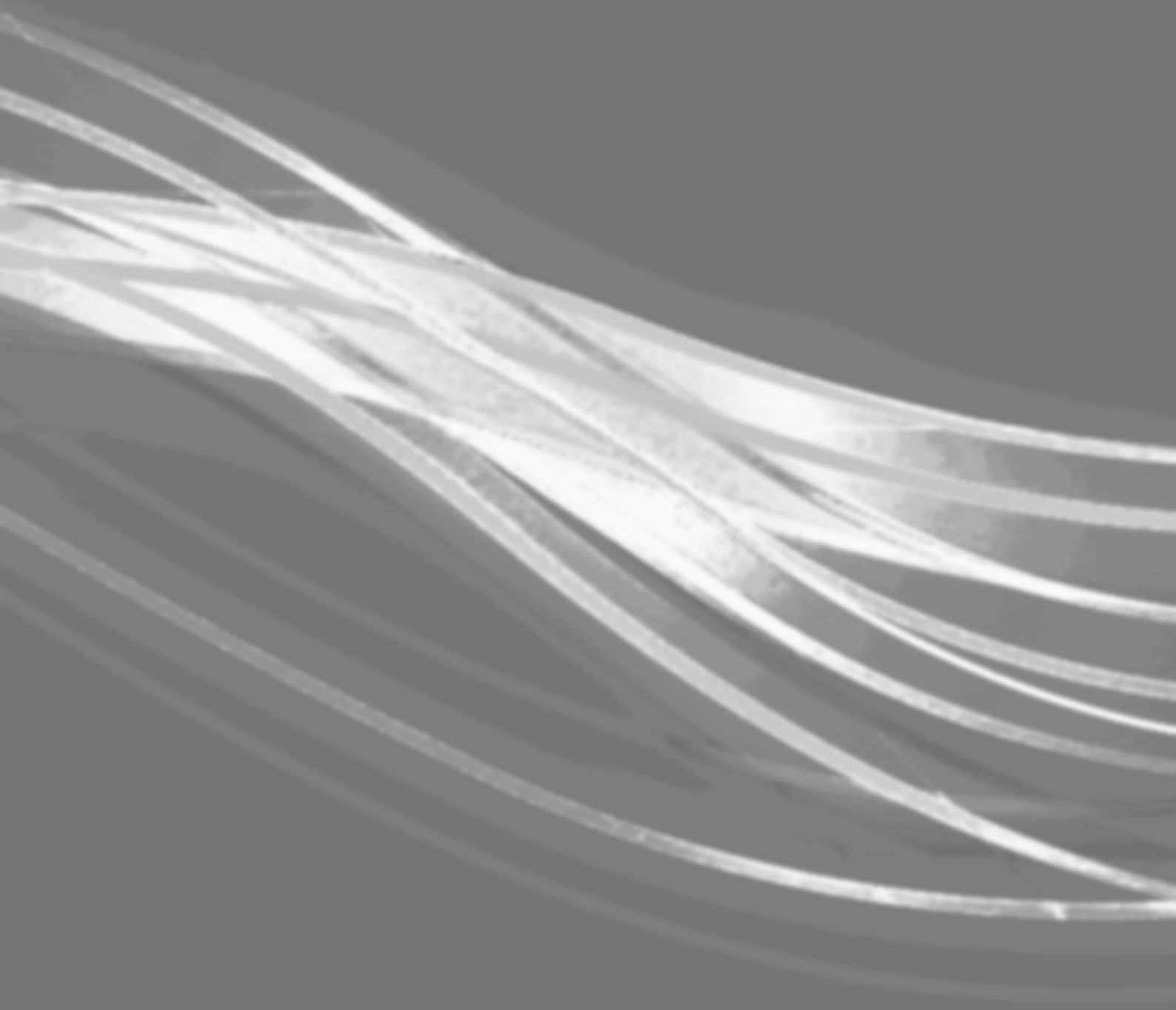
Figure 2. Relationship between the relative tag abundance of MALV-II against that of metazoans.

at the beginning defined as “unexpected” in terms of high and novel diversity. Now that we expect such amazing complexity of taxa composition, the attention is moving to the investigation of the function of the different taxa in the ecosystem. The assignation of a clear ecological role, through the classical method of culturing and new single-cell genomic approaches, will soon improve the vision that we have now on such important environmental players.

Conclusions

- 1) The V4 region of the 18S rDNA is better than the V9 region for representing the variability of the entire gene. On average, the variability contained in the V4 region is 1.4 times that of the complete 18S rDNA gene.
- 2) Typically, the maximal genetic distance among sequences belonging to a same eukaryotic taxonomic class is 0.25. This value can be used to assess the taxonomic equivalence of environmental ribogroups.
- 3) Epipelagic microeukaryotes communities are typically formed by Alveolata (47% of sequences), Stramenopiles (19%), and Rhizaria (13%). Additional groups belong to CCTH and Archaeplastida. Often Fungi and Excavata are very low (less than 1%) or undetected.
- 4) The abundance of microeukaryotes averages 54 cells mL⁻¹ in mesopelagic samples and 14 cells mL⁻¹ in bathypelagic samples. This variability is explained mainly by depth, prokaryotes abundance and oxygen concentration.
- 5) The cell size of planktonic microeukaryotes increases with depth. Cells larger than 4 μ m represent 12% at 200 m and 22% at 4000 m. Total biomass ranges from 280 pg C mL⁻¹ in the upper mesopelagic layer to 50 pg C mL⁻¹ in the deepest layer.
- 6) The diversity assessed by 454 pyrosequencing compares very well with a parallel metagenomic survey. In general, the percentage of eukaryotic supergroups is very similar by the two approaches (Figure 7a) and the same OTUs are retrieved. The main difference is a much higher representation of Excavata in miTags and a lower representation of Alveolata.
- 7) Four abundant classes generally compose the bathypelagic community of microeukaryotes: Collodaria, Chrysophyceae, MALV-II and Basidiomycota. However, the relative abundance of these classes varies a lot among samples.
- 8) The variability in community composition between samples is well explained by the water mass they belong (26% of variability) and by the abundance ratio between prokaryotes and microeukaryotes in the respective samples (34%).

Resumen de la tesis



Una breve historia sobre la investigación de los protistas marinos

La definición de “protista” abarca una gran diversidad de organismos. En términos generales, los protistas son microorganismos eucarióticos. De hecho, esta definición no tiene un verdadero significado evolutivo ya que incluye todos los eucariotas que no son animales, plantas u hongos. La primera observación de protistas registrada fue la de Leeuwenhoek en el año 1674, pero el término fue acuñado y popularizado en 1866 por Haeckel (famoso por sus ilustraciones detalladas de estos organismos, Figura 1). Al principio este término comprendía también los organismos procariotas.

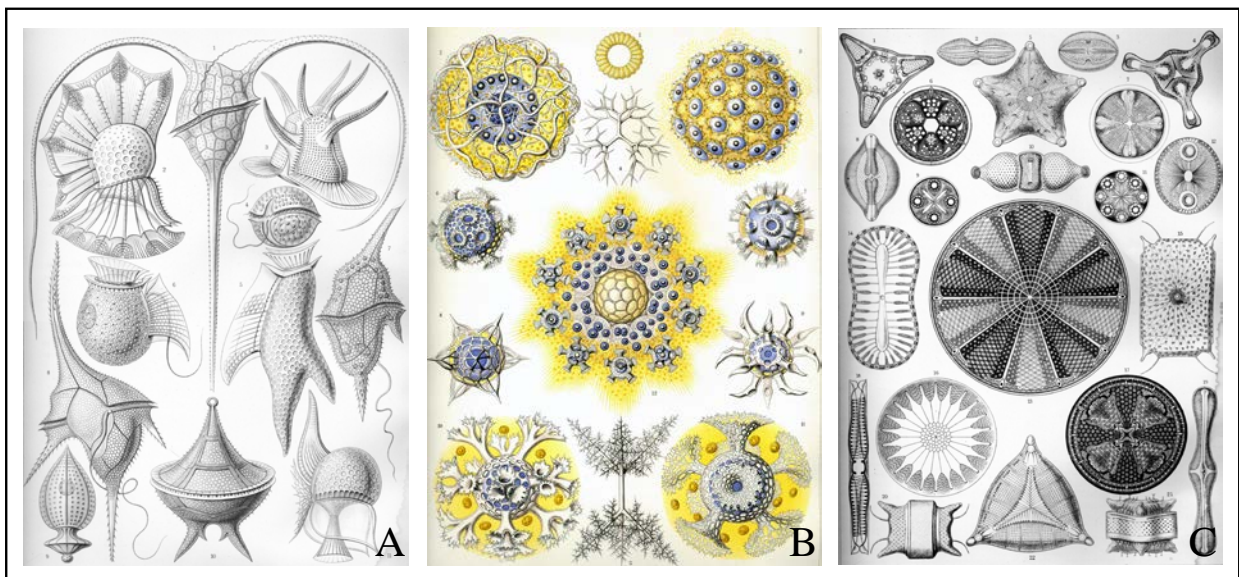


Figura 1. La diversidad morfológica en Alveolata (A), Rhizaria (B) y Stramenopiles (C). Dibujos de Ernst Haeckel, 1904.

Protista fue entonces considerado como un nuevo reino, de la misma manera que *Animalia* y *Plantae*, aparentemente menos importante y con menos categorías. Hoy en día, gracias a numerosos estudios que empezaron en la segunda mitad del siglo XX, sabemos que animales y plantas son dos pequeñas hojas en el árbol filogenético de los eucariotas, que está compuesto sobretodo por formas de vida unicelulares (Figura 2).

El progreso tecnológico ha hecho posible el paso del estudio de lo visible hacia el descubrimiento

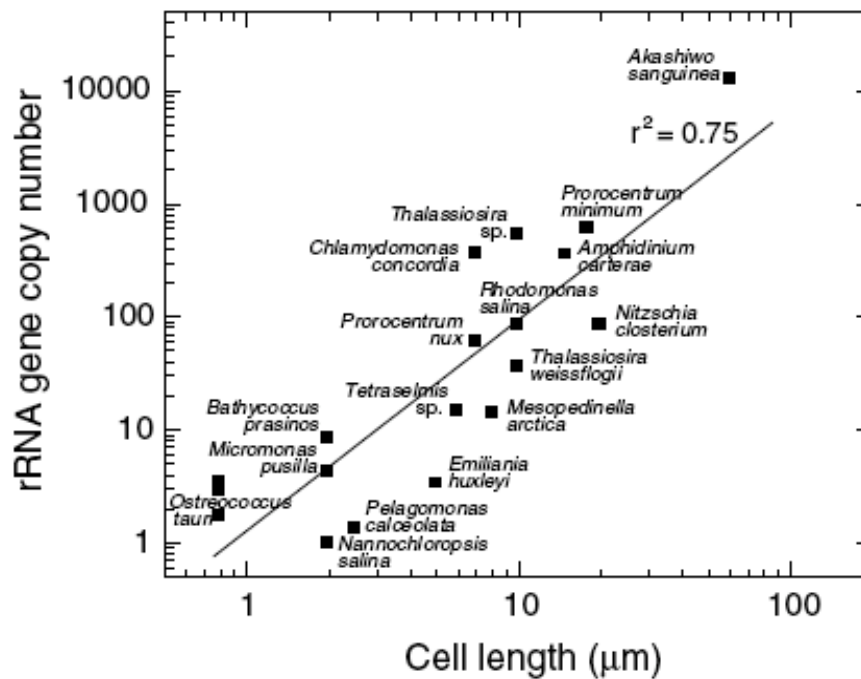


Figura 3. Correlación entre el tamaño de la célula y el número de copias de ADNr en diferentes especies de protistas. Tomado de Zhu et al. (2005).

ficar la abundancia microbiana es la citometría de flujo (Zubkov *et al.* 2006 and 2007; Christaki *et al.*, 2011). Esta técnica es muy útil en caso de numerosas muestras, pero aún necesita ser mejorada y optimizada por algún grupo funcional como el de los microeucariotas heterotróficos. A pesar de la posibilidad de identificar alguna topología celular, la microscopia de epifluorescencia no es bastante precisa para el estudio de la diversidad. De hecho, los caracteres morfológicos son inútiles para identificar células de tamaño *pico*-, también para la mayoría de tamaño *nano*- y a menudo no permite investigar en profundidad la taxonomía a nivel de especie.

Con el desarrollo de los métodos moleculares el estudio de la diversidad microbiana aumentó exponencialmente. Estudios pioneros en esta dirección fueron los de Díez *et al.*, 2001, López-García *et al.*, 2001 y Moon-van der Staay *et al.*, 2001. Estas investigaciones tenían que enfrentarse todas con un único problema general: para estimar la diversidad es fundamental identificar grupos de

Reference	Species	Sim	Origin
Rooney <i>et al.</i> (2004)	<i>Cryptosporidium parvum</i>	92.1	Genome
	<i>Plasmodium Falciparum</i>	89.5	Genome
	<i>Plasmodium berghei</i>	92.0	GeneBank
Alverson <i>et al.</i> (2005)	<i>Skeletonema grethae</i>	99.2	Strain
	<i>Skeletonema japonicum</i>	99.4	Strain
	<i>Skeletonema menzeleii</i>	99.4	Strain
	<i>Skeletonema pseudocostatum</i>	99.5	Strain
	<i>Skeletonema subsalsum</i>	99.5	Strain
Simon <i>et al.</i> (2008)	<i>Phoma exigua</i>	99.5	Strain
	<i>Mycospharella punctiformis</i>	99.6	Strain
	<i>Teratospheria microspora</i>	99.6	Strain
	<i>Davidiella tassiana</i>	99.4	Strain
	<i>Aspergillus nidulans</i>	99.6	Strain
Gong <i>et al.</i> (2013)	<i>Tintinnopsis sp.</i>	99.1	Individual
	<i>Pseudotontonia sp.</i>	99.3	Individual
	<i>Strombidium sp.</i>	99.7	Individual
	<i>Vorticella sp.</i>	99.1	Individual

Tabla 1. Variabilidad intragenómica del gen 18S ADNr (SSU). Se muestran los valores de similitud genética intragenómica en diferentes especies de microeukaryotes. Los valores bajos de similitud en *Plasmodium spp.* Y *Cryptosporidium parvum* se explica por la presencia efectiva de las diferentes formas ribosomales que se activan en diferentes huéspedes de estos parásitos.

organismos similares definidos como “especies” en la taxonomía clásica. Diferentes conceptos de especie han sido aplicados a los microorganismos en general y a los protistas en particular (Roselló-Mora and Amann 2001, Schlegel and Meisterefeld 2003). La definición más pragmática propone que una especie es “un grupo de organismos que comparten características morfológicas similares”. La definición biológica, quizás la más útil para animales y plantas, define una especie como “un grupo de organismos capaces de cruzarse sexualmente y producir una progenie fértil”. Aunque la división celular asexual es la más común entre los protistas, también hay diferentes casos de reproducción sexual (Amato *et al.*, 2007), pero existe poca información acerca de cuán distribuida se encuentre entre los diferentes clados y con qué frecuencia ocurre. Por supuesto, estudios acerca del ciclo de vida de las especies de protistas para encontrar la incidencia de la reproducción asexual o sexual son muy necesarios. La principal limitación de este tipo de estudios es que relativamente pocas especies de protistas están cultivadas y bien caracterizadas e incluso

algunos grupos carecen por completo de representantes cultivados. Por lo tanto, el uso del concepto biológico de especie no es muy práctico para estudiar la diversidad de los protistas.

Afortunadamente para los microbiólogos, en los años 70 del siglo pasado Carl Woese se dio cuenta de la posibilidad de identificar todas las formas de vida comparando sus secuencias de ADN (Woese *et al.*, 1977). Esta operación se compone básicamente de dos pasos: el alineado de las secuencias de ADN del mismo gen y la medida de sus distancias genéticas. La diana favorita de esta operación desde el principio fue el gen ribosomal del ARN (rDNA). Este gen está presente en todos los organismos y está suficientemente conservado para ser usado en una filogenia entre cualquier forma de vida. La taxonomía molecular presenta varias ventajas: se puede aplicar a una amplia gama de taxones, a todas las etapas de la vida y a grandes volúmenes de datos que son típicos de la mayoría de los estudios ecológicos (Caron *et al.* 2009). El rADN es muy útil pero no es un blanco perfecto, considerando que es un gen multicopia, particularmente en los eucariotas. En cepas de algas el número de copias varía de 1 a 10.000 (Zhu *et al.* 2005), lo que implica que la abundancia relativa de genes puede desviarse considerablemente de la abundancia relativa de células. El número de copias es proporcional al tamaño celular y del genoma (Figura 3), por lo tanto la posibilidad de grandes variaciones es más baja para las células de tamaño pico- y nano-. Además, es posible que estas copias tengan una gran variabilidad a nivel intragenómico. El riesgo de la variabilidad intragenómica es que se podrían detectar dos o más secuencias diferentes donde en realidad sólo hay un organismo. Una vez más, en la mayoría de los casos, esta variabilidad intragenómica es muy baja (Tabla 1).

La importante novedad de las técnicas moleculares fue la posibilidad de un estudio más realista de la diversidad microbiana marina, particularmente a nivel del nano- y pico plancton. El método clásico fue la construcción de bibliotecas de clones del gen 18S rDNA, las secuencias se amplificaban

a partir de ADN genómico ambiental, con una reacción en cadena de la polimerasa (PCR, Mullis *et al.* 1983). Típicamente, desde una biblioteca de clones se conseguían entre 100 y 500 secuencias. Estas secuencias ribosomales se convirtieron en la base de una nueva taxonomía molecular; de hecho fue creado un nuevo concepto de especie más pragmático que el biológico o morfológico: la unidad taxonómica operativa (OTU). Siguiendo este criterio, las secuencias fueron agrupadas en unidades contables que tenían un cierto criterio de divergencia genética, elegida por el investigador, en una operación normalmente llamada “clustering” (agrupación). Es importante subrayar que la manera de agrupar las secuencias en OTU es un paso crucial para determinar nuestra visión de la diversidad en muestras marinas.

A pesar de alguna limitación, el gen rDNA (y particularmente el 18S rDNA) sigue siendo el mejor compromiso para estudiar la diversidad de protistas y ha sido elegido como diana en la emergente tecnología de secuenciación de alto rendimiento (*high-throughput sequencing*, HTS), 454 e Illumina, que ha evolucionado tan rápidamente que la definición de “secuenciación de próxima generación” ha quedado obsoleta en menos de 5 años. El número de secuencias recogidas con las HTS es de varios órdenes de magnitud mayor que las conseguidas mediante bibliotecas de clones, lo que conlleva a la aparición de nuevos problemas relacionados con la gestión de este gran número de datos. Sin embargo, hay un gran entusiasmo relacionado con la posibilidad de poder secuenciar el océano (Venter *et al.* 2004) y muchos científicos están trabajando en la optimización del método para aumentar la confianza en él (Kunin *et al.* 2009, Quince *et al.* 2009). Las nuevas tecnologías de secuenciación dan la posibilidad de estudiar la diversidad de manera más profunda. Es importante recordar que, cuando están combinadas con la PCR, las HTS sufren los mismos tipos de errores que las bibliotecas de clones (Winzingerode *et al.* 1997). Hoy en día la metagenómica, aunque tiene por objetivo principal más el estudio de las funciones metabólicas que el de la diversidad de las

especies, es una posible alternativa para recolectar secuencias de 18S rDNA desde comunidades naturales de microorganismos (Logares *et al.* 2013). El uso de las técnicas metagenómicas es independiente del paso por la técnica de PCR, eliminando así esta fuente de errores. Para un estudio profundizado de las características de la diversidad de los protistas, todos los métodos descritos han sido utilizados en esta tesis.

