
Coupling between taxonomic and functional diversity in protistan coastal communities

Ramond Pierre^{1,2}, Sourisseau Marc², Simon Nathalie¹, Romac Sarah¹, Schmitt Sophie², Rigaut-Jalabert Fabienne³, Henry Nicolas¹, De Vargas Colomban¹, Siano Raffaele^{1,2,*}

¹ Sorbonne Université, CNRS - UMR7144 - Station Biologique de Roscoff; Place Georges Teissier, 29688 Roscoff ,FRANCE

² Dyneco Pelagos, IFREMER; BP 70, 29280 Plouzané ,France

³ Sorbonne Université, CNRS - FR2424, Station Biologique de Roscoff, Place Georges Teissier; 29688 Roscoff ,FRANCE

* Corresponding author : Raffaele Siano, email address : raffaele.siano@ifremer.fr

Abstract :

The study of protistan functional diversity is crucial to understand the dynamics of oceanic ecological processes. We combined the metabarcoding data of various coastal ecosystems and a newly developed trait-based approach to study the link between taxonomic and functional diversity across marine protistan communities of different size-classes. Environmental DNA was extracted and the V4 18S rDNA genomic region was amplified and sequenced. In parallel, we tried to annotate the Operational Taxonomic Units (OTUs) from our metabarcoding dataset to 30 biological traits using published and accessible information on protists. We then developed a method to study trait correlations across protists (i.e. trade-offs) in order to build the best functional groups. Based on the annotated OTUs and our functional groups, we demonstrated that the functional diversity of marine protist communities varied in parallel with their taxonomic diversity. The coupling between functional and taxonomic diversity was conserved across different protist size classes. However, the smallest size-fraction was characterized by wider taxonomic and functional groups diversity, corroborating the idea that nano- and pico-plankton are part of a more stable ecological background on which larger protists and metazoans might develop.

Introduction

Pelagic protists represent the majority of the eukaryotic diversity in the oceans (de Vargas *et al.*, 2015) and fluctuations in protist community composition modulate global ecosystem processes (Worden *et al.*, 2015; Guidi *et al.*, 2016). Since the end of the last century, a large share of protistan diversity has been unveiled by molecular methods (Caron *et al.*, 2012), while recent high-throughput sequencing of genomic markers (barcodes) among complex communities (metabarcoding) has provided a greater insight into both oceanic (de Vargas *et al.*, 2015; Pernice *et al.*, 2016) and coastal protists (Christaki *et al.*, 2014; Massana *et al.*, 2015; Hu *et al.*, 2016). In parallel, the functional role of protistan ecological strategies (e.g. mixotrophy or parasitism) has gradually been recognized as crucial in the oceans (Jephcott *et al.*, 2016; Mitra *et al.*, 2016; Scholz *et al.*, 2016; Ward and Follows, 2016; Stoecker *et al.*, 2017). Yet, only few efforts have been made to translate molecular diversity (i.e. Molecular Operational Taxonomic Units, MOTUs or OTUs) into functional diversity (but see de Vargas *et al.*, 2015; Genitsaris *et al.*, 2015). Understanding the relationship between functional and taxonomic diversity in plankton remains a major challenge of modern microbiology, particularly in the face of climate change and its impact on the pelagic ecosystem (Beaugrand and Kirby, 2018).

Metagenomic and metatranscriptomic analyses of marine prokaryotes have shown that communities differing in their taxonomic diversity can express similar functional roles by means of shared metabolic pathways (Louca *et al.*, 2016; Coles *et al.*, 2017; Haggerty and Dinsdale, 2017). A supposed outcome of this functional redundancy would be that changes in taxonomy and functions would be decoupled in time and space. Such a decoupling was observed recently for marine prokaryotes (Louca *et al.*, 2016; Haggerty and Dinsdale, 2017), leading to the foundation of a new microbial paradigm in which the functional roles would be selected by the environment while the identity of the species playing the roles (i.e. taxonomic diversity) would be independent and driven by dispersal and biotic interactions (Louca *et al.*, 2018). This paradigm is still being debated (Galand *et*

al., 2018) and, as for now, has excluded protists due to the present limitations in the genome analysis of micro-eukaryotes (Keeling and del Campo, 2017).

However, the functional diversity of marine protists could be studied using a biological trait approach, following the examples of research on terrestrial plants (Lavorel and Garnier, 2002), or benthic (Rigolet *et al.*, 2014) and pelagic macroorganisms (Mouillot *et al.*, 2014). Trait-based approaches are a field of functional ecology that consist in using functional traits to predict the performances of species within ecosystems (i.e. the fitness of species) (Violle *et al.*, 2007). Functional traits represent any metric (morphology, physiology, life history, trophic strategy) that influences or relates to the fitness of an organism by impacting its reproduction, survival or resource acquisition (Violle *et al.*, 2007; Litchman and Klausmeier, 2008; Litchman *et al.*, 2013). Similar trait patterns across species highlight physiological laws, compromises and constraints (i.e. trade-offs), which reveal ecological strategies (i.e. repeated trait combinations across organisms) that are often involved in species functional roles (the role of an organism in its ecosystem; Diaz *et al.*, 2013). Regarding protists, functional studies of phototrophic organisms (phytoplankton) date back to the works of Margalef (1978) and Reynolds (1984) and for heterotrophs to Fenchel's studies (1982a, 1982b). These frameworks have been successfully used to predict phytoplankton successions (Smayda and Reynolds, 2003; Alves-De-Souza *et al.*, 2008; Kruk *et al.*, 2011) and to describe the functional responses of heterotrophic protists (Massana *et al.*, 2009; Yang *et al.*, 2013). Recent reviews of relevant functional traits in plankton laid down the baseline for a trait-based approach to marine protistan functional diversity (Litchman and Klausmeier, 2008; Litchman *et al.*, 2013).

This study focuses on the relationship between protistan functional and taxonomic diversity within marine coastal ecosystems. Taxonomic diversity was assessed using the deep resolution of metabarcoding and was associated with functional diversity using a newly created trait approach. Coastal ecosystems were targeted because they encompass physical and hydrodynamic processes (e.g.

tides, currents, upwellings, pulses of nutrients, changes in salinity, oxygen or temperature, due to seasonal cycles of freshwater inputs, and exchanges with the atmosphere or the sea bottom) that shape the taxonomic diversity of marine protists, potentially corresponding to differing ecological strategies (Cloern, 1996; Barton *et al.*, 2010; Telesh *et al.*, 2013; Lallias *et al.*, 2015; Pearman *et al.*, 2017). In this study, the relationship between protistan functional and taxonomic diversity is detailed for the first time across marine micro-, nano- and pico-plankton communities.

Results

Environmental characteristics of the sampled ecosystems

A total of 277 water samples were collected in a temporal and/or spatial manner across the coasts of the north Atlantic Ocean (Fig. 1), representing various environmental gradients (Table 1). Chemo-physical variables (temperature, salinity, and nutrients NO_3^- , NO_2^- , PO_4^{3-} , NH_4^+ and $\text{Si}(\text{OH})_4^-$) were measured in all datasets, with the exception of Senegalese samples which lacked temperature and salinity measurements, and the 2015 samples of the PI and PH cruises which lacked the whole environmental set. The Principal Component Analysis (PCA; Fig. S1) of this environmental dataset showed two major gradients. On the first PCA axis (PCA1, 39.92 % of the explained variance) samples were distributed along a gradient of nutrient concentrations. On the second axis of the PCA (PCA2, 34.14% of the explained variance) the samples were separated along a salinity-temperature gradient, distinguishing notably oceanic from estuarine waters.

Genetic diversity

Plankton samples consisted of 1145 distinct filters with different pore sizes ([20 or 10 μm], [3 μm] and [0.2 μm]), which separated protist communities into three size-fractions, here called micro-, nano- and pico-plankton. DNA from the water filters was extracted and the V4 domain of 18S-rDNA was

Accepted Article

amplified and sequenced. After filtration steps, sequences were clustered into OTUs with Swarm 2 (Mahé *et al.*, 2014, 2015) and taxonomically annotated with PR² (Guillou *et al.*, 2013). We retrieved 111 089 distinct OTUs, which accounted for 3.5 x10⁶ reads. Each OTUs was annotated with a taxonomy corresponding to 2007 single taxonomic references (many references were repeated across distinct OTUs). Finally, we created a taxonomic community table based on the relative abundances of each OTU in each sample.

Rarefaction curves based on the OTUs present in our dataset did not reach the theoretical asymptotic shape (Fig. S2), neither when computed on separate datasets nor when performed by size fractions. Following taxonomic assignment, OTUs with low taxonomic level (i.e. annotated with a taxonomic reference only at the family, supergroup or lower) constituted on average 29% reads per sample (min = 0.3%; 1st quartile = 17%; 3rd quartile = 40%; max = 95%). Merging all our samples, coastal communities were mostly dominated by Dinophyta, Bacillariophyta, Chlorophyta, the marine Alveolate (MALV) and Stramenopile (MAST) groups, and Cryptophyta (Fig. 2a). The distribution of these taxa was uneven across size-fractions. Dinophyta (45%) and Bacillariophyta (18%) constituted most of the read number within the micro-plankton (Fig. 2a). The nano-plankton was characterized by a more diversified assemblage comprising not only Dinophyta (23%) and Bacillariophyta (17%) but also Chlorophyta (6%), Cryptophyta (5%) and MALV (3%). Finally, the pico-plankton contained Chlorophyta (22%), Dinophyta (10%), MALV (8%), Bacillariophyta (7%), MAST (4%), Cryptophyta (4%) and Picomonadida (3%). OTUs associated to fungi and Radiolaria were present at very low abundance in all samples. Ciliates (“Ciliophora”) were often observed at low abundances, and across all size-fractions (micro-: 3%, nano-: 2% and pico-: 4%).

