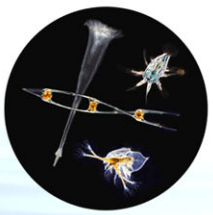
Journal of
Plankton Research

academic.oup.com/plankt



J. Plankton Res. (2018) 00(00): 1–18. doi:10.1093/plankt/fby035

REVIEW

Molecular analyses of protists in long-term observation programmes—current status and future perspectives

ROWENA STERN[1]*, ALEXANDRA KRABERG[2], EILEEN BRESNAN[3], WIEBE H.C.F. KOOISTRA[4], CONNIE LOVEJOY[5], MARINA MONTRESOR[4], XOSÉ ANXELU G. MORÁN[6], FABRICE NOT[7,8], RAFAEL SALAS[9], RAFFAELE SIANO[10], DANIEL VAULOT[7,8], LINDA AMARAL-ZETTLER[11], ADRIANA ZINGONE[4] AND KATJA METFIES[12,13]

[1]MARINE BIOLOGICAL ASSOCIATION, CITADEL HILL, PLYMOUTH PL1 2PB, UK, [2]BIOLOGISCHE ANSTALT HELGOLAND, ALFRED WEGENER INSTITUTE HELMHOLTZ CENTRE FOR POLAR AND MARINE RESEARCH, KURPROMENADE 201, HELGOLAND 27498, GERMANY, [3]MARINE SCOTLAND MARINE LABORATORY, 375 VICTORIA ROAD, ABERDEEN AB11 9DB, UK, [4]DEPARTMENT OF INTEGRATIVE MARINE ECOLOGY, STAZIONE ZOOLOGICA ANTON DOHRN, 80121 NAPOLI, ITALY, [5]QUÉBEC-OCÉAN, DÉPARTEMENT DE BIOLOGIE, INSTITUT DE BIOLOGIE INTÉGRATIVE ET DES SYSTÈMES (IBIS), UNIVERSITÉ LAVAL, QUÉBEC QC G1V 0A6, CANADA, [6]RED SEA RESEARCH CENTER, BIOLOGICAL AND ENVIRONMENTAL SCIENCES AND ENGINEERING DIVISION, KING ABDULLAH UNIVERSITY OF SCIENCE AND TECHNOLOGY, THUWAL 23955-6900, SAUDI ARABIA, [7]SORBONNE UNIVERSITÉS, UPMC UNIV PARIS 06, CNRS, UMR 7144, FRANCE, [8]STATION BIOLOGIQUE, PLACE GEORGES TEISSIER, 29680 ROSCOFF, FRANCE, [9]MARINE INSTITUTE, MARINE ENVIRONMENT AND FOOD SAFETY SERVICES, RINVILLE, ORANMORE H91 R673, REPUBLIC OF IRELAND, [10]IFREMER, CENTRE DE BREST DEPARTMENT DYNECO, BP70, PLOUZANE 29280, FRANCE, [11]DEPARTMENT OF MICROBIOLOGY AND BIOGEOCHEMISTRY, NIOZ, TEXEL, PO BOX 59, 1790 AB DEN BURG (TEXEL), THE NETHERLANDS, [12]ALFRED-WEGENER INSTITUTE, HELMHOLTZ CENTRE FOR POLAR AND MARINE RESEARCH, BREMERHAVEN 27515, GERMANY AND [13]HELMHOLTZ INSTITUTE FOR FUNCTIONAL MARINE BIODIVERSITY AT THE UNIVERSITY OF OLDENBURG, OLDENBURG 26129, GERMANY

*CORRESPONDING AUTHOR: rost@mba.ac.uk

Received August 9, 2017; editorial decision August 7, 2018; accepted August 14, 2018

Corresponding editor: John Dolan

Protists (microbial eukaryotes) are diverse, major components of marine ecosystems, and are fundamental to ecosystem services. In the last 10 years, molecular studies have highlighted substantial novel diversity in marine systems including sequences with no taxonomic context. At the same time, many known protists remain without a DNA identity. Since the majority of pelagic protists are too small to identify by light microscopy, most are neither comprehensively or regularly taken into account, particularly in Long-term Ecological Research Sites. This potentially undermines the quality of research and the accuracy of predictions about biological species shifts in a changing environment. The ICES Working Group for Phytoplankton and Microbial Ecology conducted a questionnaire survey in 2013–2014 on methods and identification of protists using molecular methods plus a literature review of protist molecular diversity studies. The results revealed an increased use of high-throughput sequencing methods and a recognition that sequence data enhance the overall datasets on protist species composition. However, we found only a few long-term molecular studies and noticed a lack of integration between microscopic and molecular methods.

available online at academic.oup.com/plankt

Here, we discuss and put forward recommendations to improve and make molecular methods more accessible to Long-term Ecological Research Site investigators.