Taxonomía general de los microeucariotas

La taxonomía de los eucariotas es un continuo motivo de debate en el mundo científico (Burki *et al.* 2008). Gracias al uso conjunto de microscopía y biología molecular es posible identificar la mayoría de los taxones, pero el verdadero reto es entender cómo estos grupos se relacionan entre sí. En este trabajo hay una mezcla de grupos morfológicos clásicos (mejorando su definición gracias a las herramientas moleculares) y nuevos ribogrupos, que están formados por secuencias que se agrupan juntas en un árbol filogenético fuera de los grupos conocidos. Estos nuevos ribogrupos se insertan entre medio de los grupos clásicos, que son los que representan la espina dorsal del árbol de la vida eucariota y están definidos por su morfología. La taxonomía general usada por la mayoría de los grupos morfológicos está basada en Adl *et al.* (2012).

De todo el árbol de la vida (Figura 2) cuatro supergrupos de protistas merecen ser mencionados debido a su particular importancia en los estudios moleculares ambientales: Alveolata, Rhizaria, Stramenopiles y CCTH. Alveolata, a menudo el supergrupo más abundante, comprende dos de los clados más clásicos: dinoflagelados y ciliados. Rhizaria está compuesto sobretodo por radiolarios, que pueden ser solitarios o vivir en colonias y se caracterizan por una estructura compleja, Cercozoa (también conocidos como Filosa) y Foraminífera, un grupo que prefiere los sedimentos. Los estramenópilos comprenden grupos fototróficos como las diatomeas y grupos heterotróficos como los bicosoécidos. El CCTH es un grupo propuesto recientemente que incluye Cryptophyta, Centrohelioczoa, Telonemia y Haptophyta (Burki *et al.* 2009), sin embargo, estudios filogenéticos más recientes levantan dudas sobre su monofilia (Baurian *et al.* 2010, Burki *et al.* 2012). En las últimas décadas, y gracias a las investigaciones moleculares, “un grupo” ha recobrado importancia en el ecosistema oceánico, los hongos (Bass *et al.* 2007, Richards *et al.* 2012). Tradicionalmente, los Fungi han sido estudiado más por botánicos que por protistólogos. Sin embargo, considerando que

muchas especies de hongos son unicelulares, entran perfectamente en la definición de microeucariotas, objetivo de nuestro estudio. Los Fungi han sido encontrados en la columna de agua y en los sedimentos (Bass *et al.* 2007, Lepere *et al.* 2008 Jebaraj *et al.* 2009, Edgcomb *et al.* 2011, Richards *et al.* 2012). Inicialmente una práctica común era eliminar las secuencias de Fungi de las bibliotecas de clones de los eucariotas, como normalmente se hace con los Metazoa. Ahora estas secuencias son apreciadas y guardadas, de hecho no tendría sentido excluir este importante elemento del ecosistema marino.

Los ribogrupos representan una gran parte de las secuencias encontradas en los estudios moleculares. De hecho la mayoría de las secuencias de este estudio pertenecen a los alveolados marinos (MALV), que fueron ya encontrados en el primer estudio molecular del océano profundo (Lopez-Garcia *et al.* 2001) y mejorada su definición posteriormente (Groissillier *et al.* 2006). Otros ribogrupos importantes son los estramenópilos marinos (MAST), definidos en el año 2004 por Masana *et al.*, y los Picozoa (conocidos anteriormente como Picobiliphyta) que fueron identificados primero por sus secuencias ambientales (Not *et al.* 2007) y posteriormente cultivados (Seenivasan *et al.* 2013). Los Rhizaria abarcan tres ribogrupos, RAD_A, RAD-B, y RAD_C, el segundo comprende el grupo morfológico de Sticholonche conocido anteriormente como Taxopodia. Todos estos ribogrupos se encuentran frecuentemente y están ampliamente reconocidos, entrando de lleno en la taxonomía “práctica”.

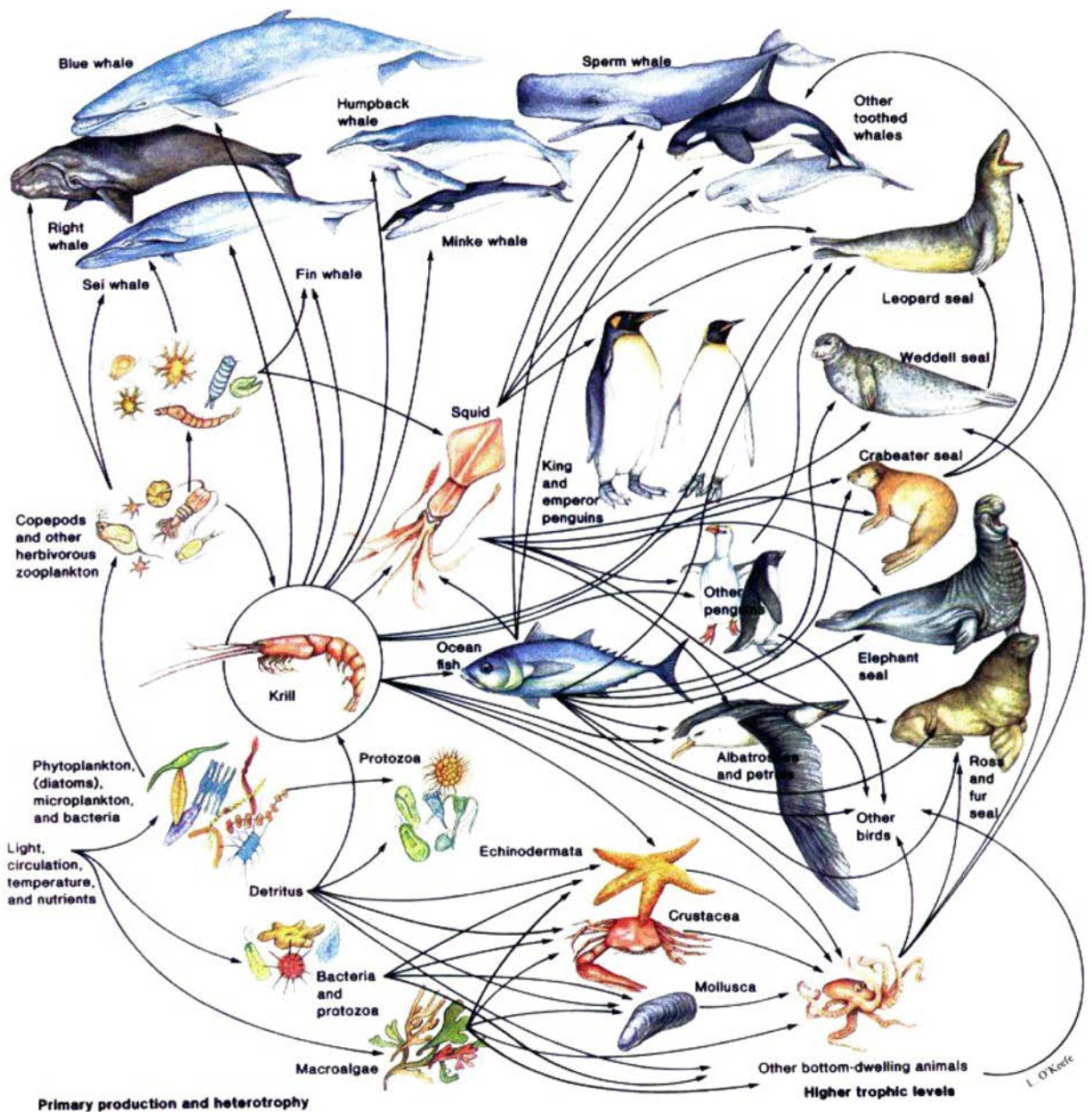


Figura 4. Una red trófica marina completa, lo que indica una gran variedad de especies y sus interacciones. La red trófica microbiana se muestra a la izquierda. Dibujo de O’Keefe L.

Papel trófico y participación en los ciclos biogeoquímicos

En un mililitro de agua epipelágica hay entre 1000 y 10.000 células de microeucariotas. Es difícil identificar la función ecológica de cada taxón pero está claro que juntos juegan un papel importante en los ciclos biogeoquímicos, tanto como autótrofos o como heterótrofos. Es importante recordar que el fitoplancton, hoy en día considerado como la suma de las bacterias fototróficas y

de los protistas fototróficos, produce el 70% del oxígeno total del planeta (Epstein *et al.* 1993), haciendo posible la vida en la tierra. Generalmente los organismos unicelulares están conectados en una compleja red trófica basada en el tamaño. El bucle microbiano, propuesto por Azam *et al.* en 1983, constituye un interesante indicio de esta red. Gracias a este bucle, la materia orgánica disuelta es consumida por los procariotas y llega a los niveles tróficos superiores debido a que los protistas fagotróficos se alimentan de procariotas y son a la vez presa de organismos zooplanctónicos más grandes (Figura 4 y Figura 5b). Los procariotas se pueden considerar como la máquina bioquímica que guía los principales ciclos biogeoquímicos (carbón, nitrógeno, azufre), pero lo que controla finalmente la velocidad de estas reacciones metabólicas son los protistas bacterívoros, probablemente junto a los virus (Boras *et al.* 2010).

La fagotrofia, la ingestión de partículas de comida mediante invaginación de la membrana celular, está muy difundida en los taxones de protistas. Es un proceso que está bastante bien estudiado tanto en el ambiente (sobre todo en agua epipelágica) como en los cultivos. Las clases importantes de fagotrofos son por ejemplo los ciliados, los cuales representan a los predadores en la clásica red trófica, las crisófitas y el ribogruppo MAST-4, quizás el bacterívoro más importante en el bucle microbiano marino (Massana 2011). Sin embargo, la fagotrofia no es la única forma de heterotrofia en el ambiente. Diferentes especies que pertenecen a los Fungi (Richards *et al.* 2012), Excavata (von Der Heyden *et al.* 2004), Chrysophyceae (Sandgren *et al.* 1995, Sanders *et al.* 2001) o Labyrinthulidae (Raghukumar *et al.* 2001) sobreviven gracias a la osmotrofia, que es la asimilación de compuestos orgánicos disueltos mediante ósmosis. Además, hay varios ejemplos de grupos que sobreviven en el océano gracias a interacciones parasitarias y abarcan un ancho abanico de organismos anfitriones. Estos parásitos marinos incluyen el ribotipo MALV-I y -II, que es el grupo más grande en término de secuencias encontradas (Siano *et al.* 2010). El estudio ambiental

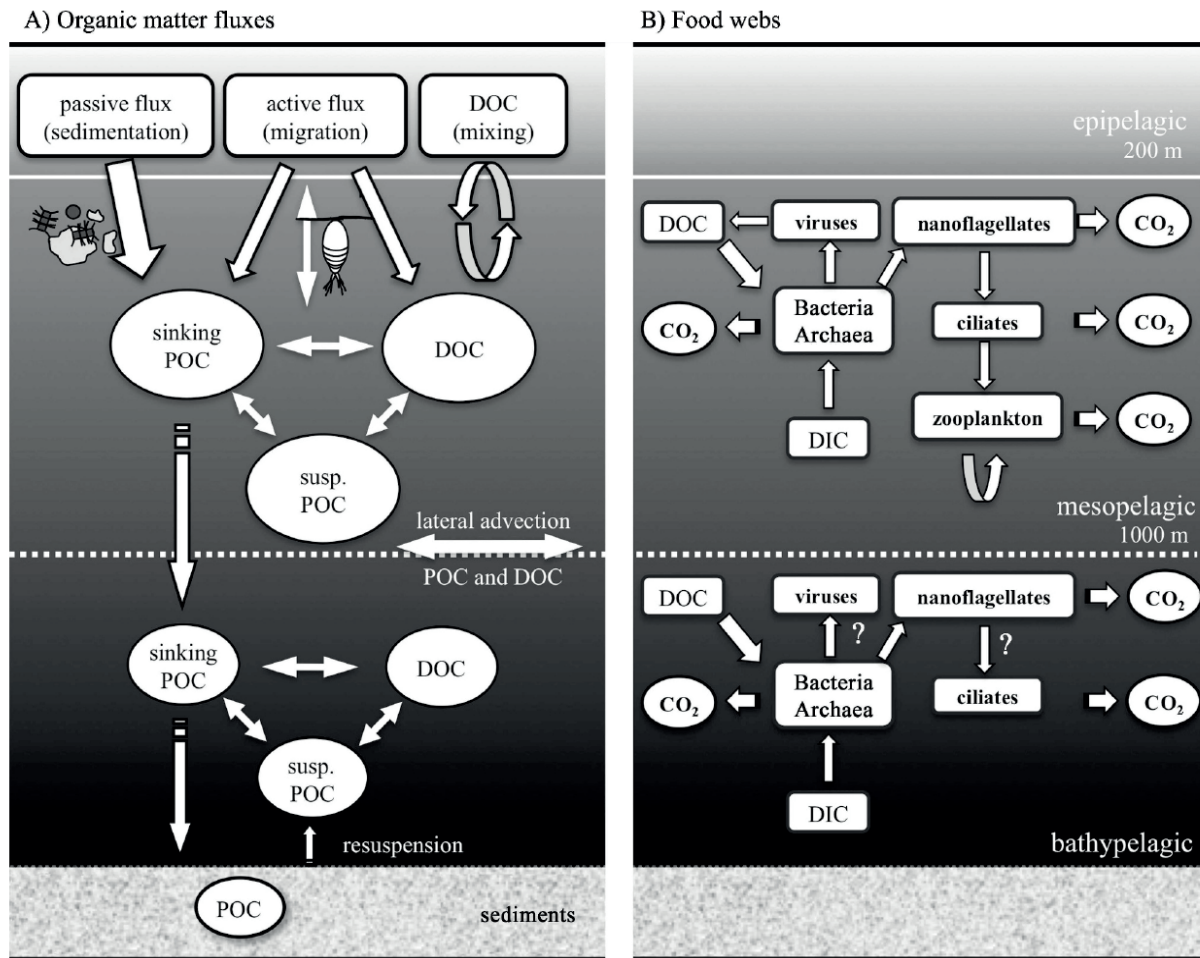


Figura 5. Representación esquemática de los flujos de materia orgánica (A) y la red trófica microbiana (B) en el ecosistema oceánico profundo (de Aristegui et al. 2009). A) Se indican tres conjuntos de carbono orgánico interconectados: carbono orgánico disuelto (DOC), carbono orgánico particulado que se hunde (POC) y POC suspendidos. B) Red trófica microbiana de los reinos mesopelágicos y batipelágicos. Los procariontes en el océano oscuro pueden sobrevivir gracias al DOC (heterotrofia) y carbono inorgánico (quimiosíntesis). En la zona batipelágica el control de los procariontes por los flagelados o los virus y el papel de los ciliados sigue siendo enigmático (signos de interrogación).

del parasitismo es bastante difícil y a pesar de los pocos casos documentados (Chambouvet *et al.* 2008), la magnitud de este fenómeno, como por ejemplo la osmotrofia, no es clara. Fuera de estas estrictas diferencias tróficas, hay que recordar que el mundo unicelular favorece una cierta plasticidad en el método trófico, y la mixotrofia, la combinación de diferentes estilos, parece ser una conducta común en los grupos de protistas (Sanders *et al.* 1991; Jones 2000, Zubkov *et al.* 2008).

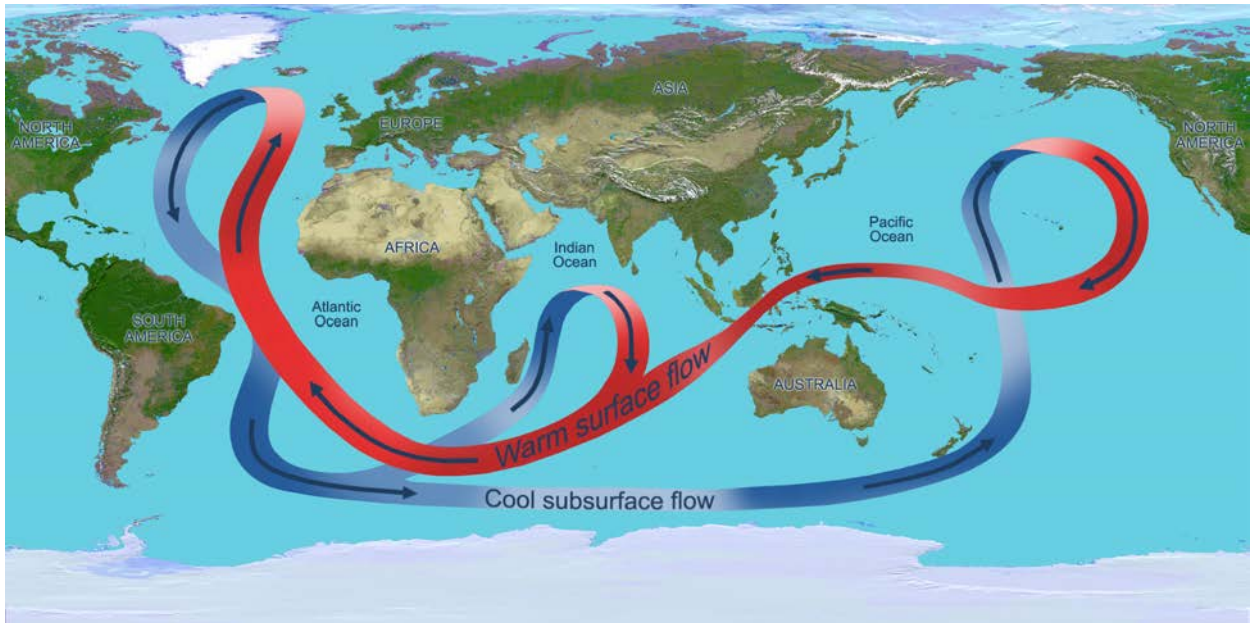
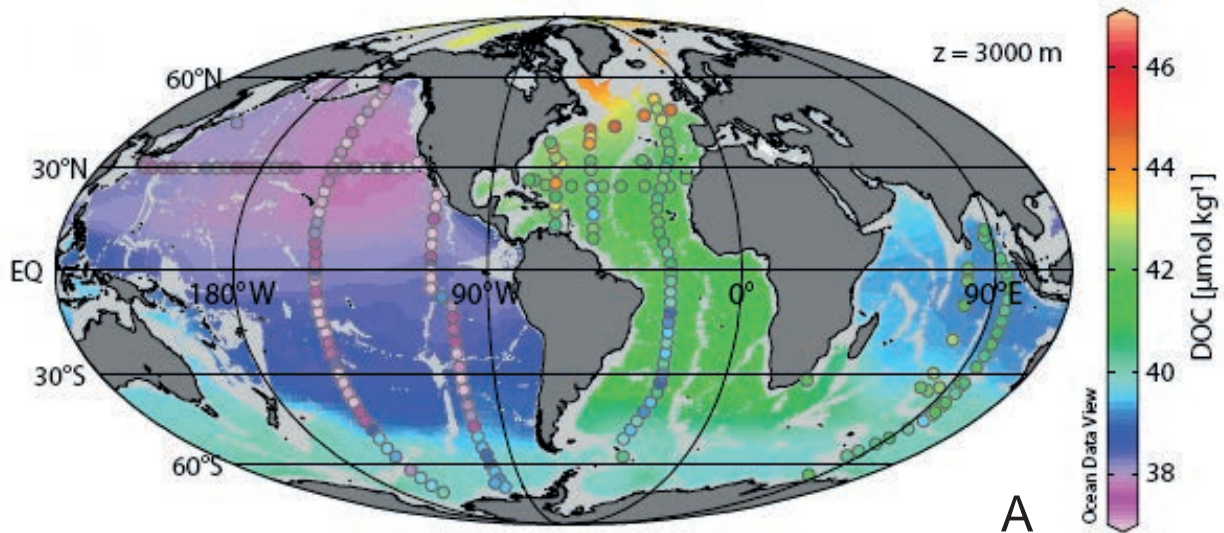