Trait approach and functional groups

Accepted Article

First, a conceptual framework of 30 biological traits distinguishing the morphology, trophic strategy, physiology, and mode of life of both photoautotrophic and heterotrophic protists was created (see Supplementary Material 1 for the ecological relevance of each trait and Fig. 3a) (Ramond *et al.*, 2018). Then, we tried to annotate these traits to our OTUs using their taxonomy. Based on a literature survey, we searched for the biological description of the 2007 single taxonomic references annotated to our OTUs and issued from PR² (Guillou *et al.*, 2013), we then tried to annotate them with a modality for each trait. Trait annotation was inferred from 717 literature sources, ranging from general protistology handbooks to specialized papers and websites (the bibliography was annotated for each taxonomic reference). The final annotated table represents the first trait annotation of marine protists; it is public and still improvable (Ramond *et al.*, 2018). In our study, we used a reduced biological trait table composed of 13 well-annotated traits; SizeMin, SizeMax, Cell Cover, Cell Shape, Presence of Spicule, Cell Symmetry, Cell Polarity, Coloniality, Motility, Chloroplast Origin, Ingestion Method, Symbiosis Type and Resting Stage during the life cycle; that were inferred for 1669 of the 2007 taxonomic references (83%; Fig. 3b). The discarded traits and taxonomic references were poorly annotated due to either a lack of available information or to the low taxonomic level of some taxonomic references.

We then developed a method to identify trait trade-offs and functional groups on the basis of our biological trait table (see Experimental Procedures for further details). Following a statistical approach developed for the study of functional diversity (Maire *et al.*, 2015), our biological trait table of 13 traits was summarized into 5 Euclidean dimensions (descriptors of our 1669 taxonomic references; Fig. S3). We investigated the correlation between our 13 traits and the 5 Euclidean dimensions, traits that were correlated on the same dimension were considered as being part of a trade-off (see experimental procedures; Fig. S4). We identified a trade-off between 8 morphological and trophic traits (Fig. S4). To construct the most coherent functional groups, only the dimensions that highlighted trade-offs were used to cluster taxonomic references (1669) with similar traits. With an

unsupervised clustering method 6 functional groups were evidenced (Fig. S5). Finally, we studied the dominant modalities of traits in each functional group which enabled to distinguish various ecological strategies (Fig. S6 to S11). The functional groups were named accordingly: 1) Parasites (PARA): characterized by their type of feeding, their symbiosis type, their host-attached life strategy and mostly naked cell surface; 2) Strict-heterotrophs (HET): characterized by phagotrophy and the absence of plastids throughout their life cycle; 3) Saprobes (SAP): characterized by their feeding strategy (osmotrophic and saprotrophic); 4) Swimmer-photoautotrophs (SWAT): characterized by swimming abilities, dominantly organic covers, phototrophy and mixotrophic tendencies; 5) Floater-photoautotrophs (FLAT) characterized by non-swimming abilities dominantly siliceous cover, phototrophy and mixotrophic tendencies; 6) and Colonial-photoautotrophs (CAT): characterized by non-swimming abilities, strict phototrophy and the ability to form colonies.

With this methodology, our 1669 taxonomic references were annotated with 13 traits and assigned with one of our 6 functional groups. The taxonomic references being descriptors of OTUs, we could then estimate the number of OTUs annotated with our traits and to a group. The 1669 annotated taxonomic references corresponded to 52 180 OTUs, a considerably reduced portion (47%) of the original taxonomic community table (111 089). The reliability of the reduced dataset was tested by comparing biodiversity patterns between the reduced and the complete dataset. The Mantel test (Mantel, 1967), calculated on the Bray-Curtis dissimilarity matrices of the two datasets (Ramette, 2007; Buttigieg and Ramette, 2014), was high (Mantel's test observation = 0.91; p-value 0.0001). In addition, two diversity proxies, OTU richness and the Shannon index H' (Piéluou, 1966), calculated for the two datasets, were highly correlated ($R^2 = 0.88$ for OTU richness, $R^2_{171} = 0.75$ for H' ; Fig. S12). Given these strong positive correlations, the reduced dataset was considered to be reliable and to carry the same cross-sample biodiversity patterns of the complete dataset. The 52 180 OTUs were distributed heterogeneously among our functional groups, parasites (PARA): 8 366, strict-heterotrophs

(HET): 19 582, saprobes (SAP): 332, swimmer-phototrophs (SWAT): 14 333, floater-phototrophs (FLAT): 4 813, and colonial-phototrophs (CAT): 4 754 OTUs.

With the taxonomic annotation of the OTUs we could also study the phylogenetic diversity of the 6 functional groups (Fig. 4). Most groups (5 out of 6) proved to be paraphyletic containing OTUs from various branches of the eukaryotic phylogenetic tree. Parasites (PARA) were mostly composed of MALV and Apicomplexans. Strict-heterotrophs (HET) were dominated by Marine Stramenopiles (MAST), Ciliophora, Picomonadida and Dinophyta. Saprobes (SAP) clustered organisms from Ascomycota, Basidiomycota, Bicoceae and Labyrinthulaceae. Swimmer-phototrophs (SWAT) were dominated by Dinophyta, Cryptophyta and Chlorophyta whereas floater-phototrophs (FLAT) were composed of Bacillariophyta (invariably from the family and orders), Acantharia and Chlorophyta. Colonial phototrophs (CAT) were only associated with Bacillariophyta (invariably from the family and orders).

Finally, to characterize the functional diversity of marine protists in coastal ecosystems, the relative abundance of each functional group was calculated on the basis of the relative abundance of the annotated OTUs over the whole dataset (1145 size-fractionated samples) (Fig. 2b). The relative abundance of non-annotated OTUs was high in all size-fractions but lower within the pico-plankton (51%, 56.5%, 40.5% for micro-, nano and pico-plankton, respectively). The contribution of the functional groups varied across the size-fractions, in parallel with a change in taxonomic composition (Fig. 2a and 2b). Micro-plankton was dominated by swimmer-phototrophs (SWAT) and strict-heterotrophs (HET), which together accounted for more than 40% of the annotated reads (30% and 10.5%, respectively), whereas within the nano- and pico-plankton the relative composition of the functional groups was more even. Floater-phototrophs (FLAT) and parasites (PARA) were more common in the pico-plankton (14% and 9%, respectively) than in the larger size-fractions (micro-plankton FLAT: 3%, PARA: 1.5%; nano-plankton and FLAT: 5%, PARA: 3.5%). In contrast,

colonial-phototrophs (CAT) relative abundance was higher in the micro- and nano-plankton (4.5% and 4%, respectively) and lower within the pico-plankton (2.5%). The relative abundance of saprobes (SAP) was very low across all size-fractions (on average $< 0.05\%$ in the micro-, nano- and pico-plankton) and more generally in all samples.

Functional vs. taxonomic diversity of marine protists

The relationships between environmental variables, taxonomic and functional diversity in pico-, nano- and micro-plankton communities were studied using the RV statistical coefficient of co-inertia, a multivariate generalization of the Pearson correlation coefficient (Borcard *et al.*, 2011; Legendre and Legendre, 2012; Husson *et al.*, 2018). Correlations (RV coefficient) between the taxonomic community table and environmental variables were low but significant across all size-fractions (values for micro-: 0.45, nano-: 0.22 and pico-plankton: 0.19, with p -value < 0.0001). Similarly, correlations between the functional community table and environmental variables were low but significant (values for micro-: 0.34, nano-: 0.16 and pico-plankton: 0.10, with p -value < 0.0001). For every size-fraction, the correlations (RV coefficient) between the functional and the taxonomic community table were high and significant (values for micro-: 0.71, nano-: 0.46 and pico-plankton: 0.75, with p -value < 0.0001) meaning that the taxonomic and the functional diversity of marine protists were tightly coupled.

In order to study if protist communities different for their taxonomic composition were characterized by similar composition of our 6 functional groups, we computed a Non-metric Multi-Dimensional Scaling (NMDS) for the samples of micro-, nano- and pico-plankton on the basis of their OTUs composition. On each NMDS, samples were clustered together by an unsupervised partitioning of samples using a k-mean method and a simple structure index (ssi) criterion. The relative abundances of the 6 functional groups within those samples and clusters were calculated. The overall

aim was to compare the functional groups diversity across protistan communities distinct for their taxonomic composition (Fig. 5). To study if there was an effect of the environment on taxonomic and functional composition, environmental variables were projected as vectors onto each NMDS ordination space.

In the NMDS built from the taxonomic tables independently for each size-fraction (Fig. 5), plankton samples were clustered in 4, 5 and 5 homogeneous groups for micro-, nano- and pico-plankton, respectively. Across all size-fractions, samples from the estuarine DA campaign were markedly isolated on the first ordination axis (simple structure index, ssi, cluster 1; Fig. 5), implying a markedly distinct community structure in this set of samples. The functional group structure of the DA samples showed a domination of swimmer-phototrophs (SWAT). This estuarine group of samples was usually opposed on the same axis to samples retrieved from the most off-shore areas (PE in the Bay of Biscay and/or MB in the Iroise Sea; ssi clusters 4, 4 and 3 for micro-, nano- and pico-plankton, respectively; Fig. 5). These clusters were dominated by the swimmer-phototrophs (SWAT) and heterotrophs (HET) groups in the micro- and nano-plankton and characterized by a more diversified assemblage in the pico-plankton, especially with a higher proportion of parasites (PARA), compared to the other size-fractions. Only within the picoplankton was this set of samples well-correlated with ammonium concentration (NH_4^+) and salinity (Sal) (Fig. 5). In comparisons with estuarine and off-shore waters, the second axis of all 3 NMDSs represented more typical coastal waters (DY, PH, PI, RA and the coastal stations of MB; into ssi clusters [2 and 3], [2, 3 and 5] and [2, 4 and 5] for micro-, nano- and pico-plankton, respectively; Fig. 5). These samples correlated well with a gradient opposing salinity and nutrient concentrations (most notably $\text{Si}(\text{OH})_4$), implying a separation between communities of winter/early spring (with enriched and fresher waters) and summer/productive communities (with saltier and depleted waters). Across size-fractions, winter samples were dominated by strict-heterotrophs (HET; ssi clusters 2, 5 and 2 for micro-, nano- and pico-plankton, respectively).