KEYWORDS: protists; Long-term Ecological Research Station; molecular; time-series; questionnaire; literature survey

INTRODUCTION

The scientific community widely recognizes that environmental change affects biological systems and their ecosystem services. Marine planktonic protists (microscopic eukaryotes living in surface waters) are one of the first group of organisms to respond to changes in physical conditions and useful indicator organisms of such change (Hays *et al.*, 2005). Rising sea surface and air temperatures have already been found to have profound consequences (potentially) affecting species' distributional ranges (Beaugrand *et al.*, 2009) and favouring species invasions (Sorte *et al.*, 2010). Increases in sea surface temperatures have been linked to increases in some marine pathogens such as *Vibrio* spp. (Motes *et al.*, 1998; Vezzulli *et al.*, 2016; Muhling *et al.*, 2017) affecting humans and marine animals and are under discussion as favouring harmful algal blooms (Wells *et al.*, 2015).

A time series is a set of regular time-ordered observations of a quantitative characteristic of an individual or collective phenomenon taken at successive, in most cases equidistant, periods/points of time (OECD glossary of statistical terms: https://stats.oecd.org/glossary/detail.asp?ID=2708). Several research institutions and monitoring stations maintain time series in order to answer ecological and systematic questions using long-term time series data. Monitoring organizations that carry out long-term time series are referred to as Long-Term Ecological Research Sites (LTER) and we focus on marine LTERs but these findings are equally applicable to freshwater LTERs. Time series require precise standardization of procedures, from sampling to analyses to make meaningful comparisons among different time intervals. Several marine research institutes across Europe maintain such time series and the data gathered at these sites fuel many research projects. We have identified 24 such sites and oceanographic time series in the North Atlantic and Mediterranean that have routinely identified and monitored planktonic protists for over a decade. They include the Continuous Plankton Recorder Survey, which has been monitoring plankton biomass and diversity in surface Atlantic waters for 85 years (Reid *et al.*, 2003), the Biological Station at Helgoland (Kraberg *et al.*, 2015) for 56 years, the LTER-MareChiara at Naples (Zingone) for 32 years, the Scottish Coastal Observatory (15–20 years, Bresnan *et al.*, 2016) and the Western Channel Observatory (Southward *et al.*, 2004), for 113 years. At present, many countries have well-established monitoring programmes with standardized procedures for collecting and counting organisms and a clear reporting structure to governmental and non-governmental stakeholders. Light microscopy is currently the most common method of data collection for monitoring plankton abundance in many plankton time series methods, being relatively rapid and inexpensive to set up and provide additional information on trophic status of plankton. The method of choice is the Utermöhl inverted microscopy method (Utermöhl, 1931; Lund *et al.*, 1958). This is considered to be a standard method, although considerable variability occurs in the way the organisms are counted in the Utermöhl chamber, despite the fact that suggestions for standardized protocols have been made, e.g in Sournia (1978), see also Zingone *et al.* (2015).

Time series data from this type of monitoring have provided the basis of our current understanding of the spatial and temporal distribution of plankton (Reid *et al.*, 1998) and have helped to link physico-chemical drivers to patterns in marine biodiversity. Planktonic taxonomic records provide a framework for indentifying shifts in species distribution over space and time. Unfortunately, due to ambiguities owing to the limitations of routine light microscopy, methodological issues, operator fatigue and even taxonomic bias (different time series can place a focus on different taxon groups based on the expertise of available analysts) taxonomic records based on microscopy are rarely complete (Culverhouse, 2015). This lack of resolution is particularly problematic for smaller single-celled species, and those living in cryptic habitats as parasites and symbionts or with inconspicuous life-cycle stages and a combination of methods are often required for detailed taxonomic assessments (Culverhouse, 2015; Jeuck *et al.*, 2017). To highlight the scale of these issues, a global study comparing catalogued planktonic morphospecies with the projected number of planktonic taxa acquired by genetics, showed that genetic methods identified an additional 138 800 taxa (De Vargas *et al.*, 2015). Most of these taxa were under 10 μm, heterotrophic or were symbionts, thus impossible to detect by light microscopy. Therefore species distributions, especially for the smaller protists, are not