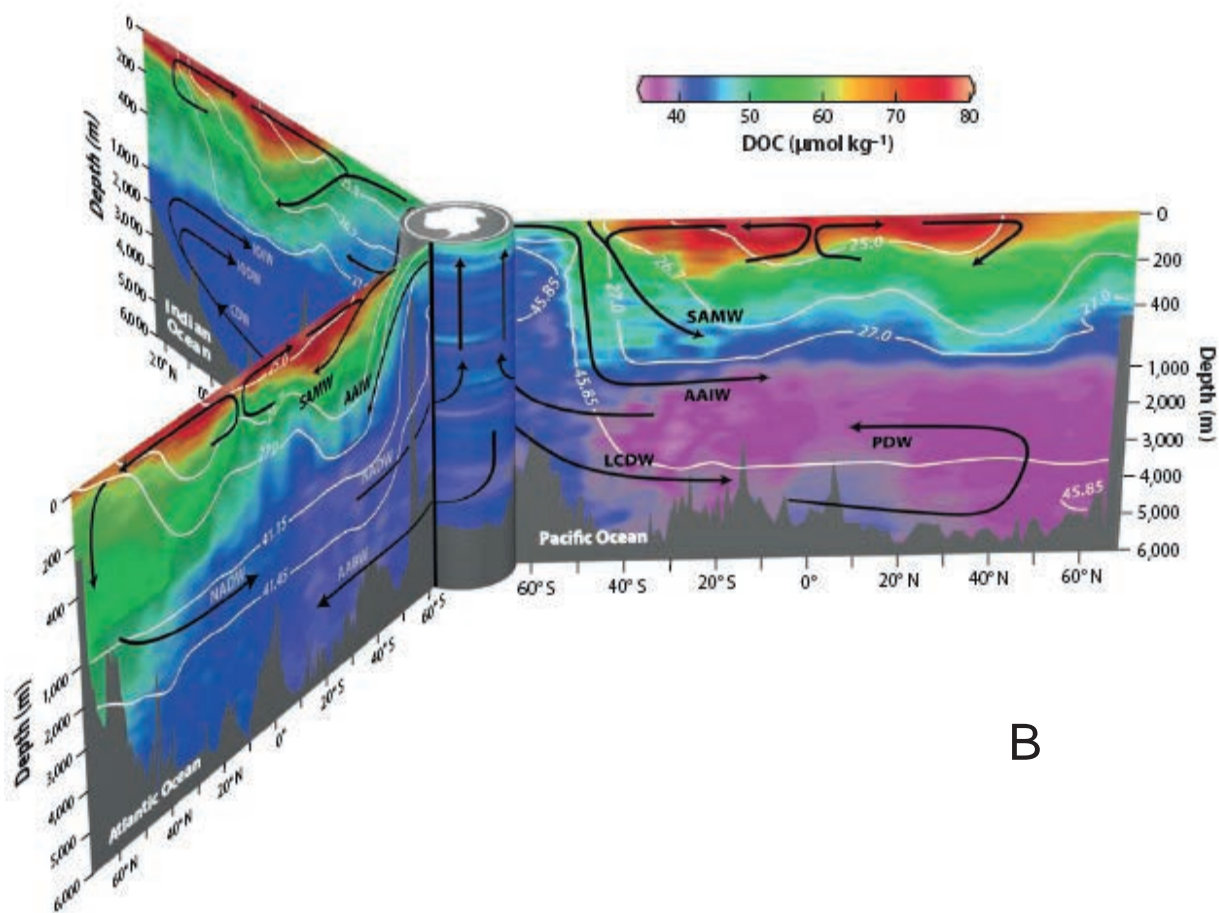
Figura 6. La circulación termohalina. Masas de aguas frías y densas se hunden en el Atlántico Norte y en los océanos del sur, creando una corriente que fluye en las cuencas oceánicas. Estas aguas vuelven a la superficie gracias al influjo de los afloramientos en los océanos Índico y Pacífico Norte, formando una corriente de agua cálida que fluye en la dirección opuesta a las capas superiores. Cerca del Polo Norte las aguas se enfrían y se hunden volviendo a reiniciar un ciclo que dura siglos.

El océano profundo: un hábitat peculiar

La capacidad de adaptación de los microeucariotas es evidente con el hecho de que están presentes en todo el planeta, incluyendo todos los tipos de ambientes extremos. En el océano, sabemos que están presentes en toda la columna de agua (Not *et al.* 2007). Sin embargo, por razones obvias, las comunidades superficiales se estudian más que las profundas. De hecho, el funcionamiento del océano profundo todavía está alejado de esclarecerse completamente. Tradicionalmente el océano profundo está dividido en tres zonas, la mesopelágica (200-1000 m), la batipelágica (1000-4000 m) y la abisopelágica (más de 4000 m). La región mesopelágica, donde a menudo se encuentra la termoclina, está más influenciada por los aportes epipelágicos (0-200 m) que los dos estratos más profundos. De hecho, una gran parte del carbono orgánico fijado por la fotosíntesis es respirado en esta zona (Aristegui *et al.* 2005).



A



B

Figura 7. Distribución de carbono orgánico disuelto (COD $\mu\text{mol kg}^{-1}$) en los océanos del mundo (Hansell et al 2009). A) Distribución del COD a 3000 m. Los puntos son los valores observados, mientras que los colores del fondo provienen de un modelo. B) Distribución del COD en el Atlántico central, Pacífico central y en el Índico oriental. Las flechas representan la circulación de masas de agua.

La zona batipelágica muestra varias diferencias en comparación con los ecosistemas superiores. Teniendo en cuenta los parámetros físicos, este sistema es más estable: el agua está generalmente bien oxigenada (aunque existan zonas anóxicas), la temperatura presenta un rango muy estrecho a nivel mundial, de 1 a 4 °C, y la salinidad es prácticamente constante alrededor de 35 ppm. En la región batipelágica, la presión es muy alta (5 a 10 MPa), pero esto no limita el desarrollo de la vida a escala macro y micro. A pesar de esta aparente homogeneidad, todavía es posible reconocer diferentes masas de agua en base a los parámetros físicos y químicos, siendo las más importantes la del agua profunda del Atlántico Norte (NADW), la del agua profunda circumpolar (CDW) y la del agua profunda del mar de Weddel (WSDW). Estos tres tipos de agua se encuentran respectivamente en el Atlántico, Pacífico e Índico.

Desde un punto de vista químico, la concentración de la materia orgánica, de los nutrientes inorgánicos y de otros compuestos químicos puede ser muy variada en diferentes regiones marinas, dependiendo del material que se hunde desde la superficie. Generalmente, el océano batipelágico es rico en las formas oxidadas de nutrientes inorgánicos (NO_3 , PO_4) y carece de compuestos reducidos tales como el amonio (Nagata *et al.* 2010). A nivel mundial las profundidades del océano son consideradas como la mayor reserva de carbono orgánico biodisponible (Libes 1992, Benner 2002), sin embargo, la concentración de carbono orgánico disuelto (COD) difiere entre las diferentes cuencas (Hansell y Carlson 1998). El papel de las profundidades del océano como sumidero de carbono inorgánico es bastante intuitivo (Figura 5a). Entre el 5 y el 15% del carbono fijado por la fotosíntesis en capas marinas superiores llegan al reino batipelágico a través de la bomba biológica (Giering *et al.* 2014), donde es respirado y secuestrado durante siglos hasta que regresa a la parte superior del océano y luego al ambiente. Por lo tanto, el sistema batipelágico tiene un papel muy importante en el balance global de CO_2 y, teniendo en cuenta su vínculo con problemas críticos,

tales como el calentamiento global y el cambio climático, es muy importante definir el destino final del COD del agua profunda (Aristegui *et al.* 2009).

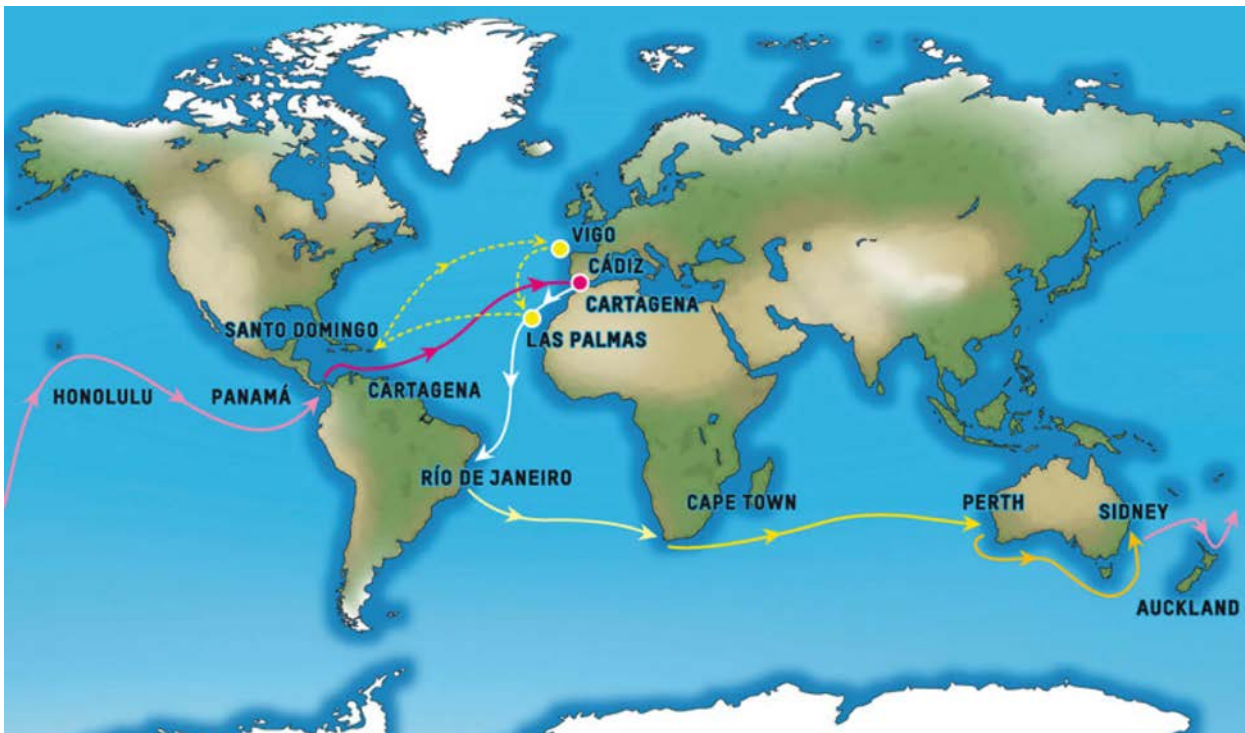


Figura 8. Itinerario de la expedición Malaspina 2010, incluyendo las rutas de la nave Hespérides (línea continua) y del Sarmiento de Gamboa (línea punteada) .

A nivel batipelágico hay una disminución general de la concentración de COD a lo largo de la circulación termohalina profunda (Figura 6). La concentración de COD es alta ($> 50 \mu\text{mol kg}^{-1} \text{C}$) en aguas del Atlántico norte recién formadas (al norte de 50°N), tiende a ser un poco más baja y constante en las regiones ecuatoriales (aproximadamente $45 \mu\text{mol kg}^{-1} \text{C}$) y desciende nuevamente en el sur a un mínimo de $39 \mu\text{mol kg}^{-1} \text{C}$. La concentración constante de COD a lo largo del sur del Océano Índico ($40 \mu\text{mol kg}^{-1} \text{C}$) sugiere una entrada neta de carbono, debida a la invasión de agua profunda circumpolar (CDW), y una eliminación posterior. La concentración de carbono orgánico en aguas profundas del océano Pacífico va disminuyendo gradualmente a medida que estas avanzan hacia el norte: el COD es de $42 \mu\text{mol kg}^{-1} \text{C}$ en las aguas circumpolares del Pacífico sur

y disminuye a $36 \mu\text{mol kg}^{-1} \text{ C}$ (Figura 7, Hansell *et al.* 2009). Probablemente esta disminución del COD se explica por el consumo biológico debido a procariontes heterotróficos y tal vez a los hongos .

Las características abióticas del océano profundo definen un hábitat muy diferente de la superficie. La información sobre la abundancia y la distribución de los microeucariotas en la columna de agua del océano oscuro es fragmentaria (Tanaka y Rassoulzadegan 2002, Yamaguchi *et al.* 2004, Fukuda *et al.* 2007, Sohrin *et al.* 2010, Morgan -Smith *et al.* 2011, Morgan - Smith *et al.* 2013). Se han realizado estudios de diversidad en la columna de agua (López-García *et al.* 2001, Stoeck *et al.* 2003, Countway *et al.* 2007, Not *et al.* 2007), en los sedimentos (Edgcomb *et al.* 2011a, Salani *et al.* 2012) y en las chimeneas marinas (Edgcomb *et al.* 2002, Sauvadet *et al.* 2010). A partir de estos estudios sabemos que la diversidad de los protistas en aguas profundas aparece dominada por Alveolata y Radiolaria mientras que los hongos predominan en los sedimentos. A pesar de no ser uno de los grupos dominantes, Excavata aparentemente prefiere las aguas profundas más que las superficiales. Siendo la fototrofia imposible en la oscuridad del océano profundo, la comunidad batipelágica de los protistas sobrevive gracias a uno de los tres estilos de vida heterótrofos mencionados anteriormente: fagotrofia, osmotrofia y parasitismo. La importancia relativa de cada modalidad trófica en el ecosistema es todavía tema de debate en la comunidad científica.

Para lograr una visión global del funcionamiento de las profundidades del océano, se realizó la campaña oceanográfica *Malaspina* en el año 2010, a bordo del BIO_Hespérides. La campaña se inició en Cádiz (España) y se tomaron muestras de 147 estaciones distribuidas por todo el mundo (Figura 8). El objetivo principal de esta expedición era el estudio del océano profundo a escala mundial, incluyendo la recopilación de datos sobre los microeucariotas. La magnitud y la multidisciplinariedad del esfuerzo de muestreo nos ha permitido comparar los datos con otros paráme-

tros, abiótico y biótico, con el fin de lograr una visión más completa de todo el sistema.

Objetivo de la tesis

El objetivo general de esta tesis es conseguir una visión global de la comunidad de microeucariotas marinos. El logro de este objetivo se estructura en cuatro capítulos. El primer capítulo (*Sequence diversity and novelty of natural assemblages of picoeukaryotes from the Indian Ocean*, ISME 2011), trata sobre el estudio de la diversidad de la comunidad epipelágica mediante bibliotecas de clones, ha sido útil como una primera aproximación a las herramientas de la biología molecular y para establecer unas directrices sobre cómo tratar los conjuntos de datos de secuencias (la alineación, la agrupación, las estimaciones de diversidad). En el segundo capítulo (*General patterns of diversity in major marine microeukaryote lineages*, PLOS ONE 2013) las secuencias derivadas de todos los trabajos publicados antes del año 2010 fueron analizadas con el fin de describir las diferentes características de la diversidad genética de los grupos microeucarióticos. Además, se realizó un estudio exploratorio del modelo evolutivo de los diferentes taxones. El fruto más precioso de este trabajo fue un conjunto de secuencias fiables, todas pertenecientes a la región V4 de 18S rADN, que fueron el núcleo de una base de datos de referencia (MAS9013) utilizada para la identificación taxonómica y la búsqueda de quimeras en los estudios sucesivos realizados por pirosecuenciación. La segunda parte de la tesis, en el marco del proyecto Malaspina-2010, se ha centrado en el ecosistema del océano profundo. El tercer capítulo (*Global abundance of planktonic heterotrophic protists in the deep ocean*, enviado a ISME J.) investiga la abundancia de flagelados heterotróficos, en las regiones globales meso y batipelágicas, con el uso combinado de la microscopía de epifluorescencia y la citometría de flujo. En el cuarto capítulo (*Diversity of marine microeukaryotes in the global deep ocean*, in preparación), se estudió la diversidad filogenética y biogeográfica de los microeucariotas, y su relación con los parámetros ambientales en la frontera entre las regiones batipelágica y abisales, a través de la pirosecuenciación de rADN y de

la metagenómica .

Los diferentes temas estudiados se pueden explicar en virtud de dos objetivos generales y varios específicos:

Objetivo 1 : *Definición de los grupos taxonómicos de los microeucariotas marinos y su estructura genética*

La primera parte de la tesis representa un esfuerzo por definir nuestra “unidad de la diversidad” a partir de los estudios basados en la clonación molecular y la secuenciación de Sanger, con el fin de establecer una base sólida para la segunda parte de la tesis. Empezamos con los datos de una campaña oceanográfica (Capítulo 1) y luego continuamos con el análisis de la base de datos de 18S rADN completa disponible en ese momento (Capítulo 2). Los objetivos específicos de esta parte fueron:

- Seleccionar la región del gen 18S rADN que mejor representa la variabilidad del gen completo
- Identificar un umbral de similitud razonable para el agrupado por OTU
- Establecer la distancia máxima en grupos taxonómicos a nivel de clase
- Destacar las clases taxonómicas típicas que forman las comunidades de superficie

Objetivo 2: *Estudio descriptivo de las comunidades del océano profundo global*

La expedición Malaspina nos permitió tener un amplio conjunto de muestras procedentes de todas partes del mundo con los parámetros abióticos y bióticos asociados. Una cantidad tan grande de

datos fue la base para el estudio de los microeucariotas de profundidad (Capítulos 3 y 4) siguiendo los siguientes objetivos específicos:

- Determinar la abundancia, biomasa y distribución de los microeucariotas en la columna de agua entre 200 y 4000 m de profundidad
- Estudiar la diversidad de microeucariotas batipelágicos mediante pirosecuenciación y metagenómica
- Identificar los parámetros abióticos y bióticos que explican la abundancia y la diversidad de los microeucariotas de profundidad, con un énfasis particular en la relación con los procariotas

Capítulo 1

A pesar de la importancia de los protistas marinos que se encuentran en la base de la cadena trófica marina, su diversidad filogenética ha sido poco estudiada, especialmente la de las células más pequeñas que son difícilmente distinguibles por sus características morfológicas. Los avances recientes obtenidos mediante la aplicación de técnicas moleculares en el estudio de la ecología de los protistas han revelado que los ensamblajes de protistas marinos están representados por grupos filogenéticos distantes e incluyen muchos taxones que son nuevos para la ciencia y han sido ignorados y han pasado desapercibidos hasta hace muy poco. En este trabajo se utilizan bibliotecas de clones del Océano Índico publicadas recientemente, con un total de 500 secuencias de 18S ADNr con alrededor de 800 pb, para descubrir el número de diferentes linajes filogenéticos y el número de OTUs (Unidades taxonómicas operativas) observados y estimados. Además ha sido indagada la novedad del conjunto de secuencias de datos. Este análisis cuantifica la magnitud de la diversidad, a diferentes escalas filogenéticas y la novedad de la señal molecular obtenida desde el plancton microbiano marino. El alto nivel de diversidad y novedad detectada tiene claras implicaciones evolutivas y ecológicas.

Capítulo 2

Los microeucariotas tienen un papel vital para el funcionamiento de los ecosistemas marinos, pero aún así algunas características generales de su diversidad y filogenia siguen sin estar claras. En este trabajo se investigaron dos aspectos de los principales linajes de microeucariotas oceánicos utilizando secuencias de 18S ADNr (las regiones hipervariables V4-V5) provenientes de bases de datos públicas y que derivan de diferentes estudios ambientales marinos. Se generó así un conjunto de datos manualmente curados de 8291 secuencias de Sanger y posteriormente se dividieron en 65 grupos taxonómicos (más o menos al nivel de clase, basándose en KeyDNATools) antes de los análisis. En primer lugar , se calcularon las distancias genéticas y el número de secuencias agrupadas en unidades taxonómicas operativas (OTUs) utilizando diferentes niveles de umbral de distancia. Se encontró que la mayoría de los grupos taxonómicos tenían una distancia genética máxima de 0.25. En segundo lugar, se utilizaron los árboles filogenéticos para estudiar los patrones evolutivos generales. Estos árboles confirmaron nuestra clasificación taxonómica y sirvieron para determinar la evolución de los linajes a través del tiempo (LTT). Los resultados de LTT indicaron diferentes dinámicas de cladogénesis entre los grupos, con algunos mostrando una diversificación temprana en su historia evolutiva y otros una más reciente. En general, nuestro estudio proporciona una descripción mejorada de la diversidad de microeucariotas en los océanos en términos de diferenciación genética dentro de los grupos, así como en la estructura filogenética general. Estos resultados serán importantes para interpretar la gran cantidad de datos de las secuencias que se van generando actualmente por las tecnologías de secuenciación de alto rendimiento.