Conversely, summer/depleted conditions coincided with an equilibrated functional composition with notably the 3 phototrophic groups (SWAT, FLAT and CAT) in higher abundance across all size-fractions (ssi clusters [3], [2] and [4 and 5] for micro-, nano- and pico-plankton, respectively). Overall, groups of samples dissimilar in their OTU composition also displayed distinct assemblages of functional groups (Adonis test, p-value = 0.0001, R^2 for micro-: 0.45, nano-: 0.35 and pico-plankton: 0.36) supposing, again, a tight coupling between taxonomy and functions for marine protists.

Interestingly, across size-fractions, functional groups seemed more evenly distributed in the smaller size-fractions while micro-plankton samples were mostly dominated by strict-heterotrophs (HET) and swimmer-phototrophs (SWAT) (Fig. 5). To investigate the functional group distribution across size-fractions, the Shannon index of equitability (Pielou, 1966) was calculated on the basis of the functional group relative abundance in each sample (Fig. 6a). The Kruskal-Wallis test (one-way analysis of variance) indicated that the equitability of functional groups was significantly higher and less variable within the nano- and pico-plankton samples than in micro-plankton (p-value < 0.001; Fig. 6a). Taxonomic equitability (calculated on the relative abundances of OTUs) and richness (number of OTUs per sample) showed similar patterns (Fig. 6b & 6c), with significantly lower values in micro-plankton (p-value < 0.001).

Discussion

By means of a taxonomic diversity analysis obtained by the metabarcoding of V4 18S-rDNA of the environmental DNA and a newly developed trait-based approach, we were able to 1) detail patterns of protistan functional diversity and 2) compare patterns of taxonomic and functional diversity of coastal protists. Our trait-based approach enabled the construction of 6 functional groups, which represented not only relevant ecological strategies but also the functional roles of marine protists. Most functional groups were paraphyletic, being composed of phylogenetically distant groups of protists. Patterns of

the taxonomic and functional diversity of coastal protist communities across various environments were described. Both functional and taxonomic diversity appeared more evenly distributed in the smaller size-fractions, while micro-plankton tended to be dominated by fewer OTUs and functional groups. Lastly, a coupling between protistan taxonomic and functional diversity was highlighted.

From a genetic to functional groups of marine protists

The metabarcoding approach was adopted in this study due to its proven efficiency in analyzing protistan taxonomic diversity (de Vargas *et al.*, 2015). Nevertheless, this approach leads to analytical problems due to methodological limitations. In our study, water-filter clogging resulted in a relatively low sequencing depth, which prevented us from obtaining the whole picture of protistan diversity from the coastal waters considered (Fig. S2; see others e.g. Pernice *et al.*, 2016). Water filtration also contributes to contamination across size-fractions. DNA from cell-breakage and small-sized gametes or resting stages of typically large organisms are often found in marine metabarcoding surveys (Massana *et al.*, 2004, 2015; Le Bescot *et al.*, 2015) and might partially contaminate smaller size-fractions. The taxonomic annotation was also imperfect as 30% of environmental reads were not annotated. More samples and taxonomic descriptions of rare protists are needed to decrease the proportion of non-annotated reads and to fully describe the nature of this microbial compartment (Caron *et al.*, 2012; Guillou *et al.*, 2013; Keeling and del Campo, 2017). Despite these limitations, our dataset still presents a valuable DNA sampling of marine coastal waters (273 water samples, 1145 water filters) and the genetic diversity analyzed in this study can be considered representative of the most abundant species of the sampled coastal protistan communities. The taxa retrieved in each size-class during this study are indeed consistent with other coastal DNA based surveys (Christaki *et al.*, 2014; Genitsaris *et al.*, 2015; Massana *et al.*, 2015; Hu *et al.*, 2016). Dinophyta (dinoflagellates) and Bacillariophyta (diatoms) were markedly dominant in the micro-plankton. These two groups co-

occurred with Chlorophyta, Cryptophyta, Picomonadida, MALV and MAST within the nano-plankton and pico-plankton, the latter being dominated by Chlorophyta.

The following step was to associate these OTUs with biological traits to describe their functional diversity. As demonstrated by statistical tests, OTUs with a good trait annotation represented a reduced (ca. 50%) but representative share of our complete taxonomic table (52 180 out of 111 089 OTUs). Well annotated traits mostly concerned the trait types of morphology and trophic strategy. Within the life cycle trait type, only the production of resting stage was relatively well-annotated. The number of traits to construct functional groups was further reduced with our trade-off approach. These analyses selected 8 out of 13 traits that were well-annotated and well-correlated. Consequently, 5 traits were not significant in our functional groups, although considered to be descriptors of protistan ecology (“resting stage” in Litchman and Klausmeier, 2008, and Lange *et al.*, 2015 ; “size” in Litchman and Klausmeier, 2008; Litchman *et al.*, 2013; Weisse *et al.*, 2016; “spicules” in Hamm, 2005; “cell shape” in Pahlow *et al.*, 1997). In fact, the excluded traits probably form trade-offs with traits absent from our functional table. As an example, cell shape and size usually correlates with growth rate, resource requirement and uptake through allometric laws (Grover, 1989; Nielsen and Sandjensen, 1990; Edwards *et al.*, 2012; Litchman *et al.*, 2013), while resting stages involve strong investment in the life cycle, longevity, stress-resistance and probably benthic-coupling (Marcus and Boero, 1998; Litchman and Klausmeier, 2008). The absence of physiological information on marine protists is thus a major outcome of our trait approach. Physiological and resource acquisition trait types are scarcely annotated since these kind of information is only found for a few cultivated species and cannot be generalized to taxonomic references at low taxonomic levels (genus, families). In future research, the combination of physiological traits and phylogenetic approaches could help to overcome this limitation by summarizing trait values, or probability of values, within larger taxonomic groups (Bruggeman, 2011).

Other biological traits could have also been included in our functional analysis, indeed single-group functional analysis can rely on more specific traits than those used in this study (e.g. for heterotrophs in Fenchel, 1980; Hansen *et al.*, 1997; Weisse *et al.*, 2016). However, in this first attempt to study the functional diversity of marine protist communities, we selected biological traits that could be generalized to the widest range of marine protist species. This functional representation of marine protistan diversity will certainly be improved over time by inputs from other protist specialists. In addition, we applied this theoretical framework exclusively to coastal communities; the study of its relevance to other ecosystems (e.g. off-shore, arctic, deep-sea) remains an interesting objective for future research.

Patterns of functional diversity of coastal protist communities

The 6 functional groups identified in this study define various protistan ecological strategies, which are acknowledged to play key roles in the structuring of pelagic communities (Worden *et al.*, 2015): phototrophs (SWAT, FLAT and CAT), heterotrophs (HET), parasites (PARA) and saprotrophs (SAP). The phototrophic groups were discernible more by their morphological adaptations, with some carrying mixotrophic potential (in SWAT and FLAT), while the three heterotrophic groups were distinguished by their ingestion methods. Except for colonial-phototrophs (CAT), all the functional groups were paraphyletic (Fig. 4), including various protistan phylogenetic branches. The paraphyly of our functional groups supposes that similar functional roles emerged along different lineages of protists. The groups of phototrophs echoed the morphological groups and life-forms first proposed by Margalef (1978) and (Reynolds *et al.*, 1983). The work of these authors indicated that phytoplankton species adapted their shape and morphology mainly to counter sedimentation, resource scarcity and predation. Our functional groups distinguished swimmers (SWAT), floaters (FLAT) and colony-forming (CAT) species, which also represent adaptations to sedimentation (i.e. swimming can

Accepted Article

counterbalance sedimentation while colonies increase cell buoyancy, e.g. in Pahlow *et al.*, 1997; Ploug *et al.*, 1999) and predation (swimmers can avoid predation while colonies may discourage predators; Margalef, 1978). The three heterotrophic groups were divided into three distinct feeding-strategies distinguished according to their prey and/or food preferences (saprobes: dissolved and detrital matter, strict-heterotrophs: prey; parasites: host type). By applying specific functional traits, the strict heterotrophs (HET) group could probably be further subdivided with the addition of phagotrophic-related traits, such as functional and numerical responses (Yang *et al.*, 2013; Weisse *et al.*, 2016), feeding mechanisms (Kiørboe, 2011) or the maximal size of engulfment apparatus.

The distribution of functional groups varied across the size-fractions (Fig. 2b). The co-existence of both phototrophic and heterotrophic functional groups in the smallest plankton can be explained by further adaptations. Pico-sized phototrophs have indeed smaller requirements in comparison with larger phototrophs (Raven, 1998; Marañón, 2015), which helps them thrive in harsher conditions (Agawin *et al.*, 2000; Worden *et al.*, 2004). The high abundance and diversification of heterotrophs (HET) within pico-plankton might be explained by the omnipresence of prey for small bacterivorous phagotrophs (Logares *et al.*, 2012; Pernice *et al.*, 2015). The persistence of parasites (PARA) could be explained by the release of numerous small-sized spores from hosts (Park *et al.*, 2004; Guillou *et al.*, 2008), which could also transform into a dormant stage, remaining resistant over time (Gleason *et al.*, 2014; Scholz *et al.*, 2016). As a consequence of these distinct strategies in the pico-nan-plankton, the relative contribution of functional groups was significantly more equitable and less variable than in micro-plankton (Fig. 6a). This pattern followed that of genetic diversity, in which pico-nano-planktonic taxonomic communities were characterized by significantly higher and less variable OTU richness and equitability than in micro-plankton (Fig. 6). The stability of phylogenetic richness as well as a higher OTU richness in samples of the smallest size-fractions have already been highlighted in coastal ecosystems (Massana *et al.*, 2004; Romari and Vaultot, 2004; Logares *et al.*,

2014; de Vargas *et al.*, 2015). Here, we demonstrate that this stability is also represented in terms of functional groups diversity. Overall, our results show distinct ecological patterns of marine protists across size-fractions, both in terms of taxonomy and functions. Protistan communities in the micro-plankton appear to be mostly dominated by assemblages with low taxonomic and functional groups diversity, while protistan communities in the nano-pico-plankton consist of more diversified assemblages in which the dominance of the different taxonomic and functional units fluctuates over space and time. This corroborates the idea that the nano- and pico-plankton are part of a more stable ecological background on which larger protists and metazoans might develop (Fenchel and Finlay, 2004; Massana, 2011), in agreement with broader ecological theories on the distribution of size and species richness (Hutchinson and MacArthur, 1959).

Coupling between functional roles and taxonomy among marine protists of coastal ecosystems

The environmental variables used in this study to characterize the sampled ecosystems enabled the identification of classic environmental gradients found in coastal environments. Heterotrophs coincided well with off-shore and winter conditions, which classically highlighted the greater influence of microbial loop processes in these environments (Azam *et al.*, 1983; Legendre and Rassoulzadegan, 1995). Phototrophs were more abundant in nutrient-depleted samples, consistent with the typical cycle of phytoplankton uptake (Cloern, 1996; Cloern and Jassby, 2008). However, our functional groups corresponded to ecological strategies that could be dependent on other environmental variables not measured in this study. For instance, water-mixing and grazing pressure could favor the floater, non-swimming, phototrophs (FLAT) and colonial phototrophs (CAT) (Landeira *et al.*, 2014). Oxygen concentrations, abundances of prokaryotic preys, particulate and dissolved organic matter concentrations, have also been shown to correlate with the large-scale

distribution of heterotrophic protists (Pernice *et al.*, 2015) and myco-plankton dynamics (Taylor and Cunliffe, 2016).

Our study showed that, in protistan coastal communities, changes in taxonomic composition corresponded to variations in the relative proportion of functional groups. For marine prokaryotes, the results are more contrasted. Some studies report a decoupling between the metabolic potential and taxonomic composition, which has been attributed to a high functional redundancy (Sunagawa *et al.*, 2015; Louca *et al.*, 2016); reversely, other authors showed strong covariations between the taxonomy of prokaryotes and their metagenome across distinct samples, challenging functional redundancy (Galand *et al.*, 2018). These divergences stress the need for a common approach, a common definition, and tests across distinct scales to answer the debate over microbial functional redundancy. Nevertheless, protists and prokaryotes could also contrast in functional redundancy due to distinct evolutionary and selection processes. Different prokaryotes developed multiple cooperating enzymes, which evolved in parallel with biogeochemical cycles (Falkowski *et al.*, 2008). As a consequence, the functional roles of prokaryotes in their environment are explained at the molecular and metabolic level. Furthermore, among prokaryotes, horizontal gene transfer is a major process of evolution (Cohan, 2002). Two distinct prokaryotic phyla can thus exchange genes (Koonin *et al.*, 2001) and their functional roles might be coded by few genes easily exchangeable by horizontal transfer. This process is thought to create a community-wide functional redundancy among prokaryotes, which could explain the decoupling between their functional and taxonomic diversity (Allison and Martiny, 2008; Falkowski *et al.*, 2008; Louca *et al.*, 2016). Micro-eukaryotes demonstrate their functional role at the cellular level, having developed various behaviors, specialized morphology, interactions, adaptations and strategies (Massana and Logares, 2013). This functional complexity is coded by multi-gene patterns (Burns *et al.*, 2018), which are more difficult to exchange by horizontal transfer across species (Massana and Logares, 2013; Keeling and del Campo, 2017). Protist functional diversity can therefore

be explained by specific morphology and trophic behaviors that evolved separately across micro-eukaryotes, justifying the close link between the taxonomic and functional diversity of their environmental communities.

Conclusions

The description of morphological characteristics and feeding behaviors of protists was enough informative to describe the functional community patterns of coastal micro-eukaryotic assemblages. A close coupling between the taxonomic and functional diversity of coastal protistan communities was evidenced in this study. This contrasts with prokaryotic oceanic communities in which functional roles are mostly played at the molecular level and easily exchangeable, blurring the borders between taxonomy and functions. Each species of protist has developed its own particular blend of morphological and behavioral features, which constitute rarely exchangeable functional roles, favoring a strong coupling between taxonomy and function. We also showed distinct functional diversity patterns between large and small protistan communities. Thus, micro-plankton seemed more prone to be dominated by one or few functional groups while the smallest size-fraction exhibited the co-existence of various phototrophs and heterotrophs in a sample. This hypothesis needs to be tested in larger experimental frameworks and beyond coastal ecosystems. Our innovative analysis, based on information gathered from the literature, is also perfectible. In fact, the dearth of crucial information on protists, especially their physiological traits, is one of the results of our analysis. This emphasizes the constant need for observations to further understand the structure of protistan communities, both *in situ* and *in vitro* and comprising e.g. microscopy, taxonomic abundances and ecological fluxes. Despite sequencing limitations, studying the genomic basis of functional roles can also represent an interesting prospect in the functional ecology of protists. As more protistan genomes become

accessible, it will become more feasible to predict phenotypic information, ecological strategies and functional roles based on DNA sequences.

Experimental Procedures

Sampling strategy

A total of 277 water samples was collected with temporal and/or spatial strategies across coastal ecosystems of France and Senegal (Fig. 1). Samples were collected at the surface water with comparable procedures (0-5 m depths). For some sites, additional samples were collected at the depth of the Deep Chlorophyll Maximum (DCM) and at the mesopelagic level, identified by CTD profiles. Water replicates (one or two) were sampled during most cruises (Table 1).

Seawater samples were collected in Niskin bottles and gradually filtered onto polycarbonate membrane filters, 47 mm in diameter with a pore size of 20 (or 10), 3 and 0.2 μm . Particles of the first two size-fractions ($>3 \mu\text{m}$) were separated by means of a peristaltic water pump and Swinnex filter holders. For the last size-fraction, 0.22 μm polyethersulfone Sterivex filters were used for the samples of MB and RA at the end of the pumping circuit. For other cruises, 0.5 to 1 L of the residual filtrate from the higher size-classes was filtered onto 0.2 μm filters. This size-fractionated sampling yielded a total of 1145 filters enabling the study of plankton size-classes. For convenience, we define here our size-fractions as micro- (>10 or $20 \mu\text{m}$), nano- (3 - 20 or $10 \mu\text{m}$) and pico- (0.22 - $3 \mu\text{m}$) plankton. Sampled water was filtered until filter clogging, which yielded a variable filtered water volume ranging from 0.5 to 10 L across the different samples. After filtration, filters were immediately frozen in liquid nitrogen and then stored at -80°C until genetic procedures. The lysis buffer was added to the RA filters before freezing. To estimate the environmental characteristic sampled, temperature and salinity were measured by CTD, and nutrient concentrations (NO_3^- , NO_2^- , PO_4^{3-} , NH_4^+ and $\text{Si}(\text{OH})_4^-$)

were measured with a Seal Analytical AA3 HR automatic analyzer following the procedures described by Aminot and K  rouel (2007)

Genetic procedures

A metabarcoding approach was adopted to characterize the taxonomic diversity of the sampled communities. The hyper-variable V4 domain of the 18S-rDNA region was chosen as a barcode for its conservative character within the eukaryotic microbial community and its relatively long length (230-520 bp; Nickrent and Sargent, 1991) which allows a relatively good genetic distinction of marine protists (Stoeck *et al.*, 2010; Behnke *et al.*, 2011; Dunthorn *et al.*, 2012). Genetic methods were the same for all cruises, except for the RA dataset (Roscoff-Astan). Genomic DNA was extracted using the DNA extraction kit NucleoSpin Plant II (Macherey-Nagel, Hoerd, France). DNA from RA filters was extracted with two buffers, one lysis buffer containing lysozyme (45 min at 37  C) and one composed of proteinase K and SDS (1 h at 55  C). The extract was treated with phenol:chloroform:isoamyl alcohol and then processed with the NucleoSpin extraction kit. Blank extractions with nuclease-free water were carried out as negative controls for contamination during the process. The quality and concentration of extracted DNA was measured using a BioTek FLx800 spectrofluorophotometer and a Quant-iT PicoGreen ds DNA quantification kit (Invitrogen, Carlsbad, CA), respectively, following the manufacturer's instructions. The final DNA concentration of all extracts was normalised to 5-10 ng/  L. PCR was then run with V4 markers assembled with the GeT-PlaGe adapters of the sequencing platform Genotoul (<http://get.genotoul.fr/>; Forward: V4f_PlaGe 5'CTT-TCC-CTA-CAC-GAC-GCT-CTT-CCG-ATC-TCC-AGC-A(C/G)C-(C/T)GC-GGT-AAT-TCC'3, Reverse: 553 V4f_PlaGe 5'GGA-GTT-CAG-ACG-TGT-GCT-CTT-CCG-ATC-TAC-TTT-CGT-TCT-TGA-T(C/T)(A/G)-A'3) and a taq polymerase (Phusion High-Fidelity PCR Master Mix with GC Buffer). PCR amplification was carried out three times for each DNA extract (representing a

single filter). The amplification protocol consisted of a denaturation step at 98°C for 30 s, followed by two sets of cycles 1) 12 x [98°C (10 s), 53°C (30 s), 74°C (30 s)] and 2) 18 x [98°C (10 s), 48°C (30 s), 74°C (30 s)]. The cycles were followed by a final elongation at 72°C for 10 min. Amplification results were verified by gel electrophoresis; triplicate reactions were pooled and purified using NucleoSpin Gel and PCR Clean-up (Macherey-Nagel, Hoerdt, France). Purified products were diluted to obtain equimolar concentrations before library construction at Genotoul for Illumina MiSeq (2x250 bp) sequencing. Six libraries were constructed, keeping separate the samples from different cruises and locations. Sequence data are available at sextant.ifremer.fr/record/16bc16ef-588a-47e2-803e-03b4acb85dca/.