Capítulo 3

El océano oscuro es uno de los más grandes biomas de la Tierra, con un papel crítico en la remineralización de la materia orgánica y en el secuestro de carbono a nivel mundial. A pesar de su reconocida importancia, poco se sabe acerca de la comunidad de los protistas heterótrofos (HP), que son probablemente los principales consumidores de biomasa procariota. Para investigar este componente microbiano a escala global, a fin de determinar su abundancia y biomasa en aguas meso y batipelágicas en las muestras de la circunnavegación Malaspina-2010, se ha usado una combinación de la microscopía de epifluorescencia y la citometría de flujo. Los HP eran claramente omnipresentes en el océano profundo global, incluso en las muestras más profundas investigadas (4000 m). Su abundancia disminuía con la profundidad, de un promedio de 72 ± 19 células mL^{-1} en aguas mesopelágicas a 11 ± 1 células mL^{-1} en aguas batipelágicas, mientras que su biomasa mundial disminuyó de 280 ± 46 pg C mL^{-1} a 50 ± 14 pg C mL^{-1} . Los parámetros que mejor explican la varianza de la abundancia de HP son la profundidad y la abundancia de procariotas, y en menor medida el oxígeno y los virus de genoma grande. Diferentes señales sugirieron la depredación activa por parte de HP en procariotas, tales como la presencia de flagelos en la mayoría de las células y la generalmente buena correlación con la abundancia procariota. El ratio entre la abundancia de procariotas y la de HP variaba a escala regional y los sitios con valores mayores aparecen relacionados con una mayor contribución de organismos osmotrofos en la comunidad. Nuestro estudio permite una mejor comprensión de la relación entre HP y su entorno, arrojando luz sobre su importancia como actores en la red trófica microbiana del océano oscuro .

Capítulo 4

El objetivo de este trabajo es estudiar la diversidad de los microeucariotas batipelágicos. Las muestras de agua de mar (desde 3000 hasta 4000 m de profundidad) procedían de 27 estaciones de la expedición global Malaspina-2010 incluyendo el Atlántico, Pacífico e Índico. Se utilizó la pirosecuenciación para obtener más de medio millón de secuencias de la región V4 del 18S ADN_r y después de varias etapas de curación (eliminación de las secuencias de baja calidad, más cortas de 250 bp, presentes en una sola muestra y quimeras) se agruparon en 2482 OTU al 97% de similitud. La abundancia relativa de las 20 OTU más abundantes coincide adecuadamente con los resultados de un análisis paralelo de metagenómica, lo que sugiere que la producción de secuencias se vio poco afectada por los sesgos de la técnica de la PCR. Apareció una tendencia débil de similitud genética entre estaciones geográficamente cercanas y entre muestras de la misma masa de agua. Además, el ratio de abundancia de células entre procariotas y microeucariotas tenía una relación significativa con la composición taxonómica. A pesar de que 42 OTUs se encontraron en todas las muestras, no se encontró una comunidad global típica. En cambio, aparecieron cuatro grupos filogenéticos principales (Collodaria, Crisófitas, MALV-II y Basidiomycota) presentes en diferentes proporciones en cada lugar. La cantidad de novedad filogenética se concentra en tres puntos, uno en cada océano y representa el 6 % de las secuencias globales. Las curvas de rarefacción señalan que aún hay especies por descubrir. Nuestro estudio es el primer paso esencial hacia una investigación más detallada de la microbiota del océano profundo.

Síntesis de los resultados y discusión general

El primer objetivo de esta tesis es mejorar la información sobre la estructura genética de los grupos de microeucariotas, principalmente a nivel de clase, con el objetivo final de construir una base de datos de secuencias depuradas y fiables. Una vez definida esta referencia, procedente sólo de los estudios moleculares realizados mediante secuenciación de Sanger, fue utilizada como marco para estudiar la diversidad global de los microeucariotas en las profundidades del océano usando secuenciación de alto rendimiento. La diversidad de microeucariotas de profundidad , más los datos recogidos de su abundancia y biomasa , permitió hacer un dibujo refinado del medio ambiente batipelágico .

El reto de la taxonomía molecular: herramientas para definir un grupo

La región V4 del 18S rDNA : el mejor compromiso

La gestión del gran número de secuencias provenientes de los estudios moleculares necesita un gran esfuerzo inicial para establecer buenos criterios de trabajo. Las secuencias deben ser organizadas en categorías (OTU) definidas por determinados niveles de similitud con el fin de obtener valores cuantitativos de diversidad. El primer paso de nuestro trabajo fue definir la región diana del 18S ADN_r, ya que actualmente la tecnología de secuenciación no está lista para analizar el gen completo. En particular, la elección era entre las regiones V4 y V9. Este debate nació dentro de “BioMarKs”, un proyecto europeo de investigación que estudia la diversidad de protistas con técnicas moleculares y microscópicas . Estudios seminales que utilizaban pirosecuenciación 454 se centraron en la región V9 (Amaral-Zettler *et al.* 2009, Cheung *et al.* 2009, Stoeck *et al.* 2009) que es una región muy corta (cerca de 150 pb) y fue óptima para la tecnología de secuenciación

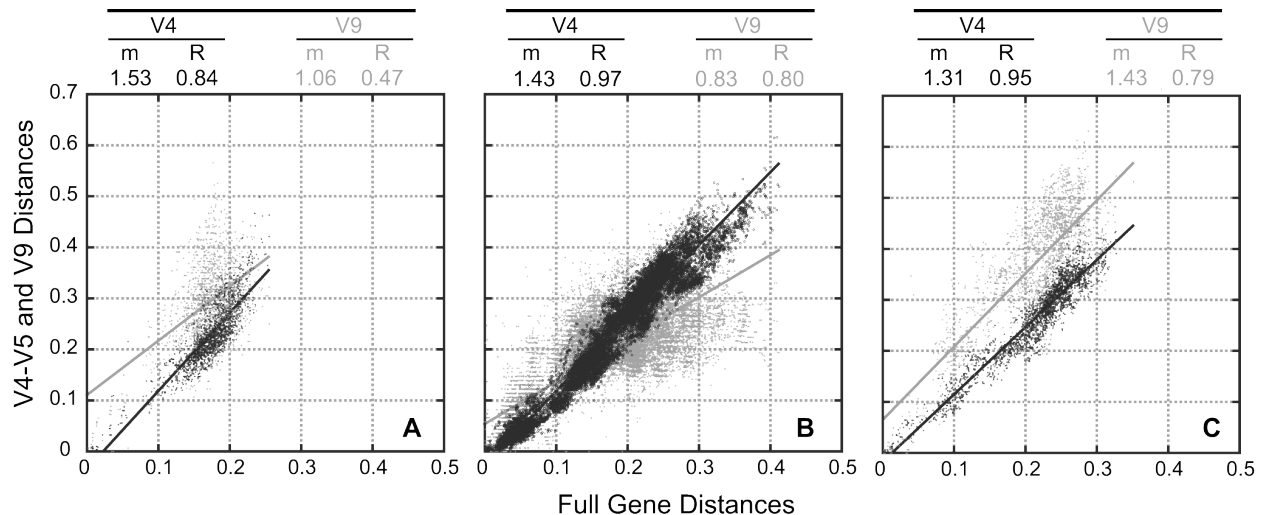


Figura 1, Capítulo 2. Comparación de secuencias parciales y completas de 18S ADNr para inferir distancias genéticas. Los tres paneles muestran las distancias genéticas (Jukes Cantor corregida) del gen completo en relación a zonas parciales (V4-V5 en gris oscuro o V9 en gris claro) para secuencias que pertenecen a Stramenopiles (A), Alveolata (B), y Rhizaria (C). Las pendientes (m) y los coeficientes (R) de las correlaciones se muestran en la parte superior de los gráficos.

en ese momento. Gracias a los avances técnicos se produjeron secuencias más largas (actualmente hasta 400 pb) y fue entonces posible analizar regiones hipervariables, como la V4. Hemos contribuido a este debate mediante la búsqueda de la región que mejor representa el gen entero. Nuestros resultados sugieren que la variabilidad detectada en la región V4 (alrededor de 500 pb) es un buen indicador de la variabilidad que se observa en el análisis de todo el gen. Las pendientes de las líneas de regresión (Figura 1, Capítulo 2) entre las distancias calculadas con el gen y la región V4 eran alrededor de 1.4 para tres supergrupos diferentes, por lo que las distancias calculadas con la región V4 se podrían traducir a distancias para todo el gen dividiendo por este valor.

Cómo definir una clase taxonómica : desde la filogenia a la agrupación

La tecnología de secuenciación mejora rápidamente, tal vez más rápido que nuestra capacidad para gestionarla. Los métodos de clonación molecular generalmente producen entre 100 y 500

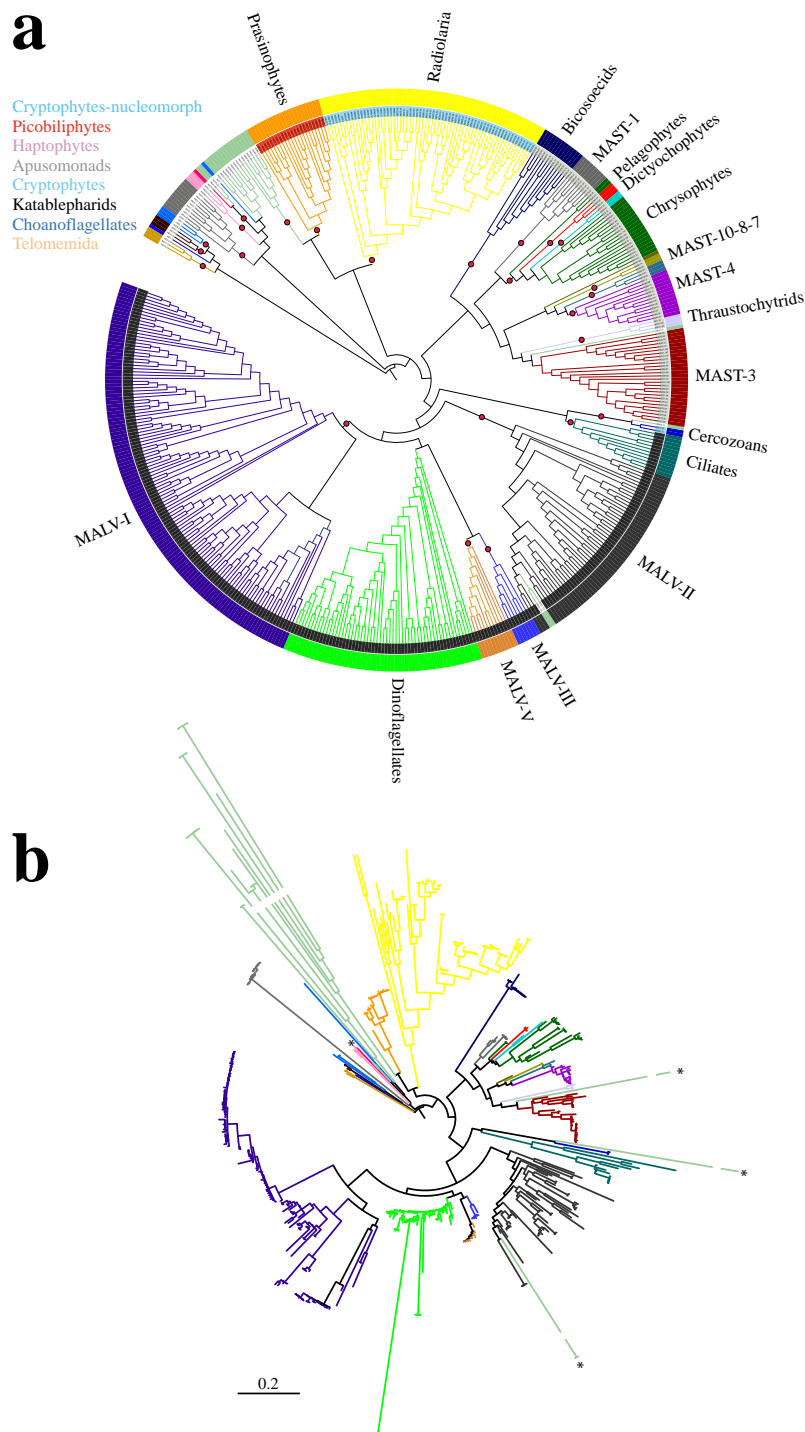


Figura 1, Capítulo 1. Árbol filogenético de secuencias de 18S ADNr de picoeucariotas muestreadas en el océano Índico. El árbol fue construido con 500 secuencias de ~ 815 bases en 961 posiciones. (a) Los colores del árbol representan el supergrupo eucariota (anillo interior) y el grupo taxonómico determinado (ramas de colores y el anillo exterior; se muestran los nombres). Las ramas que conducen a los grupos con valores de “bootstrap” por encima del 70% están marcados con un punto rojo. (b) El mismo árbol que muestra las longitudes de rama, con los mismos colores que el anterior. La longitud de las nuevas OTU (verde claro) se ha reducido a la mitad y las que se han colocado filogenéticamente están marcadas con un asterisco. La barra de escala indica 0,2 sustituciones por posición.

secuencias por muestra. A pesar del hecho de que este número es muy bajo en comparación con el rendimiento de la pirosecuenciación, el proceso global está más controlado y las secuencias obtenidas son más largas. Por lo tanto, las secuencias producidas con bibliotecas de clones son muy útiles para construir una alineación fiable, que es la base para una filogenia correcta, como hemos hecho en el primer capítulo (Figura 1). Durante la «era de las bibliotecas de clones» la construcción de un árbol era la mejor manera de definir un grupo taxonómico, pero hoy en día la superproducción de secuencias hace esta operación más difícil, especialmente el alineado de un gran número de secuencias genéticamente distantes. Ciertamente, las secuencias cortas no pueden resolver las relaciones taxonómicas entre linajes distantes. En la «era de la pirosecuenciación» hay un cambio progresivo desde los árboles filogenéticos hacia la *agrupación* mediante la cual las secuencias se agrupan en unidades taxonómicas sobre la base de similitud de secuencias. La agrupación de las secuencias en un árbol se basa en distancias patrísticas entre las secuencias (longitudes de rama), que dependen de la cantidad y variabilidad de las secuencias consideradas y comparan cada secuencia con todas las demás. Por otra parte, la *agrupación* en OTU se basa en la similitud (o distancias genéticas corregidas tales como Jukes Cantor) entre pares de secuencias. El valor absoluto de similitud o distancia es independiente del número de secuencias consideradas ya que se realizan comparaciones sólo por pares. La comunidad científica está adoptando como rutina la agrupación de secuencias basadas en la similitud y el problema es que en este proceso se pierde la forma del árbol, por lo que ahora es más difícil de identificar valores atípicos y linajes que evolucionan rápidamente. En el primer capítulo comprobamos si el número de OTU de los dos enfoques era diferente y encontramos que hasta distancias de 0.10, la agrupación utilizando distancias JC (casi equivalentes a la similitud) o distancias patrísticas (desde el árbol filogenético) mostró buena correspondencia (Figura 2, Capítulo 1). Teniendo en cuenta que a menudo el

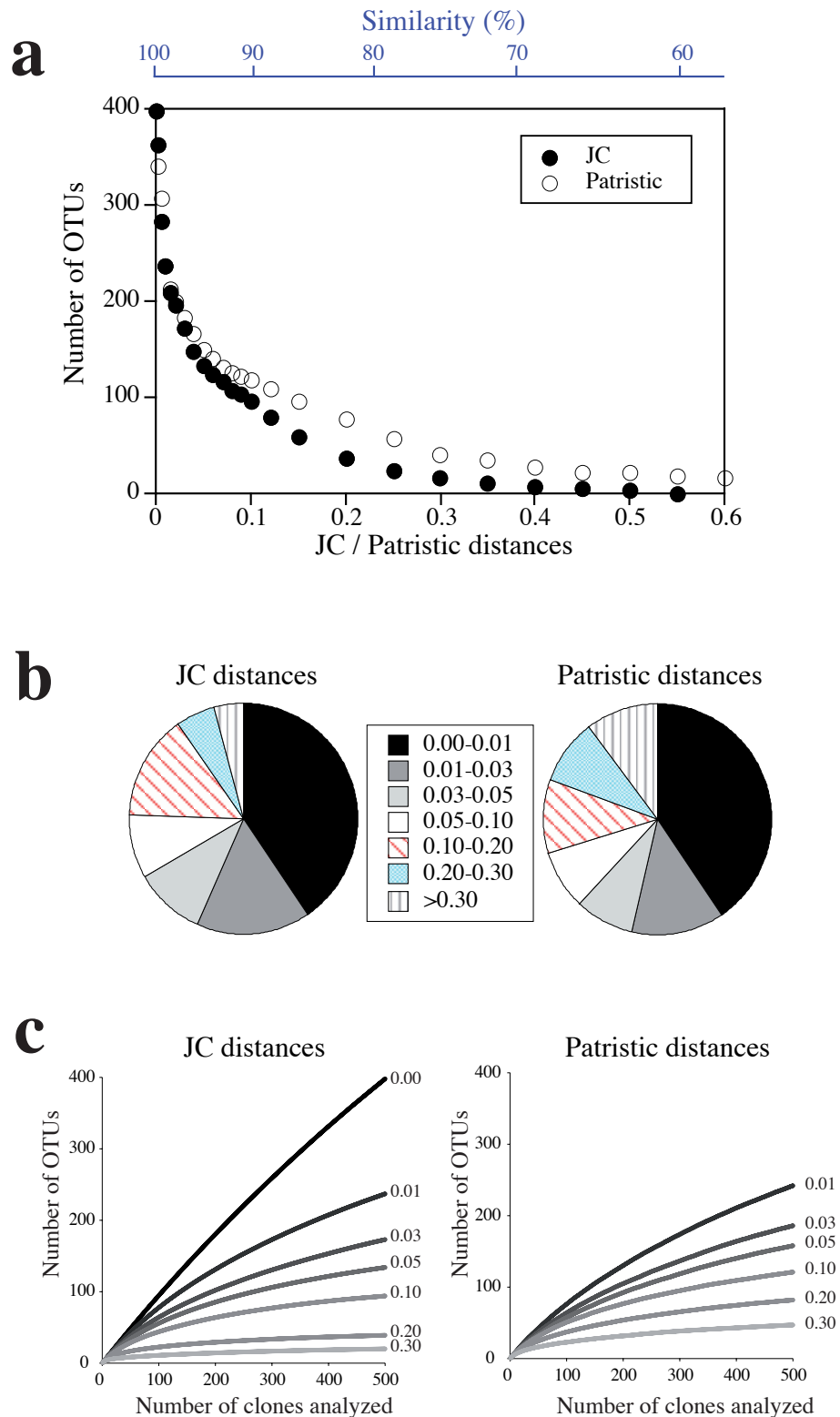


Figura 2, Capítulo 1. (a) Número de OTU observado después de agrupar las 398 secuencias únicas del Océano Índico en los diferentes niveles de agrupamiento basado en distancias Jukes-Cantor o patrística. La correspondencia entre la distancia Jc y la similitud entre secuencias se muestra en la parte superior de la gráfica para propósitos comparativos. (b) Distribución del número de OTU en clases de distancia para ambos métodos de agrupación. El área en cada clase representa la diferencia en las OTU observadas en los dos límites de la clase (así las OTU disminuyen al relajarse las condiciones de agrupamiento entre los dos límites). (c) Las curvas de rarefacción (OTU observadas versus clones analizados) a niveles discretos de distancia de agrupación (desde 0.00 a 0.30) para ambos métodos de agrupación.

agrupamiento se realiza al 95 -97 % de similitud, este se puede considerar un resultado aceptable. A pesar de que las distancias patrísticas son la base para una agrupación evolutiva más sólida y precisa, la agrupación de similitud se sigue utilizando con conjuntos de datos de 454, ya que se omite la etapa de la alineación y además proporciona una manera simple e intuitiva para analizar un número elevado de secuencias .