Sequence data cleaning, filtering and clustering into OTUs and taxa

Sequenced data were submitted to quality checking by built-in modules of the USEARCH program (Edgar *et al.*, 2011) comprising 1) removal of reads with biased nucleotides (according to the Phred score < 1%), 2) elimination of reads with incomplete or wrong primer sequences, 3) chimera removal and 4) merging of paired-end reads. In order to eliminate PCR errors and read-sample cross-contamination, a strict data cut-off was applied to the cleaned dataset. Singletons and sequences present in fewer than two samples and with a total of fewer than three reads over the whole dataset were removed (de Vargas *et al.*, 2015). All sequences were assigned with a taxonomic reference using the V4 reference database PR² (Guillou *et al.*, 2013). All sequences with a percentage of identity to the reference database $\leq 80\%$ were removed (Stoeck *et al.*, 2010; de Vargas *et al.*, 2015; Mahé *et al.*, 2017) considering that values under this threshold highlighted sequencing errors. Reads annotated to “Metazoa” and to multi-cellular plants were also removed from the database, whereas annotated fungi were kept. Metabarcodes were then clustered into Operational Taxonomic Units (OTUs) by the agglomerative, unsupervised single-linkage clustering algorithm Swarm 2 (Mahé *et al.*, 2014, 2015),

with a default clustering threshold of $d = 1$ (Mahé *et al.*, 2015). This clustering led to the creation of 111 089 OTUs cumulating into 3.5×10^6 reads. Each of these OTUs was given the taxonomic reference of its most abundant metabarcode, resulting in 2 007 distinct taxonomic references. Sampling quality was evaluated by rarefaction curves (reads vs. OTU numbers) calculated with the *rarecurve()* function of the R package “vegan” (Oksanen *et al.*, 2016; Fig. S2).

Trait approach and functional groups

The analytic steps of our newly developed method are detailed in Fig. 7. To carry out a trait annotation on our metabarcoding dataset we used the taxonomy, i.e. the 2007 taxonomic references given by PR² (Guillou *et al.*, 2013), of our OTUs. The trait annotation was carried out by searching the biological description of each taxonomic references in the literature (Ramond *et al.*, 2018). Based on this work we built a trait table composed of 1669 taxonomic references and 13 traits.

To identify relevant functional groups among our taxonomic references we developed a statistical approach inspired by a previous publication (Maire *et al.*, 2015) . The Gower distance was computed on the biological trait table to calculate a distance matrix of the selected taxonomic references (1669) according to their traits (13); this distance was then analyzed by Principal Coordinate Analysis (PCoA; Fig. S3). The Gower distance was used here because it can deal with different sorts of traits (e.g. numerical and categorical traits; Legendre and Legendre, 2012; Maire *et al.*, 2015). The following PCoA (Fig. S3) creates a linear mapping of the Gower distance onto an ordination space composed of multiple Euclidean dimensions on which taxonomic references are represented by scores (Ramette, 2007). The number of relevant Euclidean dimensions, here 5, was selected according to a neutral statistical method (Maire *et al.*, 2015). The ordination represents the 1669 taxonomic references distributed according to the similarity in their trait composition (Fig. S3). This similarity is structured on the basis of trade-offs (i.e. constraints in trait selection), e.g. colony-

Accepted Article

formation or heavy cell-covers are useful for avoiding predators but limit phagotrophic possibilities, representing a trade-off between cell cover and trophic strategy. To investigate the trade-offs among our traits, we studied the correlation between the traits and the PCoA dimensions with the Spearman Rank test (Fig. S4), the dimensions that correlated well with distinct traits represented trait trade-offs. Finally, we used the scores of each taxonomic reference on the PCoA dimensions to build functional groups. The Simple Structure Index (SSI) method of k-mean was used as an impartial criterion to select the best partitioning of our taxonomic references into functional groups (Borcard *et al.*, 2011; Laliberté *et al.*, 2015 ; Fig. S5). To build functional groups that represented cohesive ecological strategies, the clustering method was run only on the dimensions that highlighted trade-offs (Fig. S5). As an example, when including the dimension correlating only with the trait of resting stage (showing no trade-off with other traits, see PCoA 3 in Fig. S4) the number of functional group was duplicated according to the presence of a resting stage (e.g. heterotrophs with a resting stage and heterotrophs without a resting stage). We preferred to build larger functional groups, based on trade-offs, that comprised intra-group trait diversity. To characterize our functional groups, we studied the dominant modalities of traits in each functional group (Fig. S6-S11).

By re-associating the taxonomic references to their respective OTUs we could then annotate our OTUs with traits and a functional group. Functionally annotated OTUs represented a reduced portion of the complete dataset (52 180/ 111 089 OTUs). Correlations between both datasets were calculated with the Mantel test (Mantel, 1967). The method consists of comparing two dissimilarity matrices sharing the same samples but based on differing variables. Here, the dissimilarity matrices of the complete and the reduced datasets were calculated with the Bray-Curtis distance, which is suitable for datasets affected by the “double-zero problem” (Ramette, 2007; Legendre and Legendre, 2012; Buttigieg and Ramette, 2014). Mantel tests were calculated with 9,999 permutations. OTUs richness and the Shannon index were also compared between the complete and reduced dataset with the

Pearson correlation coefficient (Fig. S12). To study the phylogeny of our functional groups we used the taxonomic references of our OTUs and build a simple phylogenetic tree of eukaryotes, we then studied the proportion of the distinct phyla on the total number of OTUs inside each functional group. Finally, the read abundance of the OTUs inside each functional group was cumulated to calculate a relative abundance of the functional groups in each sample. In this way, a functional community table was built.

Statistical Analyses

A Principal Component Analysis (PCA), computed on temperature, salinity and nutrient measurements ($\text{NO}_x = \text{NO}_3^- + \text{NO}_2^-$, PO_4^{3-} and $\text{Si}(\text{OH})_4^-$) was performed (Fig. S1). This analysis enabled the characterization of environmental gradients and highlighted differences between the sampling cruises. Unfortunately, these environmental characteristics were not available for some samples, which were thus absent from the PCA. In particular, the Senegalese samples lacked temperature and salinity measurements, while the whole set of environmental variables was completely absent for the 2015 samples of the PI and PH cruises.

Relationships between environmental variables, taxonomic and functional diversity in pico-, nano- and micro-plankton communities were tested with the RV statistical coefficient of co-inertia, a multivariate generalization of the Pearson correlation coefficient (Borcard *et al.*, 2011; Legendre and Legendre, 2012; Husson *et al.*, 2018). The test was run on the taxonomic community table, the functional community table and the environmental dataset (same samples that were used in the PCA) of micro-, nano- and pico-plankton samples. The samples without environmental measurements (the same as in the PCA analysis) were discarded from this analysis and tests were conducted with 256, 254 and 271 samples for micro-, nano- and pico-plankton, respectively. For deeper investigations, the same dataset was used and a Non-Metric Multidimensional Scaling Analysis (NMDS) was calculated

on the complete taxonomic table (with the Bray-Curtis distance) to ordinate samples of each size-fraction on the basis of their OTU composition. Environmental values were fitted as explanatory vectors on the two axes of the NMDS with the *envfit()* function of the “vegan” package (Oksanen *et al.*, 2016). Clusters of samples were constructed on the basis of the ssi method of k-mean partitioning. The relative proportion of each functional group was calculated within each cluster of each size-fraction. The Adonis test, a non-parametric multivariate analysis of variance (Oksanen *et al.*, 2016; with 9,999 permutations), was used to determine if the functional groups showed a distinctive distribution across the clusters within each size-fraction.

Finally, the Shannon index (Pielou, 1966) was calculated on the basis of the functional and taxonomic community tables, as well as OTU richness. These metrics were compared across the size fractions with a Kruskal-Wallis test (a non-parametric one-way ANOVA test).

All statistical and data analyses were performed with R software (R Core Team Development, 2015); in particular, the community analyses were carried out with the “vegan” and “FactoMineR” packages (Oksanen *et al.*, 2016; Husson *et al.*, 2018).

Acknowledgments

This work was financed by the French government under the programme ‘Investissements d’Avenir’, by the projects of the initiative ECosphere Continentale et COtière (EC2CO) of the Institut National des Sciences de l’Univers/Centre National de la Recherche Scientifique (INSU/CNRS): POHEM (2016). The authors declare no conflict of interest. This research was carried out within the framework of Pierre Ramond’s PhD, co-funded by Ifremer and Region Bretagne (Allocation de REcherche Doctorale (ARED) fellowship). We thank all members of the M2BIPAT and DAOULEX consortia, as well as the cruise and Ifremer members who contributed to the collection of the PELGAS, DYNAPSE and PHYTEC/IPARO samples. We are also much obliged to the Group Plankton members who were

involved in and contributed to the SOMLIT-Astan time series off Roscoff, Western English Channel (UMR7144 - FR2424 - Station Biologique de Roscoff). Finally, we wish to thank personally Dr. Eric Machu (Institut de recherche pour le développement, IRD) for collecting and providing the Senegalese samples.

Accepted Article

References

- Agawin, N.S.R., Duarte, C.M., and Agustí, S. (2000) Nutrient and temperature control of the contribution of picoplankton to phytoplankton biomass and production. *Limnol. Oceanogr.* **45**: 591–600.
- Aleoshin, V. V., Mylnikov, A.P., Mirzaeva, G.S., Mikhailov, K. V., and Karpov, S.A. (2016) Heterokont Predator *Develorapax marinus* gen. et sp. nov. – A Model of the Ochrophyte Ancestor. *Front. Microbiol.* **7**: 1–14.
- Allison, S.D. and Martiny, J.B.H. (2008) Colloquium paper: resistance, resilience, and redundancy in microbial communities. *Proc. Natl. Acad. Sci. U. S. A.* **105 Suppl**: 11512–9.
- Alves-De-Souza, C., Gonzalez, M.T., and Iriarte, J.L. (2008) Functional groups in marine phytoplankton assemblages dominated by diatoms in fjords of southern Chile. *J. Plankton Res.* **30**: 1233–1243.
- Aminot, A. and Kérouel, R. (2007) Dosage automatique des nutriments dans les eaux marines : méthodes en flux continu, Ifremer (ed).
- Azam, F., Fenchel, T., Field, J.G., Gray, J.C., Meyer-Reil, L.A., and Thingstad, F. (1983) The ecological role of water-column microbes in the sea. *Mar. Ecol. Prog. Ser.* **10**: 257–264.
- Barton, A.D., Dutkiewicz, S., Flierl, G., Bragg, J., and Follows, M.J. (2010) Patterns of Diversity in Marine Phytoplankton. *Science (80-)*. **327**: 1509–1512.
- Beaugrand, G. and Kirby, R.R. (2018) How Do Marine Pelagic Species Respond to Climate Change? Theories and Observations. *Ann. Rev. Mar. Sci.* **10**: 169–197.
- Behnke, A., Engel, M., Christen, R., Nebel, M., Klein, R.R., and Stoeck, T. (2011) Depicting more accurate pictures of protistan community complexity using pyrosequencing of hypervariable SSU rRNA gene regions. *Environ. Microbiol.* **13**: 340–349.
- Berney, C., Romac, S., Mahé, F., Santini, S., Siano, R., and Bass, D. (2013) Vampires in the oceans:

predatory cercozoan amoebae in marine habitats. *ISME J.* **7**: 1–13.

Le Bescot, N., Mahé, F., Audic, S., Dimier, C., Garet, M.-J., Poulain, J., et al. (2015) Global patterns of pelagic dinoflagellate diversity across protist size classes unveiled by metabarcoding. *Environ. Microbiol.* **18**: 609–626.

Borcard, D., Gillet, F., and Legendre, P. (2011) Numerical Ecology with R Use R! Springer (ed).

Bruggeman, J. (2011) A Phylogenetic Approach To the Estimation of Phytoplankton Traits. *J. Phycol.* **65**: 52–65.

Burki, F., Kudryavtsev, A., Matz, M. V, Aglyamova, G. V, Bulman, S., Fiers, M., et al. (2010) Evolution of Rhizaria: new insights from phylogenomic analysis of uncultivated protists. *BMC Evol. Biol.* **10**: 377.

Burns, J.A., Pittis, A.A., and Kim, E. (2018) Gene-based predictive models of trophic modes suggest Asgard archaea are not phagocytotic. *Nat. Ecol. Evol.*

Buttigieg, P.L. and Ramette, A. (2014) A guide to statistical analysis in microbial ecology: A community-focused, living review of multivariate data analyses. *FEMS Microbiol. Ecol.* **90**: 543–550.

Caron, D.A., Countway, P.D., Jones, A.C., Kim, D.Y., and Schnetzer, A. (2012) Marine Protistan Diversity. *Ann. Rev. Mar. Sci.* **4**: 467–493.

Christaki, U., Kormas, K.A., Genitsaris, S., Georges, C., Sime-Ngando, T., Viscogliosi, E., and Monchy, S. (2014) Winter-Summer Succession of Unicellular Eukaryotes in a Meso-eutrophic Coastal System. *Microb. Ecol.* **67**: 13–23.

Cloern, J.E. (1996) Phytoplankton bloom dynamics in coastal ecosystems: A review with some general lessons from sustained investigation of San Francisco Bay, California. *Rev. Geophys.* **34**: 127.

Cloern, J.E. and Jassby, A.D. (2008) Complex seasonal patterns of primary producers at the land-sea

interface. *Ecol. Lett.* **11**: 1294–1303.

Cohan, F.M. (2002) What are Bacterial Species? *Annu. Rev. Microbiol.* **56**: 457–487.

Coles, V.J., Stukel, M.R., Brooks, M.T., Burd, A., Crump, B.C., Moran, M.A., et al. (2017) Ocean biogeochemistry modeled with emergent trait-based genomics. *Science (80-.)*. **1154**: 1–26.

Diaz, S., Purvis, A., Cornelissen, J.H.C., Mace, G.M., Donoghue, M.J., Ewers, R.M., et al. (2013) Functional traits, the phylogeny of function, and ecosystem service vulnerability. *Ecol. Evol.* **3**: 2958–2975.

Dunthorn, M., Klier, J., Bunge, J., and Stoeck, T. (2012) Comparing the hyper-variable V4 and V9 regions of the small subunit rDNA for assessment of ciliate environmental diversity. *J. Eukaryot. Microbiol.* **59**: 185–187.

Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., and Knight, R. (2011) UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27**: 2194–2200.

Edwards, K.F., Thomas, M.K., Klausmeier, C. a., and Litchman, E. (2012) Allometric scaling and taxonomic variation in nutrient utilization traits and maximum growth rate of phytoplankton. *Limnol. Oceanogr.* **57**: 554–566.

Falkowski, P.G., Fenchel, T., and Delong, E.F. (2008) The Microbial Engines That Drive Earth 's Biogeochemical Cycles. *Science (80-.)*. **320**: 1034–1039.

Fenchel, T. (1982a) Ecology of Heterotrophic Microflagellates. I. Some Important Forms and Their Functional Morphology. *Mar. Ecol. Prog. Ser.* **8**: 211–223.

Fenchel, T. (1982b) Ecology of Heterotrophic Microflagellates. II. Bioenergetics and Growth. *Mar. Ecol. Prog. Ser.* **8**: 225–231.

Fenchel, T. (1980) Relation between particle size selection and clearance in suspension feeding ciliates. *Limnol. Oceanogr.* **25**: 733–738.

Fenchel, T. and Finlay, B.J. (2004) The Ubiquity of Small Species: Patterns of Local and Global

Diversity. *Bioscience* **54**: 777.

Galand, P.E., Pereira, O., Hochart, C., Auguet, J.C., and Debroas, D. (2018) A strong link between marine microbial community composition and function challenges the idea of functional redundancy. *ISME J.* 1–9.

Genitsaris, S., Monchy, S., Viscogliosi, E., Sime-Ngando, T., Ferreira, S., and Christaki, U. (2015) Seasonal variations of marine protist community structure based on taxon-specific traits using the eastern English Channel as a model coastal system. *FEMS Microbiol. Ecol.* **91**: fiv034.

Gleason, F.H., Lilje, O., Marano, A. V., Sime-Ngando, T., Sullivan, B.K., Kirchmair, M., and Neuhauser, S. (2014) Ecological functions of zoospore hyperparasites. *Front. Microbiol.* **5**: 1–10.

Gómez, F., López-García, P., Nowaczyk, A., and Moreira, D. (2009) The crustacean parasites *Ellobiopsis* Caullery, 1910 and *Thalassomyces* Niezabitowski, 1913 form a monophyletic divergent clade within the Alveolata. *Syst. Parasitol.* **74**: 65–74.

Grover, J.P. (1989) Influence of cell shape and size on algal competitive ability. *J. Phycol.* **25**: 402–405.

Guidi, L., Chaffron, S., Bittner, L., Eveillard, D., Larhlimi, A., Roux, S., et al. (2016) Plankton networks driving carbon export in the oligotrophic ocean. *Nature*.

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.

Guillou, L., Viprey, M., Chambouvet, A., Welsh, R.M., Kirkham, A.R., Massana, R., et al. (2008) Widespread occurrence and genetic diversity of marine parasitoids belonging to Syndiniales (Alveolata). *Environ. Microbiol.* **10**: 3349–3365.

Haggerty, J.M. and Dinsdale, E.A. (2017) Distinct biogeographical patterns of marine bacterial

taxonomy and functional genes. *Glob. Ecol. Biogeogr.* **26**: 177–190.

Hamm, C.E. (2005) The Evolution of Advanced Mechanical Defenses and Potential Technological Applications of Diatom Shells. *J. Nanosci. Nanotechnol.* **5**: 108–119.

Hansen, P.J., Bjørnsen, P.K., and Hansen, B.W. (1997) Zooplankton grazing and growth: Scaling within the 2–2,000- μ m body size range. *Limnol. Oceanogr.* **42**: 687–704.

Howe, A.T., Bass, D., Scoble, J.M., Lewis, R., Vickerman, K., Arndt, H., and Cavalier-Smith, T. (2011) Novel Cultured Protists Identify Deep-branching Environmental DNA Clades of Cercozoa: New Genera Tremula, Micrometopion, Minimassisteria, Nudifila, Peregrinia. *Protist* **162**: 332–372.

Hu, S., Campbell, V., Connell, P., Gellen, A.G., Liu, Z., Terrado, R., and Caron, D.A. (2016) Protistan diversity and activity inferred from RNA and DNA at a coastal ocean site in the eastern North Pacific. *FEMS Microbiol. Ecol.* 1–39.

Husson, A.F., Josse, J., Le, S., Mazet, J., and Husson, M.F. (2018) Package “FactoMineR.” 97.

Hutchinson, G.E. and MacArthur, R.H. (1959) A Theoretical Ecological Model of Size Distributions Among Species of Animals. *Am. Nat.* **93**: 117–125.

Jephcott, T.G., Alves-de-Souza, C., Gleason, F.H., van Ogtrop, F.F., Sime-Ngando, T., Karpov, S.A., and Guillou, L. (2016) Ecological impacts of parasitic chytrids, syndiniales and perkinsids on populations of marine photosynthetic dinoflagellates. *Fungal Ecol.* **19**: 47–58.

Keeling, P.J., Burki, F., Wilcox, H.M., Allam, B., Allen, E.E., Amaral-Zettler, L.A., et al. (2014) The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): Illuminating the Functional Diversity of Eukaryotic Life in the Oceans through Transcriptome Sequencing. *PLoS Biol.* **12**:

Keeling, P.J. and del Campo, J. (2017) Marine Protists Are Not Just Big Bacteria. *Curr. Biol.* **27**: R541–R549.