Definiendo los límites de los grupos taxonómicos

Cuando se trabaja con el método de la agrupación es esencial poder responder a dos preguntas: ¿cuál es el umbral de distancia que define una OTU con significado biológico útil (es decir, una especie)? y ¿cuál es la distancia máxima que se puede encontrar dentro de un grupo determinado?. Acerca de la primera pregunta, diferentes autores han propuesto diferentes niveles de umbral(Worden 2006, Jeon *et al.* 2006 , Caron *et al.* 2009), pero hay poco apoyo para justificar cada hipotético nivel y sería importante que el mundo científico llegara a una conclusión en esta dirección. Debido a los polimorfismos intragenómicos (Introducción, Tabla 1) y a los errores de secuenciación pensamos que no es aconsejable el uso del 100% de similitud para definir las OTU. Un valor entre 97-99% de similitud parece ser un criterio más razonable, ya que es suficiente para ser bastante estricto, pero no tanto como para separar secuencias que difieren a causa del polimorfismo intragenómico o por errores de secuenciación. En el capítulo dos se abordó la segunda pregunta para los grupos más o menos equivalentes a las clases taxonómicas en la sistemática clásica, incrementada por la falta de la representación gráfica del árbol que hizo necesaria una forma alternativa para identificar rápidamente las secuencias de valores atípicos. Se realizaron árboles para comprobar la afiliación de las secuencias y luego fueron analizadas por grupo. Encontramos que el 75% de los grupos a nivel de clase tenían una distancia genética máxima (corregida) por debajo de 0.25. Esta es ahora nuestra referencia general para la distancia máxima permitida dentro de una

Tabla 1, Capítulo 2. Clasificación de las secuencias de 18S ADNr ambientales en 42 grandes grupos taxonómicos. Cada grupo se codifica de acuerdo a su rango taxonómico (S: suborden; C: clase; O: orden; G: género; R: ribogrupa). La tabla muestra el número de secuencias por grupo (SEC), la media (AVG), máximo (Max) y máximo corregido (MAXC) de distancias y el número de OTU en tres niveles de corte.

* Nassellaria comprende también el orden Collodaria

Supergroup	Group		Distances				OTUs		
			Seq	Avg	Max	Max _c	0.00	0.01	0.05
Opisthokonta	Choanoflagellata	C	100	0.13	0.30	0.24	89	56	32
Rhizaria	Acantharea	C	129	0.15	0.29	0.26	110	63	29
	Chlorarachniophyceae	C	33	0.14	0.24	0.23	29	13	7
	Larcopele	O	18	0.02	0.05	-	13	4	1
	Monadofilosa	S	81	0.11	0.30	0.22	72	56	33
	Nassellaria*	O	52	0.18	0.41	0.32	45	29	19
	RAD A	R	37	0.17	0.29	0.26	34	23	15
	RAD B	R	88	0.11	0.23	0.16	66	36	17
	Spumellaria	O	209	0.06	0.26	0.13	154	79	20
Archaeplastida	Prasinophyceae	C	551	0.09	0.31	0.21	376	130	30
	Trebouxiophyceae	C	89	0.01	0.12	0.04	26	11	6
Stramenopiles	Bacillariophyceae	C	253	0.14	0.30	0.29	207	120	57
	Bicosoecia	C	75	0.11	0.35	0.28	60	34	17
	Bolidophyceae	C	63	0.05	0.12	0.11	34	12	7
	Chrysophyceae	C	152	0.13	0.27	0.24	115	75	32
	Dictyochophyceae	C	91	0.09	0.22	0.16	65	35	16
	Eustigmatophyceae	C	15	0.01	0.03	-	11	3	1
	Labyrinthulida	C	29	0.17	0.35	0.34	26	19	17
	MAST-1	R	107	0.08	0.20	0.16	74	28	9
	MAST-2	R	20	0.01	0.05	-	13	6	2
	MAST-3	R	149	0.12	0.27	0.21	110	73	31
	MAST-4	R	92	0.03	0.07	0.06	60	24	3
	MAST-7	R	82	0.04	0.14	0.08	48	21	6
	MAST-8	R	17	0.07	0.13	-	14	9	6
	MAST-12	R	26	0.16	0.27	-	24	19	16
	Oomyceta	C	19	0.11	0.29	-	16	13	10
	Pelagophyceae	C	34	0.01	0.07	0.02	22	8	2
	Pirsonids	-	47	0.03	0.09	0.08	37	26	5
CCTH	Cryptophyceae	C	179	0.09	0.24	0.21	130	45	3
	Katablepharids	-	20	0.02	0.06	-	12	6	2
	Picobiliphyceae	R	53	0.07	0.20	0.15	42	24	8
	Prymnesiophyceae	C	193	0.08	0.30	0.14	148	90	37
	Telonemia	C	68	0.05	0.12	0.11	60	42	9
Alveolata	Ciliophora	P	956	0.18	0.42	0.37	788	434	187
	Dinophyceae	C	1018	0.07	0.50	0.24	848	463	122
	MALV-I	R	980	0.19	0.48	0.42	779	431	132
	MALV-II	R	1815	0.16	0.38	0.30	1517	900	353
	MALV-III	R	79	0.05	0.15	0.11	60	38	9
	MALV-V	R	51	0.02	0.07	0.04	41	19	3
Excavata	Diplonemea	C	58	0.11	0.21	0.21	56	51	27
	Kinetoplastea	C	40	0.23	0.39	0.37	31	22	15
Incertae sedis	Apusomonadidae	C	14	0.15	0.41	-	9	6	4

clase y es un valor útil para interpretar la equivalencia taxonómica de ribogrupos ambientales.

La importancia de una base de datos de referencia

Una buena base de datos de referencia es una herramienta esencial para cualquier estudio molecular basado en secuencias cortas, pero las bases de datos actuales comprenden sólo los procariotas (Greengenes) o dan un tratamiento menos preciso a los eucariotas (SILVA). El principal problema de SILVA es que a menudo carece de una buena asignación taxonómica de las secuencias de eucariotas, especialmente de los ribogrupos recién descubiertos. Una nueva herramienta para la identificación de eucariotas a partir del 18S ADNr se ha publicado muy recientemente, la base de datos PR2 (Base de datos Protista ribosomal de referencia, Guillou *et al.* 2013). Esta herramienta no existía al comienzo de esta tesis. En este marco, quiero destacar la importancia del conjunto de secuencias depuradas del capítulo dos (8291 secuencias), que constituye el núcleo, mejorado con PR2, de una base de datos interna de referencia (MAS9013), que se ha utilizado para la asignación taxonómica y detección de quimeras en la segunda parte de la tesis y en otras publicaciones en preparación.

Composición típica de la comunidad de microeucariotas epipelágicos

La diversidad observada en el capítulo 1, en términos de abundancia relativa de determinados linajes (Figura 1), es la típica que se encuentra en otros estudios moleculares de picoeucariotas marinos (Massana y Pedrós-Alió 2008, Vaulot *et al.* 2008) y se asemeja a la abundancia de grupos en el capítulo 2 (Tabla 1). Los alveolados, principalmente MALV-I y MALV-II, dominan la comunidad y representan el 47% de los clones, seguidos por los estramenopilos (19%) y Rhizaria (13%). Los hongos no fueron considerados en los dos primeros capítulos, ya que generalmente están poco representados en el entorno epipelágico, a nivel mundial comprenden menos del 1%

de las secuencias de las bibliotecas de clones (Massana y Pedrós-Alió 2008). Las diferencias en la composición taxonómica entre epipelágico y ecosistema profundo, incluso a niveles de supergrupos, son evidentes (Figura 6, Capítulo 4) y serán analizadas en la segunda parte de esta discusión. Curiosamente, a bajas distancias (mínimo 1300 km), las muestras diferían fuertemente cuando se analizaban por bibliotecas de clones (Figura 6, Capítulo 1) y en ese momento esto fue explicado por un efecto del submuestreo. Sin embargo, a distancias comparables y con un gran esfuerzo de secuenciación, vemos que las diferencias entre las muestras profundas también están presentes (Figura 5, Capítulo 4). Esta fue la primera señal del fuerte efecto del medio ambiente en la selección de la comunidad.

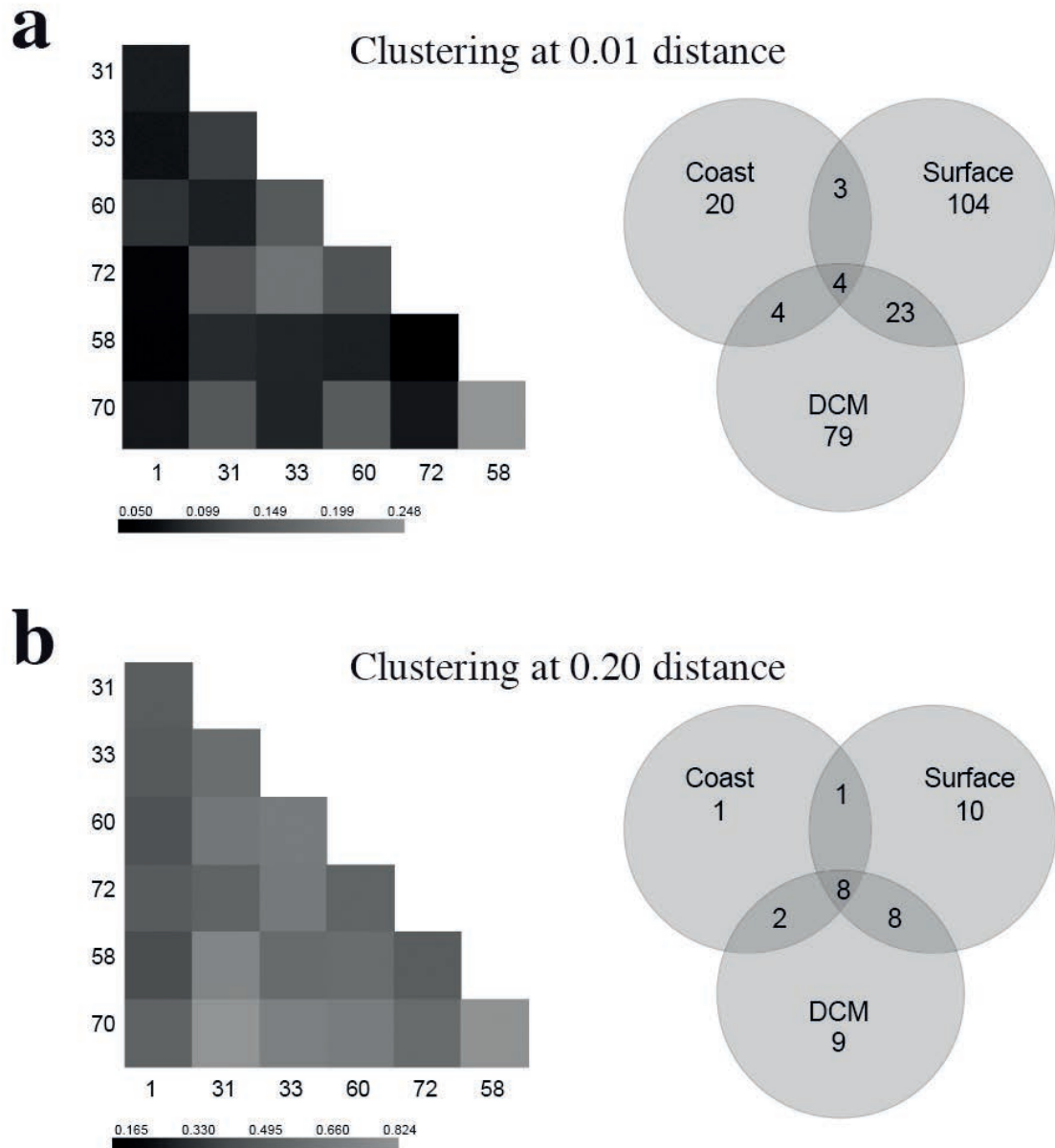


Figura 6, Capítulo 1. “Heatmaps” (izquierda) y diagramas de Venn (derecha) que comparan la diversidad de picoeucariotas marinos entre las muestras, utilizando las OTU compartidas o exclusivas definidas en la agrupación de distancias JC de 0.01 (a) o 0.20 (b). Las muestras son costeras (1), de alta mar superficiales (31, 58 y 70) o de DCM (33, 60 y 72).

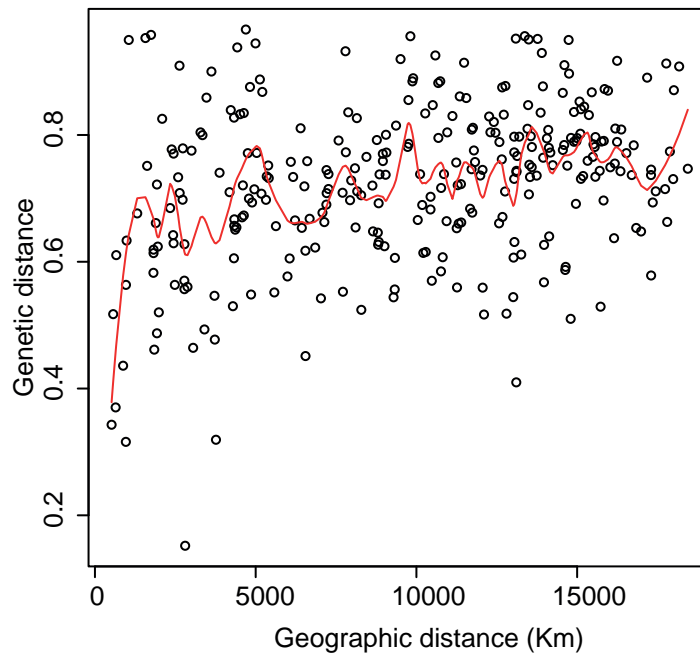


Figura 5, Capítulo 4. Test de Mantel de la relación entre las distancias genéticas (Bray-Curtis) de los conjuntos de protistas profundos y la distancia geográfica entre las muestras. También se muestra la línea de la interpolación de los valores (línea LOESS).

El océano profundo

Contando microeucariotas : la necesidad de la citometría de flujo

Por lo que sabemos este es el primer estudio que aplica la citometría de flujo, junto con la microscopía, en una investigación a gran escala (Capítulo 3). La microscopía de epifluorescencia es un trabajo que necesita mucho tiempo y es propensa a errores del operador, mientras que la citometría de flujo presenta otros tipos de problemas. Es posible identificar varias poblaciones de microeucariotas fotosintéticos gracias a sus pigmentos (Olson *et al.* 1993, Li *et al.* 1994, Marie *et al.* 2000), pero para detectar microeucariotas heterotróficos se requiere una coloración (en este caso SYBR Green). Sin embargo, esto no discrimina entre procariotas y eucariotas y puede existir una continuidad entre eucariotas de pequeño tamaño y grandes bacterias. Para resolver este problema, se utilizó la microscopía de epifluorescencia con el fin de posicionar la ventana de conteo en el software de citometría en muestras seleccionadas (a continuación, se aplicó al conjunto de datos completo) y para comprobar los valores de estaciones con abundancias poco realistas. Un método alternativo para ahorrar tiempo y recuperar el tamaño y la forma de las células, sería la microscopía de epifluorescencia automática, que se aplicará en el futuro en nuestro laboratorio. La comparación de la citometría de flujo y los recuentos microscópicos (Figura 2, Capítulo 3) fue muy buena ($R^2 = 0.82$, $p < 0.0001$). Por lo tanto, un gran número de muestras fueron procesadas por citometría de flujo. La abundancia de microeucariotas así determinados fue uno de los dos parámetros en la descripción de la comunidad global del océano profundo .

Características generales de los microeucariotas en el océano batipelágico

En la región batipelágica (1000- 4000 m), que fue también objeto del estudio la diversidad, la abundancia de microeucariotas promedio fue de 14 células mL⁻¹. Esta concentración no es cons-

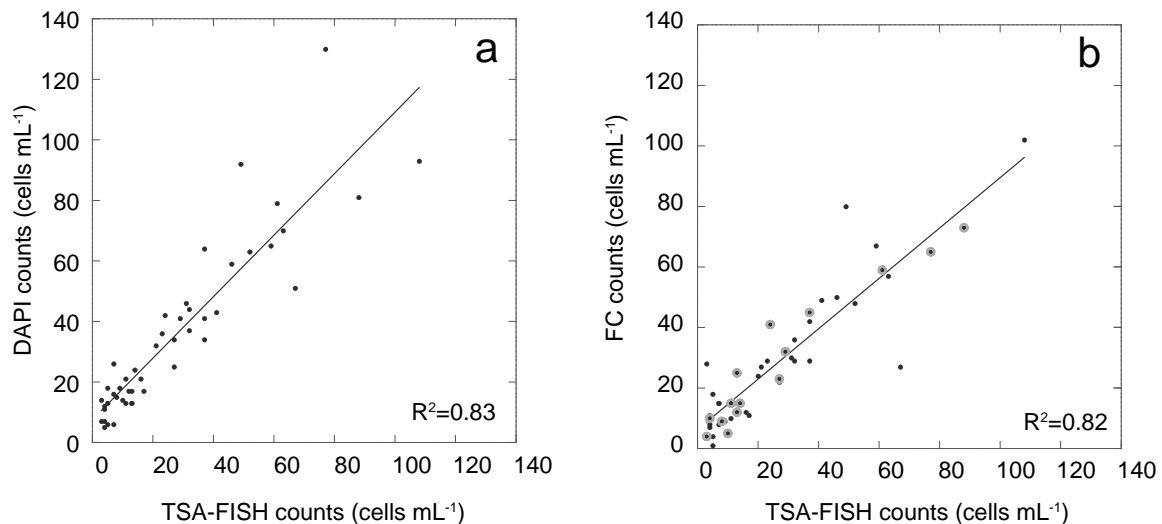


Figura 2, Capítulo 3. Comparación metodológica de recuentos de protistas heterótrofos del océano profundo. Recuentos de DAPI versus recuentos de Citometría de Flujo (FC) (a) y FC en comparación con los recuentos de TSA- FISH (b) en las muestras de diez perfiles verticales seleccionados (que se muestran como estaciones numeradas en la figura 1). Las muestras que se utilizaron para colocar la ventana FC están rodeadas por una zona gris claro en el panel B.

tante, lo cual es particularmente evidente en el Pacífico Sur, donde hay un pico de 58 células mL⁻¹ en la muestra de mayor profundidad de la estación 98. En cuanto a la estructura del tamaño celular, el porcentaje de células muy pequeñas (diámetro equivalente < 3 μm) disminuye con la profundidad (Figura 6, Capítulo 3) y a partir de las imágenes tomadas por las mediciones de biomasa, sabemos que algunas de estas células son claramente flageladas. La biomasa promedio de microeucariotas es 50 Pg C ml⁻¹ en la capa de 1400-4000 m. Para el estudio de la diversidad de las muestras más profundas (~ 4.000 m) fueron filtrados 120 L, lo que significa que hemos recogido cerca de 1.320.000 células por muestra. La mayoría de las secuencias recuperadas pertenecen a Rhizaria, seguido por Alveolata y Stramenopiles (Figura 6, Capítulo 4). A nivel local y a un rango taxonómico menor, las comunidades están dominadas fundamentalmente por 3 clases (Collodaria, Crisófitas, y Basidiomicota) y un ribogruppo (MALV -II). Las diferencias en la abundancia y la diversidad entre muestras profundas de microeucariotas están claramente relacionadas, tanto a los