- Kjørboe, T. (2011) How zooplankton feed: Mechanisms, traits and trade-offs. *Biol. Rev.* **86**: 311–339.
- Koonin, E. V., Makarova, K.S., and Aravind, L. (2001) Horizontal Gene Transfer in Prokaryotes: Quantification and Classification. *Annu. Rev. Microbiol.* **55**: 709–742.
- Kruk, C., Peeters, E.T.H.M., Van Nes, E.H., Huszar, V.L.M., Costa, L.S., and Scheffer, M. (2011) Phytoplankton community composition can be predicted best in terms of morphological groups. *Limnol. Oceanogr.* **56**: 110–118.
- Laliberté, E., Legendre, P., and Bill Shipley (2015) Measuring functional diversity (FD) from multiple traits, and other tools for functional ecology. 1–28.
- Lallias, D., Hiddink, J.G., Fonseca, V.G., Gaspar, J.M., Sung, W., Neill, S.P., et al. (2015) Environmental metabarcoding reveals heterogeneous drivers of microbial eukaryote diversity in contrasting estuarine ecosystems. *ISME J.* **9**: 1208–21.
- Landeira, J.M., Ferron, B., Lunven, M., Morin, P., Marie, L., and Sourisseau, M. (2014) Biophysical interactions control the size and abundance of large phytoplankton chains at the ushant tidal front. *PLoS One* **9**..
- Lange, K., Townsend, C.R., and Matthaei, C.D. (2015) A trait-based framework for stream algal communities. *Ecol. Evol.*
- Lavorel, S. and Garnier, E. (2002) Predicting changes in community composition and ecosystem functioning from plant traits: revisiting the Holy Grail. *Funct. Ecol.* **16**: 545–556.
- Legendre, L. and Rassoulzadegan, F. (1995) Plankton and nutrient dynamics in marine waters. *Ophelia* **41**: 153–172.
- Legendre, P. and Legendre, L. (2012) Numerical Ecology. Third English Edition Third Engl. Developments in Environmental Modelling, 24 and Numerical (eds) Elsevier, Amsterdam (The Netherlands).
- Lin, Y.C., Campbell, T., Chung, C.C., Gong, G.C., Chiang, K.P., and Worden, A.Z. (2012)

Distribution patterns and phylogeny of marine stramenopiles in the North Pacific Ocean. *Appl. Environ. Microbiol.* **78**: 3387–3399.

Litchman, E. and Klausmeier, C.A. (2008) Trait-Based Community Ecology of Phytoplankton. *Annu. Rev. Ecol. Evol. Syst.* **39**: 615–639.

Litchman, E., Ohman, M.D., and Kiørboe, T. (2013) Trait-based approaches to zooplankton communities. *J. Plankton Res.* **35**: 473–484.

Logares, R., Audic, S., Bass, D., Bittner, L., Boutte, C., Christen, R., et al. (2014) Patterns of rare and abundant marine microbial eukaryotes. *Curr. Biol.* **24**: 813–21.

Logares, R., Audic, S., Santini, S., Pernice, M.C., de Vargas, C., and Massana, R. (2012) Diversity patterns and activity of uncultured marine heterotrophic flagellates unveiled with pyrosequencing. *ISME J.*

Louca, S., Polz, M.F., Mazel, F., Albright, M.B.N., Huber, J.A., O'Connor, M.I., et al. (2018) Function and functional redundancy in microbial systems. *Nat. Ecol. Evol.*

Louca, S., Wegener Parfrey, L., and Doebeli, M. (2016) Decoupling function and taxonomy in the global ocean microbiome. *Science (80-.)*. **353**: 1272–1277.

Mahé, F., Rognes, T., Quince, C., de Vargas, C., and Dunthorn, M. (2014) Swarm: robust and fast clustering method for amplicon-based studies. *PeerJ* **2**: e593.

Mahé, F., Rognes, T., Quince, C., de Vargas, C., and Dunthorn, M. (2015) Swarm v2: highly-scalable and high-resolution amplicon clustering. *PeerJ* **1420**: 1–20.

Mahé, F., de Vargas, C., Bass, D., Czech, L., Stamatakis, A., Lara, E., et al. (2017) Parasites dominate hyperdiverse soil protist communities in Neotropical rainforests. *Nat. Ecol. Evol.* **1**: 91.

Maire, E., Grenouillet, G., Brosse, S., and Villéger, S. (2015) How many dimensions are needed to accurately assess functional diversity? A pragmatic approach for assessing the quality of functional spaces. *Glob. Ecol. Biogeogr.* **24**: 728–740.

- Mantel, N. (1967) The Detection of Disease Clustering and a Generalized Regression Approach. *Cancer Res.* **27**: 209–220.
- Marañón, E. (2015) Cell Size as a Key Determinant of Phytoplankton Metabolism and Community Structure. *Ann. Rev. Mar. Sci.* **7**: 241–264.
- Marcus, N.H. and Boero, F. (1998) Minireview: The importance of benthic-pelagic coupling and the forgotten role of life cycles in coastal aquatic systems. *Limnol. Oceanogr.* **43**: 763–768.
- Margalef, R. (1978) Life-Forms of Phytoplankton As Survival Alternatives in An Unstable Environment. *Oceanologia* **1**: 493–509.
- Massana, R. (2011) Eukaryotic Picoplankton in Surface Oceans. *Annu. Rev. Microbiol.* **65**: 91–110.
- Massana, R., Balagué, V., Guillou, L., and Pedrós-Alió, C. (2004) Picoeukaryotic diversity in an oligotrophic coastal site studied by molecular and culturing approaches. *FEMS Microbiol. Ecol.* **50**: 231–243.
- Massana, R., Gobet, A., Audic, S., Bass, D., Bittner, L., Boutte, C., et al. (2015) Marine protist diversity in European coastal waters and sediments as revealed by high-throughput sequencing. *Environ. Microbiol.* **17**: 4035–4049.
- Massana, R. and Logares, R. (2013) Eukaryotic versus prokaryotic marine picoplankton ecology. *Environ. Microbiol.* **15**: 1254–1261.
- Massana, R., Unrein, F., Rodríguez-Martínez, R., Forn, I., Lefort, T., Pinhassi, J., and Not, F. (2009) Grazing rates and functional diversity of uncultured heterotrophic flagellates. *ISME J.* **3**: 588–596.
- Mitra, A., Flynn, K.J., Tillmann, U., Raven, J.A., Caron, D., Stoecker, D.K., et al. (2016) Defining Planktonic Protist Functional Groups on Mechanisms for Energy and Nutrient Acquisition; Incorporation of Diverse Mixotrophic Strategies. *Protist.*
- Mouillot, D., Villeger, S., Parravicini, V., Kulbicki, M., Arias-Gonzalez, J.E., Bender, M., et al.

- (2014) Functional over-redundancy and high functional vulnerability in global fish faunas on tropical reefs. *Proc. Natl. Acad. Sci.* **111**: 13757–13762.
- Nickrent, D.L. and Sargent, M.L. (1991) An overview of the secondary structure of the V4 region of eukaryotic small-subunit ribosomal RNA. *Nucleic Acids Res.* **19**: 227–235.
- Nielsen, S.L. and Sandjensen, K. (1990) Allometric scaling of maximal photosynthetic growth-rate to surface volume ratio. *Limnol. Oceanogr.* **35**: 177–181.
- Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., et al. (2016) vegan: Community Ecology Package.
- Pahlow, M., Riebesell, U., and Wolf-Gladrow, D. a. (1997) Impact of cell shape and chain formation on nutrient acquisition by marine diatoms. *Limnol. Oceanogr.* **42**: 1660–1672.
- Park, M.G., Yih, W., and Coats, D.W. (2004) Parasites and phytoplankton, with special emphasis on dinoflagellate infections. *J. Eukaryot. Microbiol.* **51**: 145–155.
- Pearman, J.K., Ellis, J., Irigoien, X., Sarma, Y.V.B., Jones, B.H., and Carvalho, S. (2017) Microbial planktonic communities in the Red Sea: high levels of spatial and temporal variability shaped by nutrient availability and turbulence. *Sci. Rep.* **7**: 6611.
- Pernice, M.C., Forn, I., Gomes, A., Lara, E., Alonso-Sáez, L., Arrieta, J.M., et al. (2015) Global abundance of planktonic heterotrophic protists in the deep ocean. *ISME J.* **9**: 782–792.
- Pernice, M.C., Giner, C.R., Logares, R., Perera-Bel, J., Acinas, S.G., Duarte, C.M., et al. (2016) Large variability of bathypelagic microbial eukaryotic communities across the world's oceans. *ISME J.* **10**: 945–958.
- Pielou, E.C. (1966) Shannon's Formula as a Measure of Specific Diversity: Its Use and Misuse. *Am. Nat.* **100**: 463–465.
- Ploug, H., Stolte, W., and Jørgensen, B.B. (1999) Diffusive boundary layers of the colony-forming plankton alga, *Phaeocystis* sp. --implications for nutrient uptake and cellular growth. *Limnol.*

Oceanogr. **44**: 1959–1967.

R Core Team Development (2015) R: A language and environment for statistical computing.

Ramette, A. (2007) Multivariate analyses in microbial ecology. *FEMS Microbiol. Ecol.* **62**: 142–160.

Ramond, P., Siano, R., and Sourisseau, M. (2018) Functional traits of marine protists. *SEANOE*.

<https://doi.org/10.17882/51662>

Raven, J. a. (1998) Small is beautiful: the picophytoplankton. *Funct. Ecol.* **12**: 503–513.

Reynolds, C.S. (1984) Phytoplankton periodicity: the interaction of form, function and environmental variability. *Freshw. Biol.* **14**: 111–142.

Reynolds, C.S., Wiseman, S.W., Godfrey, B.M., and Butterwick, C. (1983) Some effects of artificial mixing on the dynamics of phytoplankton in large limnetic enclosures. *J. Plankt. Res.* **5**: 203–234.

Rigolet, C., Dubois, S.F., and Thiébaud, E. (2014) Benthic control freaks: Effects of the tubicolous amphipod *Haplopus nirae* on the specific diversity and functional structure of benthic communities. *J. Sea Res.* **85**: 413–427.

Romari, K. and Vaultot, D. (2004) Composition and temporal variability of picoeukaryote communities at a coastal site of the English Channel from 18S rDNA sequences. *Limnol. Oceanogr.* **49**: 784–798.

Scholz, B., Guillou, L., Marano, A. V., Neuhauser, S., Sullivan, B.K., Karsten, U., et al. (2016) Zoosporic parasites infecting marine diatoms-A black box that needs to be opened. *Fungal Ecol.* **19**: 59–76.