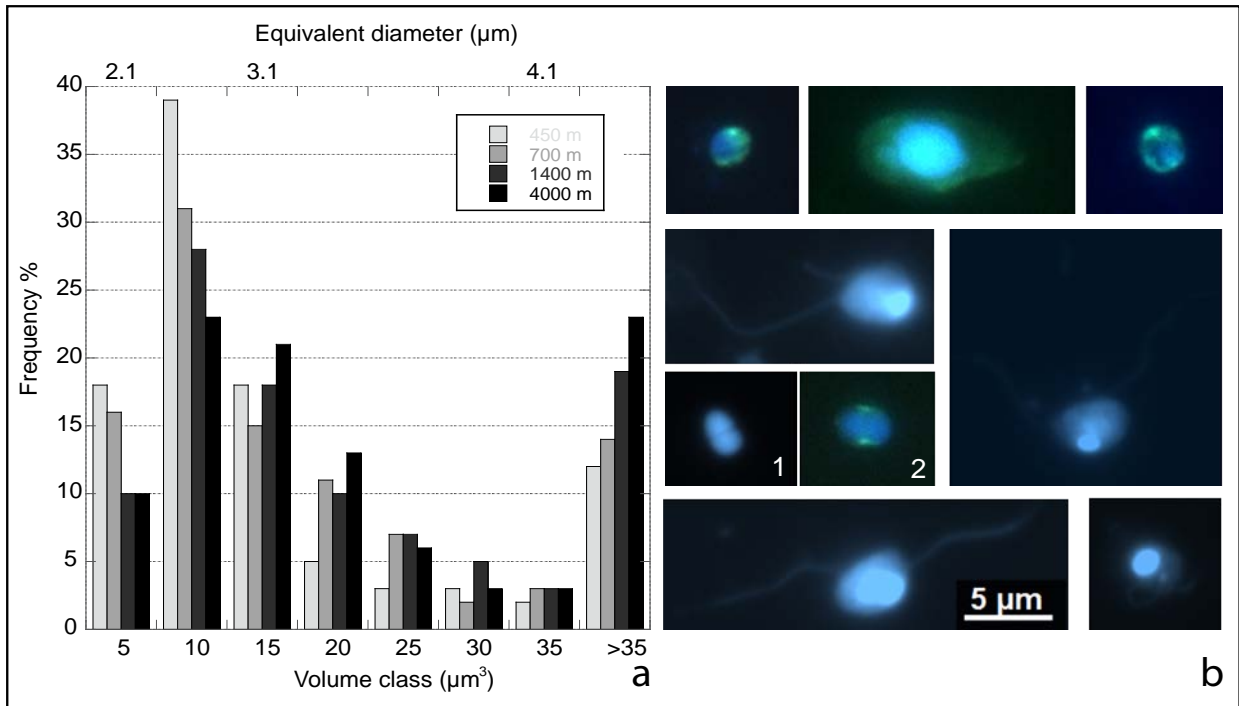


Figura 6, Capítulo 3 (a) Espectros del biovolumen celular de las células HP en diferentes capas de profundidad. También se indica para cada biovolumen celular el diámetro esférico equivalente. (b) Algunas micrografías de células HP batipelágicas, que muestran diferentes formas de células y la presencia de flagelos. La señal azul corresponde al núcleo teñido con DAPI y la señal verde al citoplasma teñido con TSA-FISH. Morfotipos parciales se muestran en las figuras 1 y 2.

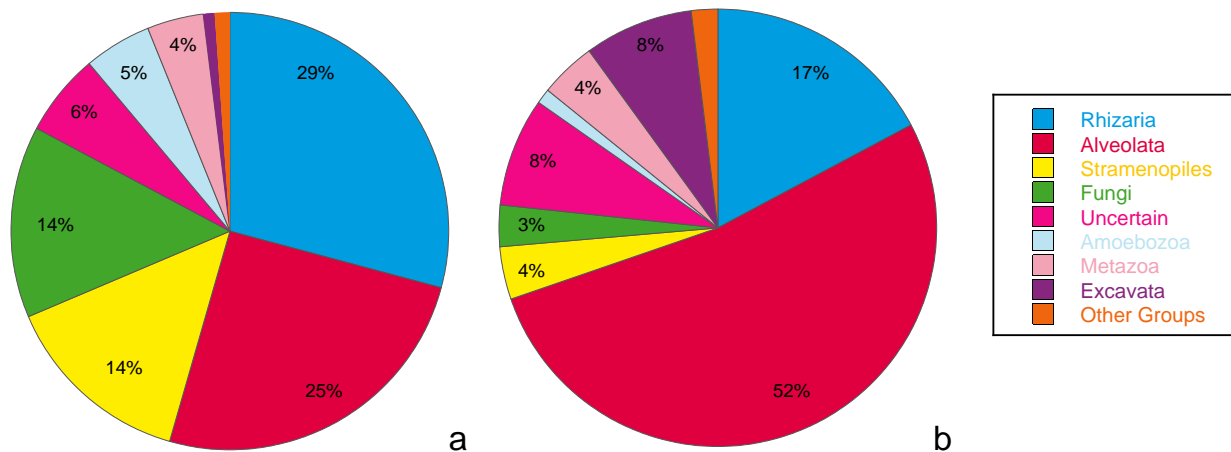


Figura 6, Capítulo 4. Cuadro general de la diversidad de los protistas de profundidad a nivel taxonómico de supergrupo. (a) Número de pyrotags por cada supergrupo, con un promedio de la abundancia relativa de cada muestra. (b) Número de OTU97 por cada supergrupo.

parámetros abióticos (oxígeno, temperatura) como bióticos (abundancia procariótica y viral).

Tabla 3, Capítulo 4. Las veinte OTU más abundantes con su número de secuencias, ocurrencia, identificación taxonómica y la similitud con la referencia medioambiental más cercana (CEM) y la correspondencia mayor con una especie cultivada (CCM).

OTU ID	Pyrotags	OCC	Group	CEM	% SI	CCM	% SI
146	45261	27	<i>Collodaria</i>	GU825331	90	<i>Collophidium ellipsoidae</i>	90
6539	29099	27	<i>Basidiomycota</i>	HQ438183	99	<i>Tilletiopsis minor</i>	98
941	23512	27	<i>Chrysophyceae</i>	JQ782092	99	<i>Pedospumella encystans</i>	98
3736	16125	24	<i>Spumellaria</i>	EF172914	99	<i>Cladococcus viminalis</i>	96
2627	13645	27	<i>Dinophyceae</i>	EU500130	100	<i>Lepidodinium chlorophorum</i>	99
309	11748	27	<i>Ascomycota</i>	GQ120160	99	<i>Engyodontium album</i>	99
1730	11131	27	<i>Amoebozoa</i>	GU320596	99	<i>Platyamoeba contorta</i>	90
2006	8958	20	<i>Uncertain</i>	JX194706	77	<i>Collozoum Serpentinum</i>	85
2275	8364	25	<i>Chrysophyceae</i>	KC306509	98	<i>Ochromonas distigma</i>	97
1165	8228	26	<i>Collodaria</i>	GU219126	99	<i>Collophidium ellipsoidae</i>	94
2418	8151	27	<i>Metazoa</i>	AY937332	99	<i>Gilia reticulada</i>	98
7646	6784	27	<i>Chrysophyceae</i>	HM749946	99	<i>Mallomonas Tonsurada</i>	92
2825	5694	27	<i>MALV-II</i>	FN598288	100	<i>Amoebophyra sp.</i>	89
6489	5597	25	<i>Uncertain</i>	GU824572	82	<i>Collozoum Serpentinum</i>	88
4675	4604	25	<i>Chrysophyceae</i>	KC306509	98	<i>Ochromonas distigma</i>	97
3936	4472	21	<i>Collodaria</i>	GU825728	96	<i>Collophidium ellipsoidae</i>	96
149	4404	20	<i>Collodaria</i>	AY046728	96	<i>Collophidium ellipsoidae</i>	84
4324	3565	27	<i>MALV-II</i>	JX194526	98	<i>Amoebophyra sp.</i>	90
5203	3552	15	<i>Collodaria</i>	GU824619	82	<i>Collozoum Serpentinum</i>	94
7437	2568	27	<i>Amoebozoa</i>	FN598227	98	<i>Platyamoeba contorta</i>	89

Protistas cultivados en aguas profundas

Sorprendentemente, el análisis de las primeras veinte OTU, que representan juntas el 50% del total de las secuencias, muestran varios ejemplos con una similitud alta con los organismos cultivados (Tabla 3, Capítulo 4). Esto no es nuevo por lo que se refiere a los hongos (Bass *et al.* 2007, Richards *et al.* 2012), pero fue inesperado para los otros grupos. Por ejemplo, 3 OTU de crisófitas que explican el 79% de las secuencias totales de esta clase son muy similares a las especies cultivadas. Sabemos que en el ambiente epipelágico normalmente existe poco acuerdo entre la diversidad detectada por estudios moleculares y la basada en cultivos (Massana *et al.* 2004), y por lo tanto considerando las características ambientales del océano profundo y su aislamiento relativo, esperábamos encontrar un consenso menor. Sin embargo, un tercio de las OTU más abundantes tienen una similitud mayor al 97% con una especie cultivada y representan el 28% de las secuencias totales. Además dos de estas OTU son en un 99% similares a especies cultivadas. Teniendo

en cuenta que la mayoría de especies cultivadas derivan de muestras epipelágicas, se puede deducir que al menos una cuarta parte de las secuencias están compartidas entre la superficie y las comunidades profundas. Esto demuestra la gran capacidad de adaptación de los microeucariotas a diferentes ambientes. Además, la observación de un impacto diferente del sesgo de cultivo entre comunidades de superficie y de profundidad es un tema que merece ser examinado más a fondo.

Ubicuidad : Todo está en todas partes?

Finlay *et al.* (2004) definió la diferencia entre la *ubicuidad* y la *dispersión ubicua* afirmando que la mayoría de los protistas están caracterizados por la segunda, lo que implica que no necesariamente se encuentran en todas partes, sino que deben estar presentes en los hábitats adecuados en todo el mundo (Caron *et al.* 2009). Otro aspecto importante de este debate es que la presencia de la misma secuencia en océanos separados nos puede dar información acerca de la ubicuidad de las respectivas especies, mientras que su ausencia puede ser simplemente debido al submuestreo. En el conjunto de datos Malaspina, 42 OTU están presentes en todas las estaciones (27) y representan el 80% de las secuencias. Esta visión cualitativa de las comunidades se acerca a la idea de que “todo está en todas partes”. Sin embargo, siguiendo el concepto de “el medio ambiente selecciona”, la distribución de estas OTU no es uniforme y una OTU que pertenece a la biosfera rara en una muestra a menudo es la OTU dominante en otra, como se ha observado en varios ejemplos, como es el caso de los hongos (Figura 7, capítulo 4). Como se ha visto en la introducción, la homogeneidad ambiental del océano profundo es una vieja y engañosa idea y, a pesar de la supuesta capacidad de dispersión alta, hay un fuerte efecto ambiental sobre la composición de la comunidad. Es probable que las propiedades intrínsecas de las masas de aguas profundas, además de la presencia de presas o alimento alternativo, permitan la existencia de diferentes nichos tróficos.

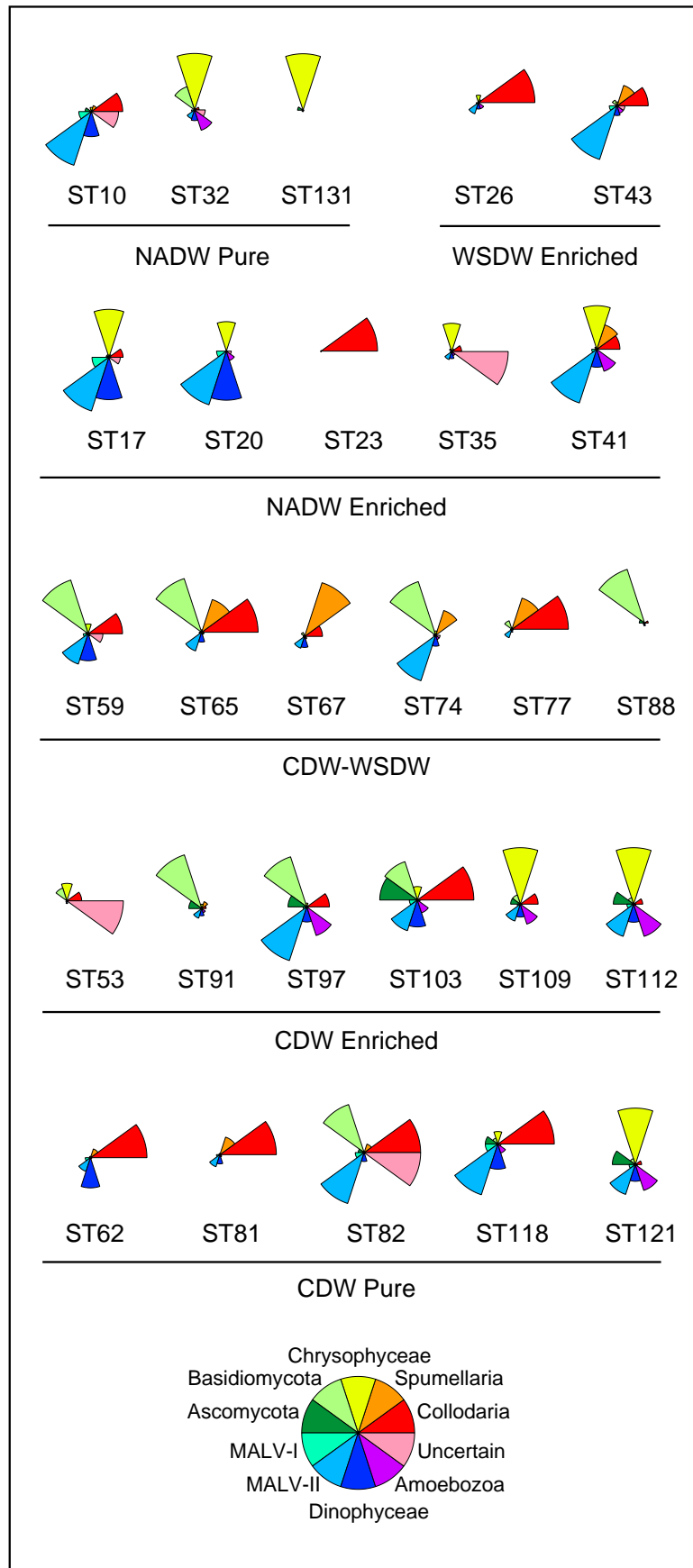


Figura 7, Capítulo 4. La abundancia relativa de los diez grupos filogenéticos más abundantes en todas las muestras profundas. Las estaciones se agrupan por su respectiva masa de agua.

Fagotrofia: la relación con los procariotas

Se esperaba que la fagotrofia fuera la vía trófica principal para los microeucariotas en el océano profundo. Para probar esta hipótesis, primero se analizó la relación entre la abundancia de procariotas y microeucariotas y después la relación entre su ratio y la diversidad. Teniendo en cuenta toda la columna de agua profunda, la abundancia de los microeucariotas se correlaciona bien con la de los procariotas ($R^2 = 0.50$, $p = 0.0001$), excepto en las estaciones del Pacífico Sur ($R^2 = 0.08$, $p > 0.05$, Figura 5, capítulo 3). Posteriormente, los análisis de regresión múltiple demostraron que este resultado era independiente de la profundidad (Capítulo 3). Sin embargo, a pesar de esta relación significativa, los procariotas explican sólo una parte de la varianza de la abundancia de

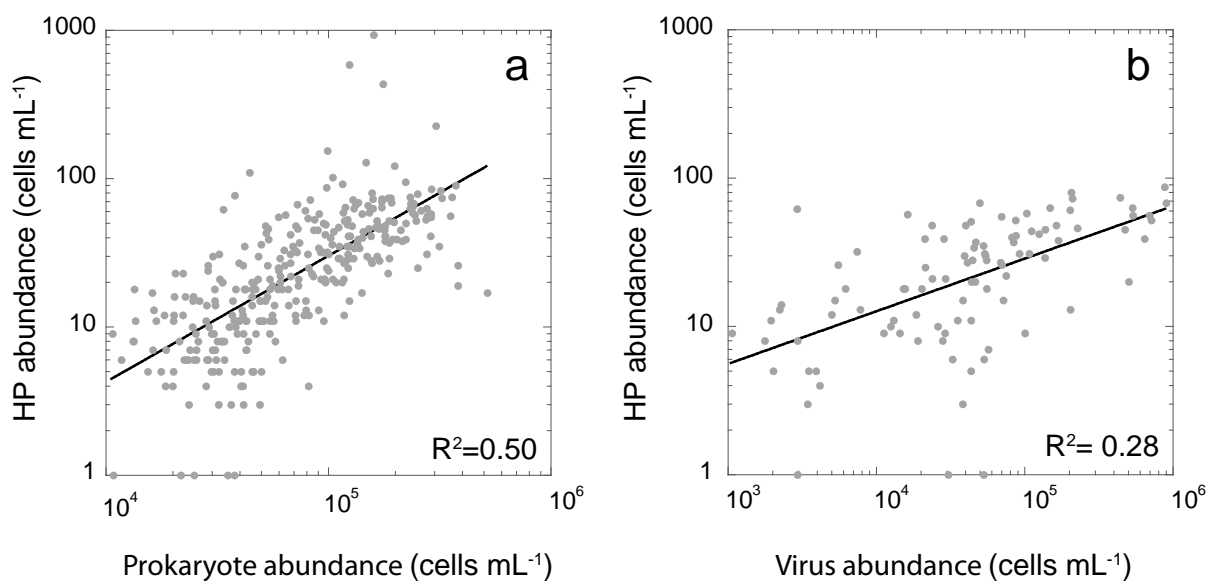


Figura 5, Capítulo 3. Abundancia de los protistas heterotróficos frente a la abundancia de los procariotas (a) y frente a la abundancia de virus grandes (b), en muestras derivadas de 71 y 20 perfiles verticales, respectivamente.

los microeucariotas. La abundancia de los procariotas ha sido utilizada también para explicar la variabilidad en la diversidad, pero el resultado no fue significativo ($p > 0.058$), aunque el 34% de la variabilidad fue explicada significativamente por el ratio entre los procariotas y los microeucariotas.

riotas ($p = 0.001$). Los valores bajos de esta proporción, similares a los de aguas epipelágicas (ca. 2000), corresponden a comunidades dominadas por Collodaria y Chrysophyceae, lo que sugiere un papel putativo de depredadores para estas dos clases. Teniendo en cuenta que menos de la mitad de la variabilidad de la abundancia y de la diversidad se explica por los procariotas y que el ratio de abundancia tiene un valor promedio más alto que en la superficie, se puede concluir que la fagotrofia en las profundidades del océano parece dar también un espacio importante a las demás vías tróficas tales como la osmotrofia y el parasitismo.

Osmotrofia: el papel de los hongos

Comparando el ratio entre procariotas y microeucariotas con la abundancia relativa de varios taxones, sólo los hongos han presentado una relación significativa (Figura 7, Capítulo 3). Teniendo en cuenta que los hongos son ciertamente osmotrofos y probablemente incapaces de realizar fagotrofia (Richards *et al.* 2012), se podría sugerir que cuando la comunidad estaba dominada por los hongos la presión de depredación sobre los procariotas era menor, favoreciendo así valores altos en el ratio de las abundancias.

Como se ve en la introducción (Figura 7b) la distribución del COD no es constante en la región batipelágica, de hecho el COD disminuye a lo largo de la cinta transportadora profunda, resultando más concentrado en aguas atlánticas que pacíficas. Es posible asociar parte de la disminución del COD desde el Océano Antártico hacia el Pacífico Norte a la presencia de hongos en estas aguas. Sin embargo, teniendo en cuenta que el COD se concentra más en el océano Atlántico, especialmente en el norte, es difícil entender por qué los hongos no prosperan en estas aguas (excepto en la estación 32). Hay varias explicaciones posibles:

- *Hongos versus procariotas*. Una primera hipótesis simple es que los hongos compiten con los

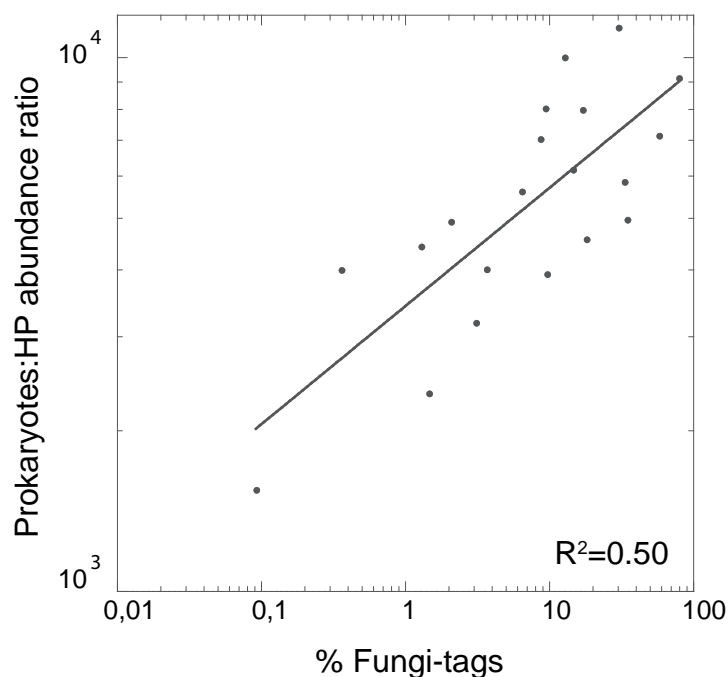


Figura 7, Capítulo 3. Relación entre el ratio de las abundancias de procariotas respecto a las células de HP y el porcentaje de secuencias de hongos en las muestras correspondientes. Los porcentajes derivan de un estudio paralelo sobre la diversidad de protistas de profundidad (Pernice et al, en preparación).

procariotas por el COD en las aguas del Atlántico. Una relación antagonista de los hongos y bacterias se observó en varios experimentos, como por ejemplo en Moller *et al.* (1999), donde los hongos y procariotas compiten claramente por el COD.

- *Hongos versus Crisófitas*. Las Crisófitas pueden ser fagotrofas (Massana 2011) y osmotrofas (Sandgren *et al.* 1995, Sanders *et al.* 2001), y pueden sobrevivir gracias a una combinación de estas dos estrategias. Salvo en la estación 32, nunca comparten el dominio con los hongos. Våge *et al.* (2013) construyeron un modelo con el fin de probar la importancia de la mixotrofia en comparación con la osmotrofia pura. Demostraron que a baja proporción de tamaño entre presas (procariotas) y depredadores (crisófitas), como sucede en las profundidades del océano, para un mixotrofo es muy conveniente la estrategia de “comerse al competidor” (Thingstad *et al.* 1996).

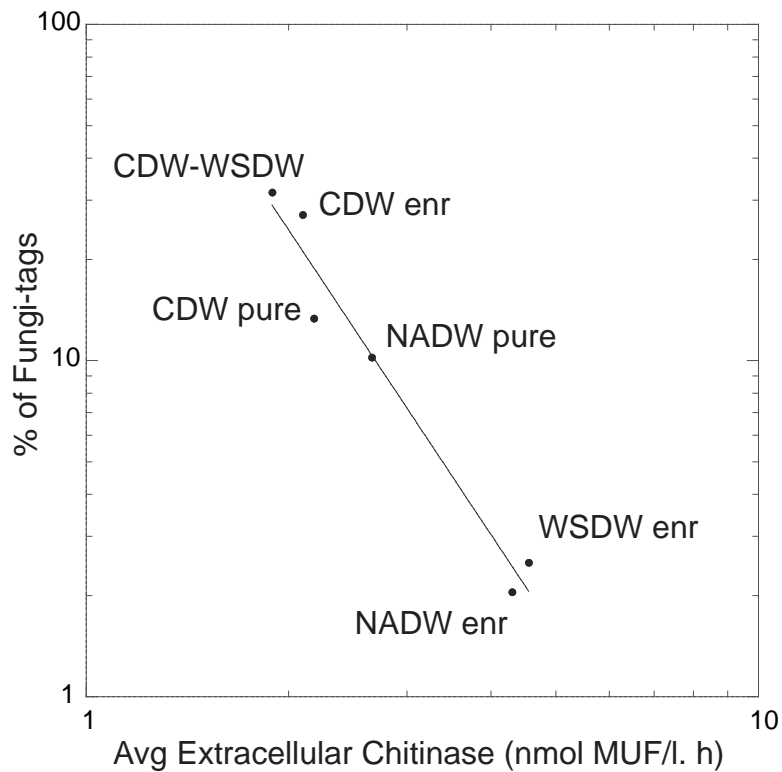


Figura 9. Relación entre la concentración de quitinasa extracelular y la abundancia media de secuencias de hongos en las diferentes masas de agua definidas.

Así que en este caso los mixotrofos (crisófitas) podrían suprimir los osmotrofos puros (hongos) depredándolos, como podría ocurrir en aguas del Atlántico y del Pacífico Norte.

- *Hongos y carbono recalcitrante.* La presencia mayor de hongos en aguas pobres en COD podría explicarse por una especialización a la asimilación de carbono recalcitrante. Un posible mecanismo es la secreción de moléculas de superóxido, en particular, formas oxidadas de Mn, que oxida carbono recalcitrante convirtiéndolo en formas más biodisponibles (Hansel *et al.* 2012). Este mecanismo está probablemente compartido con las bacterias. Por lo tanto, otra explicación de un ratio mayor entre procariontas y microeucariotas es que en una comunidad dominada por Hongos también los procariontas podrían prosperar con el COD recalcitrante.

- *Hongos frente a la quitinasa*. Una de las características típicas de los hongos es la presencia de quitina en la pared exterior (Richard *et al.* 2012). Durante la expedición Malaspina se midió la actividad enzimática extracelular presente en el agua, incluyendo la quitinasa, la enzima que digiere la quitina. La actividad de la quitinasa resulta ser mayor en el Atlántico que en el Pacífico e Índico. Teniendo en cuenta valores promedios para cada masa de agua, la abundancia relativa de las secuencias de Hongos muestran una relación inversa con la actividad de la quitinasa (Figura 9) en una correlación altamente significativa ($p=0.0005$; R^2 de 0.95). Las quitinasas podrían producirse por procariontes así como por microeucariotes (Cottrell *et al.* 2000). No está claro si los hongos son los principales objetivos de esta enzima, de todos modos el efecto de la quitinasa extracelular es un ambiente no favorable para su vida.

La osmotrofia también está presente en otros grupos taxonómicos, como los Labyrinthulidae (Raghukumar *et al.* 2001) o Excavata (Lara *et al.* 2009). De hecho, el alcance del proceso osmotrófico debería estudiarse mejor para entender el impacto de los protistas heterótrofos en el balance global de carbono.

Parasitismo: las relaciones ocultas

Inferir interacciones parasitarias desde los datos de secuenciación es bastante difícil porque no existe una correspondencia clara entre el anfitrión y la abundancia del parásito (Skovgard *et al.* 2014). Varios clados de microeucariotes marinos se consideran sobre todo parásitos, siendo el más abundante el MALV-I y el MALV-II. Numerosas especies dentro de los dinoflagelados y hongos también pueden ser parásitos. Los supuestos anfitriones de todos estos parásitos son otros microeucariotes y también la macrofauna. En nuestra base de datos, la abundancia relativa de las secuencias de MALV-II tiene una correlación significativa con la abundancia relativa de las se-

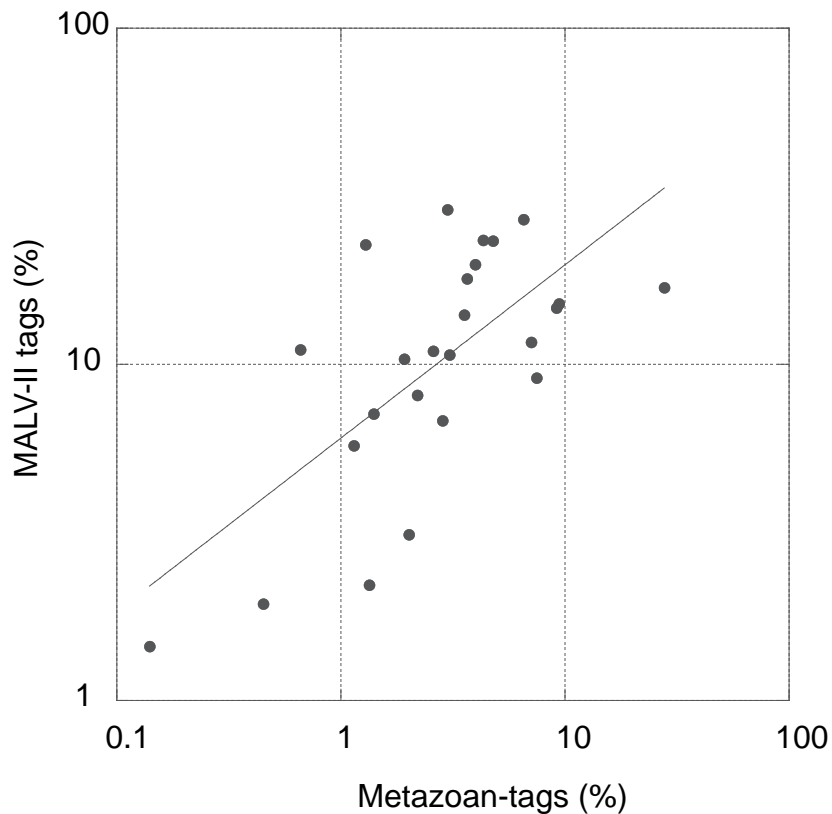


Figura 10. Relación entre la abundancia relativa de la secuencias de MALV-II frente a la de los metazoos.

cuencias de metazoos ($R^2 = 0.45$, $p = 0.0005$) (Figura 10). Esta relación es mejor para MALV-I ($R^2 = 0.60$) y menos fuerte para dinoflagelados ($R^2 = 0.41$), pero no se detectó para los hongos u otras clases. Las tres relaciones son particularmente evidentes en las muestras del Atlántico, en la que R^2 es, respectivamente, 0.85, 0.89 y 0.76. Otros candidatos putativos como parásitos del océano profundo son los hongos, especialmente considerando que la OTU principal es muy similar a una especie de parásito reconocida, *Engyodontium álbum*, que parasita *Felis domesticus* (Dennis 1995). Sin embargo, asumiendo una distribución aleatoria de los anfitriones se espera una distribución aleatoria de las secuencias de parásitos. En este sentido, la distribución mejor de secuencias de alveolados encaja con este escenario en comparación con la distribución no aleatoria de hongos.

Las últimas décadas de la investigación de los protistas se han centrado en el estudio de la diversidad de los microeucariotas, al principio definido como “inesperada” en términos de diversidad alta y novedosa. Ahora que esperamos esta increíble complejidad de la composición de los taxones, la atención se está dirigiendo hacia la investigación de la función de los diferentes taxones en el ecosistema. La asignación de un papel ecológico claro, a través del método clásico de cultivo y los nuevos enfoques genómicos de una sola célula, pronto van a mejorar la visión que tenemos ahora sobre este conjunto tan importante del medio ambiente.

Conclusiones

- 1) La región V4 del gen 18S ADNr representa mejor la variabilidad de todo el gen que la región V9. La pendiente media es de 1.4, este factor podría ser utilizado para obtener la variabilidad de todo el gen.
- 2) El valor normal de la distancia genética máxima para las secuencias que pertenecen a una misma clase es de 0.25, este valor podría ser útil para evaluar el nivel taxonómico de un ribogrupo.
- 3) Una típica comunidad epipelágica está constituida por Alveolata (47%), Estrapilós (19%) y Rhizaria (13%). El resto de la comunidad está compuesto por Archaeplastida y CCTH. A menudo, los Fungi y los Excavata son realmente pocos (menos del 1 %) o no tienen representación.
- 4) La abundancia de microeucariotas es de 54 ± 5 células mL⁻¹ para la capa mesopelágica y de 14 ± 1 de células mL⁻¹ para la batipelágica, su variabilidad se explica principalmente por la profundidad, la abundancia de procariotas y la concentración de oxígeno (en orden de importancia).
- 5) El tamaño de las células promedio aumenta con la profundidad, el número de células de más de $35 \mu\text{m}^3$ ($> 4 \mu\text{m}$ de diámetro) representa el 12% a 200 m y el 22% a 4000 m. La biomasa total varía desde 280 ± 46 pg mL C⁻¹ en la capa superior mesopelágica a 50 ± 14 pg mL C⁻¹ en la capa más profunda.
- 6) Los resultados del análisis de la pirosecuenciación 454 se compararon con un análisis metagenómico paralelo. En general, el porcentaje de los supergrupos era muy similar con los dos métodos (Figura 7A). Las diferencias que se encontraron fueron una presencia inferior de Alveolados y una muy superior de Excavata en el metagenoma.
- 7) Cuatro clases abundantes componen principalmente la comunidad batipelágica: Collodaria, Chrysophyceae MALV-II y Basidiomycota. Mientras que la composición de la comunidad es bastante homogénea entre las muestras la distribución de estas clases es heterogénea.
- 8) La diferencia en la composición de la comunidad entre las muestras se explica bien debido a la pertenencia a una masa de agua (26 %) y al ratio entre procariotas y microeucariotas (34%).

General references

- Adl SM, Simpson AGB, Lane CE, Lukeš J, Bass D, Bowser SS *et al* (2012). The Revised Classification of Eukaryotes. *J Euk Microbiol* **59**: 429-514.
- Amaral-Zettler LA, McCliment EA, Ducklow HW, Huse SM (2009). A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA genes. *PLoS ONE* **4**: e6372, doi:10.1371/journal.pone.0006372.
- Amato A, Kooistra W, Ghiron JHL, Mann DG, Proschold T, Montresor M (2007). Reproductive isolation among sympatric cryptic species in marine diatoms. *Protist* **158**: 193-207.
- Arístegui J, Duarte CM, Gasol JM, Alonso-Sáez L (2005). Active mesopelagic prokaryotes support high respiration in the subtropical northeast Atlantic Ocean. *Geophys Res Lett* **32**: L03608, doi:10.1029/2004gl021863
- Arístegui J, Duarte CM, Gasol JM, Herndl GJ (2009). Microbial oceanography of the dark ocean's pelagic realm. *Limnol Oceanogr* **54**: 1501-1529.
- Bass D, Richards T, Matthai L, Marsh V, Cavalier-Smith T (2007). DNA evidence for global dispersal and probable endemism of protozoa. *BMC Evol Biol* **7**: 162.
- Baurain D, Brinkmann H, Petersen J, Rodríguez-Ezpeleta N, Stechmann A, Demoulin V *et al* (2010). Phylogenomic Evidence for Separate Acquisition of Plastids in Cryptophytes, Haptophytes, and Stramenopiles. *Mol Biol Evolution* **27**: 1698-1709.
- Benner R (2002). Chemical composition and reactivity. In DA Hansell and CA Carlson (Eds.), *Biogeochemistry of marine dissolved organic matter*. Elsevier Science. pp 59-90.
- Boras JA, Montserrat Sala M, Baltar F, Arístegui J, Duarte CM, Vaqué D (2010). Effect of viruses and protists on bacteria in eddies of the Canary Current region (subtropical northeast Atlantic). *Limnol Oceanogr* **55**: 885-898.
- Burki F, Schalchian-Tabrizi K, Pawlowski J (2008). Phylogenomics reveals a new “megagroup” including most photosynthetic eukaryotes. *Biology Lett* **4**: 366-369.
- Burki F, Inagaki Y, Bråte J, Archibald JM, Keeling PJ, Cavalier-Smith T *et al* (2009). Large-Scale Phylogenomic Analyses Reveal That Two Enigmatic Protist Lineages, Telonemia and Centroheliozoa, Are Related to Photosynthetic Chromalveolates. *Genome Biol Evol* **1**: 231-238.
- Burki F, Okamoto N, Pombert J-F, Keeling PJ (2012). The evolutionary history of haptophytes and

- cryptophytes: phylogenomic evidence for separate origins. *Proc R Soc B* **279**: 2246-2254.
- Caron DA, Countway PD, Savai P, Gast RJ, Schnetzer A, Moorthi SD (2009). Defining DNA-based operational taxonomic units for microbial-eukaryote ecology. *Appl Environ Microbiol* **75**: 5797-5808.
- Chambouvet A, Morin P, Marie D, Guillou L (2008). Control of Toxic Marine Dinoflagellate Blooms by Serial Parasitic Killers. *Science* **322**: 1254-1257.
- Cheung MK, Au CH, Chu KH, Kwan HS, Wong CK (2010). Composition and genetic diversity of picoeukaryotes in subtropical coastal waters as revealed by 454 pyrosequencing. *ISME J* **4**: 1053-1059.
- Christaki U, Courties C, Massana R, Catalá P, Lebaron P, Gasol JM *et al* (2011). Optimized routine flow cytometric enumeration of heterotrophic flagellates using SYBR Green I. *Limnol Oceanogr: Methods* **9**: 329-339.
- Cottrell MT, Wood DN, Yu L, Kirchman DL (2000). Selected Chitinase Genes in Cultured and Uncultured Marine Bacteria in the α - and γ -Subclasses of the Proteobacteria. *Appl Environ Microbiol* **66**: 1195-1201.
- Countway PD, Gast RJ, Dennett MR, Savai P, Rose JM, Caron DA (2007). Distinct protistan assemblages characterize the euphotic zone and deep sea (2500 m) of the western North Atlantic (Sargasso Sea and Gulf Stream). *Environ Microbiol* **9**: 1219-1232.
- Díez B, Pedrós-Alió C, Massana R (2001). Study of genetic diversity of eukaryotic picoplankton in different oceanic regions by small-subunit rRNA gene cloning and sequencing. *Appl Environ Microbiol* **67**: 2932-2941.
- Dennis, RWG (1995) *Fungi of South East England*. Royal Botanic Gardens, Kew.
- Edgcomb V, Kysela D, Teske A, de Vera Gomez A, Sogin M (2002). Benthic eukaryotic diversity in the Guaymas Basin hydrothermal vent environment. *Proc Natl Acad Sci USA* **99**: 7658 - 7662.
- Edgcomb V, Beaudoin D, Gast R, Biddle JF, Teske A (2011). Marine subsurface eukaryotes: the fungal majority. *Environ Microbiol* **13**: 172-183.
- Epstein PR, Ford TE, Colwell RR (1993). Health and climate change: Marine ecosystems. *The Lancet* **342**: 1216-1219.