Schüßler, A., Schwarzott, D., and Walker, C. (2001) A new fungal phylum, the Glomeromycota: phylogeny and evolution. *Mycol. Res.* **105**: 1413–1421.

Smayda, T.J. and Reynolds, C.S. (2003) Strategies of marine dinoflagellate survival and some rules of assembly. *J. Sea Res.* **49**: 95–106.

Smetacek, V. (2002) The ocean's veil. *Nature* **419**: 2002.

Stoeck, T., Bass, D., Nebel, M., Christen, R., Jones, M.D.M., Breiner, H.W., and Richards, T.A.

(2010) Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol. Ecol.* **19**: 21–31.

Stoecker, D.K., Hansen, P.J., Caron, D.A., and Mitra, A. (2017) Mixotrophy in the Marine Plankton.

Ann. Rev. Mar. Sci. **9**: 311–335.

Sunagawa, S., Coelho, L.P., Chaffron, S., Kultima, J.R., Labadie, K., Salazar, G., et al. (2015)

Structure and function of the global ocean microbiome. *Science (80-.)*. **348**: 1261359–1261359.

Taylor, J.D. and Cunliffe, M. (2016) Multi-year assessment of coastal planktonic fungi reveals

environmental drivers of diversity and abundance. *ISME J.* **10**: 2118–2128.

Telesh, I., Schubert, H., and Skarlato, S. (2013) Life in the salinity gradient: Discovering mechanisms

behind a new biodiversity pattern. *Estuar. Coast. Shelf Sci.* **135**: 317–327.

de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., et al. (2015) Eukaryotic

plankton diversity in the sunlit ocean. *Science (80-.)*. **348**: 1261605.

Violle, C., Navas, M.L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I., and Garnier, E. (2007) Let

the concept of trait be functional! *Oikos* **116**: 882–892.

Ward, B.A. and Follows, M.J. (2016) Marine mixotrophy increases trophic transfer efficiency, mean

organism size, and vertical carbon flux. *Proc. Natl. Acad. Sci.* 201517118.

Weisse, T., Anderson, R., Arndt, H., Calbet, A., Hansen, P.J., and Montagnes, D.J.S. (2016)

Functional ecology of aquatic phagotrophic protists - Concepts, limitations, and perspectives.

Eur. J. Protistol. **55**: 50–74.

Worden, A.Z., Follows, M.J., Giovannoni, S.J., Wilken, S., Zimmerman, A.E., and Keeling, P.J.

(2015) Rethinking the marine carbon cycle: Factoring in the multifarious lifestyles of microbes.

Science (80-.). **347**: 1257594.

Worden, A.Z., Nolan, J.K., and Palenik, B. (2004) Assessing the dynamics and ecology of marine picophytoplankton: The importance of the eukaryotic component. *Limnol. Oceanogr.* **49**: 168–179.

Yabuki, A., Eikrem, W., Takishita, K., and Patterson, D.J. (2013) Fine Structure of *Telonema subtilis* Griessmann, 1913: A Flagellate with a Unique Cytoskeletal Structure Among Eukaryotes. *Protist* **164**: 556–569.

Yang, Z., Lowe, C.D., Crowther, W., Fenton, A., Watts, P.C., and Montagnes, D.J.S. (2013) Strain-specific functional and numerical responses are required to evaluate impacts on predator-prey dynamics. *ISME J.* **7**: 405–16.

Figure Legends

Fig. 1: Map of sampling sites. Shapes and colors of dots represent respectively the geolocation of samples from distinct oceanographic cruises and their sampling strategy.

Fig. 2: Coastal protist community structure in terms of a) genetic diversity (total relative read number associated to the taxa in the legend) and b) diversity of functional groups (total relative read number associated to the 6 functional groups in the legend), across planktonic size-fractions. a) Only taxa present above 10% in at least one sample are represented, other taxa are cumulated into ‘Others’. The group “Undetermined” cumulates the relative abundance of OTUs with low taxonomic affiliation (in a) and unresolved group annotation (in b). b) Functional groups identified are named with acronyms: PARA: Parasites, HET: Strict Heterotrophs, SAP: Saprobes, SWAT: Swimmer photo-autotrophs, FLAT: Floater photo-autotrophs, CAT: non-swimmer, strict-photoautotrophs, colony-forming photo-autotrophs.

Fig. 3: a) Theoretical framework of traits used to describe marine protists functional ecology and b) quality of the trait annotation for each of the 2007 taxonomic references associated to the OTUs of this study. a) The 30 traits chosen are ordered by trait type (Cell Morphology and Structure, Trophic Strategy, Physiology and Life Cycle) and associated to ecological/survival functions (Resources Acquisition, Reproduction, Predator Avoidance). Each trait is associated to different modalities. b) The proportion of the 2007 isolated taxonomic reference that have been annotated with the respective trait is represented in green. The references for which the chosen functional traits are not annotated are represented in red. References with “Low Taxonomic Resolution” corresponding to poorly determined OTUs not assignable functionally (i.e. super-group or family level) are represented in dark brown.

Fig. 4: Phylogenetic composition of the 6 functional groups. A simplified phylogenetic reconstruction among taxonomic groups was built inspired by a selected bibliography (Schübler *et al.*, 2001; Gómez *et al.*, 2009; Burki *et al.*, 2010; Howe *et al.*, 2011; Lin *et al.*, 2012; Berney *et al.*, 2013; Yabuki *et al.*, 2013; Keeling *et al.*, 2014; Worden *et al.*, 2015; Aleoshin *et al.*, 2016). The OTUs of each functional group were associated with a classified taxon. The relative contribution (number of OTUs on total OTUs number in the functional group) of the distinct taxa to the pool of OTU from each functional group has been represented by colors (from grey to red to represent low to high contribution). Functional groups identified are named with acronyms: PARA: Parasites, HET: Strict Heterotrophs, SAP: Saprobes, SWAT: Swimmer photo-autotrophs, FLAT: Floater photo-autotrophs, CAT: non-swimmer, strict-photoautotrophs, colony-forming photo-autotrophs.

Fig. 5: Taxonomic gradients across samples and size-fractions, with associated functional group composition. At the left, and from top to bottom, Non-Metric Multidimensional Scaling analyses (NMDS with Bray-Curtis distance) based on the genetic diversity (OTUs) of each sample, are represented for micro-, nano-, and pico-plankton. Dot shapes identifies the sample's dataset. Stress values of each NMDS plot, represented at top-right, indicates that two axes were sufficient to represent community dissimilarity between samples. Arrows represent environmental variables fitted onto the two NMDS axes. Samples clustering (color of the dots) was calculated impartially through the k-mean partitioning of samples in different number of clusters, followed by computation of the simple structure index (ssi) to select the best partitioning. Environmental variables were fitted onto NMDS ordinations with function *envfit()* of R package "vegan" (Osaksen *et al.*, 2016). At the right, the mean relative importance of each functional group and their standard deviations values (error bars) within each cluster of samples in each size fraction are represented. Functional groups identified are named with acronyms: PARA: Parasites, HET: Strict Heterotrophs, SAP: Saprobes, SWAT: Swimmer

photo-autotrophs, FLAT: Floater photo-autotrophs, CAT: non-swimmer, strict-photoautotrophs, colony-forming photo-autotrophs.

Fig. 6: Boxplots comparing 3 diversity metrics calculated for all samples of micro-, nano- and pico-plankton. a) Shannon index H' calculated on the relative abundances of the 6 functional groups and b) relative OTU abundance; c) OTUs richness with micro-plankton containing a total of 56 655 OTUs, nano-plankton 85 373 and pico-plankton 64 404. The Significance of the differences in metric values (Kruskall-Wallis test) between each size-fractions is shown above the boxplots (ns: non-significant; ***: significant with $p.value < 0.001$).

Fig. 7: Explanatory scheme of the workflow and methods used in our study.

Table Legend

Table 1: Information on the ecosystems sampled in this study.

	Sampled sites	Environmental forcing	Depth	Year(s)	Season	Dataset type	Spatio-temporal sample frequency	Filters	Sampled size-fractions (μm)	Replicate(s)
DA	Bay of Brest (Daoulas River)	Regular dinoflagellate bloom occurring in estuary typical mid-latitude ecosystem, influenced by the Loire plume	Surface	2013 - 2014 - 2015	Spring-summer (May to August)	Temporal	Every 3 days	245	0.2-3; 3-20; >20	1 (for 3-20 and >20)
DY	South Brittany (Bay of Concarneau)		Surface	2012	Spring-summer (March to July)	Temporal	Weekly and twice weekly during bloom	78	0.2-3; 3-20; >20	None
MB	Iroise Sea	Ushant tidal front	Surface/ DCM/ Mesopelagic	2014 - 2015	Spring-summer (March, July and September)	Spatio-temporal	5 stations (coast to off-shore radial)	414	0.2-3; 3-10; >10	2 (all)
PE	Bay of Biscay (Gironde River plume)	Gironde plume	Surface/ DCM	2012 - 2013	Spring (May)	Spatio-temporal	8 stations (coast to off-shore radial)	78	0.2-3; 3-20; >20	None
PH	Bay of Brest	Tidal-influenced semi-enclosed ecosystem	Surface/ DCM	2013 - 2015	Summer-autumn (June, September)	Spatio-temporal	16 stations	87	0.2-3; 3-20; >20	1 (for 3-20; >20 μm during 2015)
PI	Bay of Biscay (Loire River plume)	Loire plume	Surface/ DCM	2013 - 2015	Summer-autumn (June, September)	Spatio-temporal	23 stations (north-south radial)	87	0.2-3; 3-20; >20	1 (for 3-20; >20 μm during 2015)
RA	North Brittany	Permanent water-column mixing	Surface	2009 - 2010 - 2011	All	Temporal	Twice monthly	138	0.2-3	None
SE	South-West Senegal	Upwelling ecosystem	Surface	2015	Spring (November)	Spatial	6 stations (coast to off-shore radial)	26	0.2-3; 3-20; >20	1 (for 3-20 and >20)

Table 1: Information on the ecosystems sampled in this study.