- Finlay B (2004). Protist taxonomy: an ecological perspective. *Phil Trans Roy Soc Lond B* **359**: 599 - 610.
- Fu Y, O’Kelly C, Sieracki M, Distel DL (2003). Protistan Grazing Analysis by Flow Cytometry Using Prey Labeled by In Vivo Expression of Fluorescent Proteins. *Appl Environ Microbiol* **69**: 6848-6855.
- Fukuda H, Sohrin R, Nagata T, Koike I (2007). Size distribution and biomass of nanoflagellates in meso- and bathypelagic layers of the subarctic Pacific. *Aquat Microb Ecol* **46**: 203-207.
- Giering SLC, Sanders R, Lampitt RS, Anderson TR, Tamburini C, Boutrif M *et al* (2014). Reconciliation of the carbon budget in the ocean’s twilight zone. *Nature* **507**: 480-483.
- Groissillier A, Guillou L, Massana R, Valentin K, Vaultot D (2006). Genetic diversity and habitats of two enigmatic marine alveolate lineages. *Aquat Microb Ecol* **42**: 277-291.
- Guillou L, Bachar D, Audic S, Bass D, Berney C, Bittner L *et al* (2013). The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. *Nucleic Acids Res* **41**: 597-604.
- Hampl V, Hug L, Leigh JW, Dacks JB, Lang BF, Simpson AGB *et al* (2009). Phylogenomic analyses support the monophyly of Excavata and resolve relationships among eukaryotic “super-groups”. *PNAS* **106**:3859-3864.
- Hansel CM, Zeiner CA, Santelli CM, Webb SM (2012). Mn(II) oxidation by an ascomycete fungus is linked to superoxide production during asexual reproduction. *PNAS* **109**: 12621-12625.
- Hansell DA, Carlson CA (1998). Deep-ocean gradients in the concentration of dissolved organic carbon. *Nature* **395**: 263-266.
- Hansell DA, Carlson CA, Repeta DJ, Schlitzer R (2009). Dissolved organic matter in the ocean: A controversy stimulates new insights. *Oceanography* **22**:202–211
- Holen DA, Biraas ME (1996) Mixotrophy in chrysophytes. *In*: D Craig, CD Sandgren, JP Smol, J Kristiansen (Eds.) *Chrysophyte algae. Ecology, phylogeny and development*, University Press, Leiden (1996), pp 119–140.
- Jebaraj CS, Raghukumar C, Behnke A, Stoeck T (2010). Fungal diversity in oxygen-depleted regions of the Arabian Sea revealed by targeted environmental sequencing combined with cultivation. *FEMS Microbiol Ecol* **71**: 399-412.

- Jeon SO, Bunge J, Stoeck T, Barger KJA, Hong SH, Epstein SS (2006). Synthetic statistical approach reveals a high degree of richness of microbial eukaryotes in an anoxic water column. *Appl Environ Microbiol* **72**: 6578-6583.
- Jezbera J, Hornák K, Simek K (2005). Food selection by bacterivorous protists: insight from the analysis of the food vacuole content by means of fluorescence in situ hybridization. *FEMS Microbiol Ecol* **52**: 351-363.
- Jones RI (2000). Mixotrophy in planktonic protists: an overview. *Freshwat Biol* **45**: 219-226.
- Kunin V, Engelbrektson A, Ochman H, Hugenholtz P (2010). Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. *Environ Microbiol* **12**: 118-123.
- Lara E, Moreira D, Vereshchaka A, López-García P (2009). Pan-oceanic distribution of new highly diverse clades of deep-sea diplomonads. *Environ Microbiol* **11**: 47-55.
- Li W (1994). Primary production of prochlorophytes, cyanobacteria, and eucaryotic ultraphytoplankton: measurements from flow cytometric sorting. *Limnol oceanog*: 169-175.
- Libes SM (1992). An introduction to marine biogeochemistry. Wiley
- Logares R, Sunagawa S, Salazar G, Cornejo-Castillo FM, Ferrera I, Sarmiento H *et al* (2013). Metagenomic 16S rDNA Illumina tags are a powerful alternative to amplicon sequencing to explore diversity and structure of microbial communities. *Environ Microbiol* doi: 10.1111/1462-2920.12250
- Lopez-Garcia P, Rodriguez-Valera F, Pedros-Alio C, Moreira D (2001). Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. *Nature* **409**: 603 - 607.
- Marie D, Simon N, Guillou L, Partensky F, Vaulot D (2001). DNA/RNA Analysis of Phytoplankton by Flow Cytometry. *Current Protocols in Cytometry*. John Wiley & Sons, Inc.
- Massana R, Balagué V, Guillou L, Pedrós-Alió C (2004a). Picoeukaryotic diversity in an oligotrophic coastal site studied by molecular and culturing approaches. *FEMS Microbiol Ecol* **50**: 231-243.
- Massana R, Castresana J, Balagué V, Guillou L, Romari K, Groisillier A *et al* (2004b). Phylogenetic and ecological analysis of novel marine stramenopiles. *Appl Environ Microbiol* **70**: 3528-3534.

- Massana R, Guillou L, Terrado R, Forn I, Pedros-Alio C (2006). Growth of uncultured heterotrophic flagellates in unamended seawater incubations. *Aquat microb ecol* **45**: 171-180.
- Massana R, Pedrós-Alió C (2008). Unveiling new microbial eukaryotes in the surface ocean. *Curr Opin Microbiol* **11**: 213-218.
- Massana R (2009). Pycoeukaryotes. In: M Schaechter (Eds.) *Encyclopedia of Microbiology*. Elsevier Science (2009), pp 674–687.
- Massana R, Unrein F, Rodríguez-Martínez R, Forn I, Lefort T, Pinhassi J *et al* (2009). Grazing rates and functional diversity of uncultured heterotrophic flagellates. *ISME J* **3**: 588-596.
- Massana R (2011). Eukaryotic picoplankton in surface oceans. *Annu Rev Microbiol* **65**: 1-47.
- Massana R, Pernice M, Bunge JA, Campo Jd (2011). Sequence diversity and novelty of natural assemblages of picoeukaryotes from the Indian Ocean. *ISME J* **5**: 184-195.
- Møller J, Miller M, Kjølner A (1999). Fungal–bacterial interaction on beech leaves: influence on decomposition and dissolved organic carbon quality. *Soil Biol Biochem* **31**: 367-374.
- Moon-van der Staay SY, De Wachter R, Vaulot D (2001). Oceanic 18S rDNA sequences from picoplankton reveal unsuspected eukaryotic diversity. *Nature* **409**: 607-610.
- Morgan-Smith D, Herndl GJ, van Aken HM, Bochdansky AB (2011). Abundance of eukaryotic microbes in the deep subtropical North Atlantic. *Aquat Microb Ecol* **65**: 103-115.
- Morgan-Smith D, Clouse MA, Herndl GJ, Bochdansky AB (2013). Diversity and distribution of microbial eukaryotes in the deep tropical and subtropical North Atlantic Ocean. *Deep Sea Res Pt-I* **78**: 58-69.
- Nagata T, Tamburini C, Arístegui J, Baltar F, Bochdansky AB, Fonda-Umani S *et al* (2010). Emerging concepts on microbial processes in the bathypelagic ocean – ecology, biogeochemistry, and genomics. *Deep Sea Res Pt-II* **57**: 1519-1536.
- Not F, Valentin K, Romari K, Lovejoy C, Massana R, Tobe K (2007a). Picobiliphytes: a marine picoplanktonic algal group with unknown affinities to other eukaryotes. *Science* **315**: 252-254.
- Not F, Gausling R, Azam F, Heidelberg JF, Worden AZ (2007b). Vertical distribution of picoeukaryotic diversity in the Sargasso Sea. *Environ Microbiol* **9**: 1233-1252.
- Olson RJ, Vaulot D, Chisholm SW (1985). Marine-phytoplankton distributions measured using

- shipboard flow-cytometry. *Deep Sea Res* **32**: 1273-1280.
- Pernice MC, Logares R, Guillou L, Massana R (2013). General patterns of diversity in major marine microeukaryote lineages. *PLoS ONE* **8**: e57170.
- Pernthaler J, Pernthaler A, Amann R (2003). Automated enumeration of groups of marine picoplankton after fluorescence in situ hybridization. *Appl Environ Microbiol* **69**: 2631-2637.
- Porter K, Feig Y (1980). The use of DAPI for identifying and counting aquatic microflora. *Limnol oceanog* **25**: 943-948.
- Quince (2009). Accurate determination of microbial diversity from 454 pyrosequencing data. *Nature Methods* **6**: 639-641.
- Raghukumar S, Ramaiah N, Raghukumar C (2001). Dynamics of thraustochytrid protists in the water column of the Arabian Sea. *Aquat Microb Ecoly* **24**: 175-186.
- Richards TA, Jones MDM, Leonard G, Bass D (2012). Marine Fungi: Their Ecology and Molecular Diversity. *Ann Rev Mar Sci* **4**: 495-522.
- Rosselló-Mora R, Amann R (2001). The species concept for prokaryotes. *FEMS Microbiol Rev* **25**: 39-67.
- Saiki R, Scharf S, Faloona F, Mullis K, Horn G, Erlich H *et al* (1985). Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* **230**: 1350-1354.
- Salani FS, Arndt H, Hausmann K, Nitsche F, Scheckenbach F (2012). Analysis of the community structure of abyssal kinetoplastids revealed similar communities at larger spatial scales. *ISME J* **6**: 713-723.
- Sanders RW (1991). Mixotrophic Protists In Marine and Freshwater Ecosystems. *J Protozool* **38**: 76-81.
- Sanders RW, Caron DA, Davidson JM, Dennett MR, Moran DM (2001). *Nutrient Acquisition and Population Growth of a Mixotrophic Alga in Axenic and Bacterized Cultures*, vol. 42.
- Sauvadet A-L, Gobet A, Guillou L (2010). Comparative analysis between protist communities from the deep-sea pelagic ecosystem and specific deep hydrothermal habitats. *Environ Microbiol* **12**: 2946-2964.

- Schlegel M, Meisterfeld R (2003). The species problem in protozoa revisited. *Eur J Protistol* **39**: 349-355.
- Schnepf E, Drebes G, Elbrachter M (1990). *Pirsonia guinardiae*, gen. et spec. nov.: a parasitic flagellate on the marine diatom *Guinardia flaccida* with an unusual mode of food uptake, vol. 44. Biologische Anstalt Helgoland: Hamburg, Germany.
- Seenivasan R, Sausen N, Medlin LK, Melkonian M (2013). *Picomonas judraskeda* Gen. Et Sp. Nov.: The First Identified Member of the Picozoa Phylum Nov., a Widespread Group of Picoeukaryotes, Formerly Known as ‘Picobiliphytes’. *PLoS ONE* **8**: e59565 doi:10.1371/journal.pone.0059565
- Siano R, Alves-de-Souza C, Foulon E, Bendif EM, Simon N, Guillou L *et al* (2011). Distribution and host diversity of Amoebozoa parasites across oligotrophic waters of the Mediterranean Sea. *Biogeosciences* **8**: 267-278.
- Skovgaard A (2014). Dirty Tricks in the Plankton: Diversity and Role of Marine Parasitic Protists. *Acta Protozool* **53**:51-62
- Sohrin R, Imazawa M, Fukuda H, Suzuki Y (2010). Full-depth profiles of prokaryotes, heterotrophic nanoflagellates, and ciliates along a transect from the equatorial to the subarctic central Pacific Ocean. *Deep Sea Res Pt-II* **57**: 1537-1550.
- Stoeck T, Taylor G, Epstein S (2003). Novel eukaryotes from a permanently anoxic Cariaco Basin (Caribbean Sea). *Appl Environ Microbiol* **69**: 5656 - 5663.
- Stoeck T, Behnke A, Christen R, Amaral-Zettler L, Rodriguez-Mora MJ, Chistoserdov A (2009). Massively parallel tag sequencing reveals the complexity of anaerobic marine protistan communities. *BMC Biology* **7**: 72.
- Tanaka T, Rassoulzadegan F (2002). Full-depth profile (0–2000 m) of bacteria, heterotrophic nanoflagellates and ciliates in the NW Mediterranean Sea: Vertical partitioning of microbial trophic structures. *Deep Sea Res Pt-II* **49**: 2093-2107.
- Thingstad TF, Havskum H, Garde K, Riemann B (1996). On the Strategy of “Eating Your Competitor”: A Mathematical Analysis of Algal Mixotrophy. *Ecology* **77**: 2108-2118.
- Våge S, Castellani M, Giske J, Thingstad TF (2013). Successful strategies in size structured mixotrophic food webs. *Aquat Ecol* **47**: 329-347.

- Vaulot D, Eikrem W, Viprey M, Moreau H (2008). The diversity of small eukaryotic phytoplankton ($\leq 3 \mu\text{m}$) in marine ecosystems. *FEMS Microbiol Rev* **32**: 795-820.
- Venter J, Remington K, Heidelberg J, Halpern A, Rusch D, Eisen J *et al* (2004). Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304**: 66-74.
- von der Heyden S, Chao E, Cavalier-Smith T (2004). Genetic diversity of goniomonads: an ancient divergence between marine and freshwater species. *Eur J Phycol* **39**: 343 - 350.
- Wintzingerode F, Göbel UB, Stackebrandt E (1997). Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. *FEMS Microbiol Rev* **21**: 213-229.
- Woese CR, Fox GE (1977). Phylogenetic structure of the prokaryotic domain: The primary kingdoms. *PNAS* **74**: 5088-5090.
- Worden A (2006). Picoeukaryote diversity in coastal waters of the Pacific Ocean. *Aquat microb ecol* **43**: 165-175.
- Yamaguchi A, Watanabe Y, Ishida H, Harimoto T, Furusawa K, Suzuki S *et al* (2004). Latitudinal Differences in the Planktonic Biomass and Community Structure Down to the Greater Depths in the Western North Pacific. *J Oceanogr* **60**: 773-787.
- Zhu F, Massana R, Not F, Marie D, Vaulot D (2005). Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. *FEMS Microbiol Ecol* **52**: 79-92.
- Zubkov M, Tarran G (2008). High bacterivory by the smallest phytoplankton in the North Atlantic Ocean. *Nature* **455**: 224-226.
- Zubkov MV, Burkill PH (2006). Syringe pumped high speed flow cytometry of oceanic phytoplankton. *Cytometry Part A* **69A**: 1010-1019.
- Zubkov MV, Burkill PH, Topping JN (2007). Flow cytometric enumeration of DNA-stained oceanic planktonic protists. *J Plankton Res* **29**: 79-86.

Agradecimientos



Agradecimientos

Más de una vez me han llamado afortunado y probablemente lo soy. No es el tipo de fortuna que te hace ganar la lotería, o que te hace encontrar dinero en el suelo, es más de esta suerte que hace que las cosas sean leves y que vayan en la dirección correcta, hacia lo bonito de la vida. Estos años han sido una gran muestra de esta fortuna, un río de regalos, estímulos y experiencias que ni siquiera me imaginaba. Esta fortuna no es una diosa con ojos vendados, esta fortuna tiene caras y nombres y hoy quiero agradecerlos.

Ringrazio prima di tutto i miei genitori per avermi dato la possibilità di arrivare dove sono adesso, sono due persone bellissime e sono fortunato di essere figlio loro. Un ringraziamento per il sostegno e l'allegria a mia sorella Cristina, e ai miei fratelli Marco e Francesco, vi auguro che i vostri sogni e progetti si realizzino.

Agraeixo al meu “cap” Ramon, per haver-me donat la possibilitat de fer un doctorat, i per haver-me ajudat moltíssim en el desenvolupament d’aquesta tesi. Tot i la normal tensió que comporta la fase final segueixo pensant que ha estat una sort tenir un coordinador amb els peus a terra considerant que el meu cap se’n va als núvols.

Agradezco, a la gran familia ICM, a la que estoy contento y orgulloso de pertenecer. La especial calidad de las personas que se pasean en este edificio sigue sorprendiéndome, he tenido la posibilidad de relacionarme con diferentes micro-clusters sociales, que como en mi tesis a menudo se distinguen por su comportamiento trófico. . .

Agradezco a los de “comemos dentro que es más cómodo, a ser posible en la mesa redonda”: Raquel por su soporte y su presencia constante, Montse pel riure i l’especial complicitat, Bea por su maravillosa y gran energía, Eli por cuidar de mis aportes de vitaminas, Ero por las sabrosas castañadas, Irene F. por su macroscópica ayuda al microscopio, Pedro muy presente en los domingos de esta última fase y Juanlu por las charlitas naturalísticas.

Agradezco a los de “comemos al solete que es más chulo y por supuesto carajillo de Baileys”: Sara por las charlas desahogantes en el despacho, Guillem por las conversaciones científicas y por las conversaciones absurdas, Elena por nuestras indispensables pausas juntos, Fran por su talento pa’ cantar y reírse, Anamari por su alma sureña llena de sol, Eli A. por su sonrisa.

Agradezco a los de “hoy como donde me da la gana y sábado hay barbacoa”: Rosana que por suerte se quedó más de seis meses, Roy por ser una cascada de alegría, María d F. que sabe lo mucho que me encanta, Ana G. por su dulzura, Rodrigo por haber sido acogedor conmigo, Sarah Jeanne por su estilo loco que me llena de alegría, Suso por su simpatía y Xavi Leal por su buen humor. Y “los nuevos”, que nuevos ya no son: Elisa, Mireia, Fran (hola), Encarna y Caterina que va a llegar lejos. Un gracias especial a “los italianos”: Rachele, Miriam, Stefano, Elisabetta, parlare ogni tanto italiano é un sollievo.

Agradezco a los “gourmet del tupper” siempre disponibles a echar una mano, Ramiro que es indispensable, Pablo por su gran simpatía y sus muchos postres, Isabel por la alegría en las fiestas, Marta por ser majísima y Bibiana por su ayuda estadística.

Agradezco a los de “vamos al restaurante”, todos los jefes del departamento, importante referencia en estos años en particular: Marta Estrada, Carlos Pedrós, Pep Gasol, Silvia Acinas, Rafel Simó, Celia Marrasé, Montse Sala, Dolors Vaqué, Elisa Berdalet, Esther Garcés, Albert Calbet, Miquel Alcaraz y Enric Saiz. Y aprovecho también para agradecer a las técnicos, indispensables en cada paso, Clara e Vanessa, por toda su ayuda en esta tesis.

Agradezco a los de “Jamón, I miss you” que ya se fueron al otro lado del mar: Clara, Juancho, Arancha, Hugo, Ivo, Pati, Thomas, Javier y Marionna, a la que quiero muchísimo. Ha sido bonito compartir todo este tiempo, pienso mucho en vosotros ahora que ya me toca marchar a mí.

Agradezco a los de “Había un Jamón en el Hespérides”, mis queridos amigos malaspinianos, Lara por su gran corazón, Xiker por nuestro viaje, Víctor por su sentido del humor, Laura y Belén por las cervecitas.

Quedan muchas personas más y espero no olvidarme a nadie, gracias a Lorena, Eva Flo, Estela, Sdena, Isabel(miniMi), Dafne, Sergio, Conchita B., Nuria A., Eva L., Uxue, Carlos D., Jordi G., Elena G. Ariadna y Marieta.

Estos años han sido leves también por la presencia de otros amigos que me han apoyado. Quiero agradecer ante todo a “mis chicas”, mi primera familia en Barcelona, con quien siempre estaré en deuda por haberme acogido: a mi bellísima Esthel que está siempre presente, a Carol que vale mucho, a Nuria por su espíritu guerrero, a Rosa que sabe bien qué es un doctorado y Angie por su dulzura. Un agradecimiento particular va para mi “familia” actual que se ha tragado todo el estrés final de la tesis, a Gemma por el cariño y la buena vida y a Manuel por la buena vida y el cariño. Agradezco también a “mis chicos” compañeros de numerosos desayunos, meriendas y cenitas: Javi y Quim, Samuel, Enrique, Iñaki, Valentino, Rubén, Lluís y Miguel, hacéis mi vida mucho más divertida e interesante. Gracias también a los Bencinis, Alessia y la Familia Blau que colorean mis mañanas. Un gracias muy muy grande a Irene que me ha dado muchísimo. Un gracias especial a Roberta, que de alguna manera ha sido testigo de todos estos años, eres preciosa en mi vida.

Un grazie agli amici dell'università, Emanuele, Daniele, Florenza, Tiziana, Teresa e Pina con cui iniziai a studiare biologia. E non poteva mancare un ringraziamento agli amici di tutta una vita: Olga, Antonino, Ilaria, Emanuela, Valeria e Sabina, grazie per essere stati presenti in questi anni e perdonatemi il poco che mi sono fatto sentire.

Os agradezco a todos *mis fortunas*, gracias por haber paseado por mi vida.

“E quindi uscimmo a riveder le stelle.”

Dante alighieri, *Inferno* XXXIV, 139

*Has vist que bé que he parlat?
Quin discurs tan ben travat.
Quins principis clars i fermes,
dignes d'un home de seny.
Però, un avís per navegants:
Fes-me cas els dies senars,
i els parells, fes com qui sent
que a la platja hi xiula el vent.*

(Manel)