effectively known despite long-term routine assessments in a number of European time series programmes. However, it is for these morphologically indistinct organisms that molecular technologies hold great promise for the future (Epstein and Lopez-Garcia, 2008).

Over the past decade, numerous publications have shown the power of the analysis of ribosomal genes (rRNA genes) to gain new insights into the phylogeny and biogeography of prokaryotic and eukaryotic microorganisms (Sunagawa *et al.*, 2015). The genes coding for the rRNA are particularly well suited for phylogenetic analysis and taxonomical identification, because they are universally present in all cellular organisms and thus have been suggested as a primary barcode marker (Pawlowski *et al.*, 2012). Furthermore, rRNA genes are of relatively large size and contain both highly conserved and variable regions with no evidence for lateral gene transfer (Woese, 1987). Molecular methods targeting rRNA genes often take advantage of target-specific molecular probes that can be used in combination with a wide variety of hybridization-based methods, such as nucleic acid biosensors (Diercks *et al.*, 2008; Ussler *et al.*, 2013) quantitative PCR (Bowers *et al.*, 2010; Toebe *et al.*, 2013) or fluorescence *in situ* hybridization (FISH) (Thiele *et al.*, 2014) to generate information on occurrence and abundance of selected taxa. In contrast, high-throughput sequencing methods (HTS) using next-generation sequencing platforms, such as Illumina, Ion Torrent and related technologies (Scholz *et al.*, 2012) provide high resolution, taxon-specific information on variability and composition of whole microbial communities that is independent of the size or morphology of target organisms, including their smallest size fractions and the rare biosphere (Kilias *et al.*, 2014b; De Vargas *et al.*, 2015). As sequencing costs have plummeted and HTS technology has become common, the discovery rate on novel protist lineages and potentially new species has continued to climb. HTS exceeds Sanger sequencing data ouput by orders of magnitude, allowing for greater species detection extending to rare types. Such methods have uncovered an enormous variety of hitherto unknown diversity (López-García *et al.*, 2001; Moon-Van Der Staay *et al.*, 2001; Amaral-Zettler *et al.*, 2009, 2010; Kilias *et al.*, 2014b; De Vargas *et al.*, 2015). Thus, routine molecular monitoring of key microbial components is vital to fill this knowledge gap to provide meaningful, long-term data needed to judge the environmental status of key marine systems (Moffat *et al.*, 2011).

During the 2013 annual meeting of the ICES Working Group Phytoplankton and Microbial Ecology (WGPME) held at Helgoland (Germany), it was decided that a review current practices in molecular microbial marine surveys was needed through a questionnaire (answered by 14 researchers engaged in molecular marine biodiversity studies) and a literature review of research involving 67 studies in order to elucidate the extent to which molecular methods are currently used at monitoring and Long-term Ecological Research Sites.24

Overall the questionnaire was expected to: (i) provide an overview on the kind and location of recent molecular time series projects; (ii) elucidate which technological approaches are used in the context of the molecular time series projects; (iii) summarize bioinformatics approaches to analyse molecular time series data; and (iv) elucidate the understanding of the community on the benefit of molecular methods for long-term observation of marine microbes. The detailed methods used are described in Supplementary document S1 and detailed results of the literature review and questionnaire are described in Supplementary Tables S1 and S2, respectively. Here we review the survey data on molecular diversity and evaluate the extent of complementary with conventional biological time series to determine if they can be harmonized.

MATERIAL AND METHODS

In order to accomplish the aims of this study, we carried out a literature review of molecular methodology that has been applied over the last 15 years. In addition, as molecular technology is rapidly evolving, we devised a questionnaire for current methodologies to assess how methodologies are changing and current issues faced by researchers.

Literature search

A review of the available literature on molecular methods and their use in protist monitoring was performed. Searches were carried out in Web of Science in 2014–2015 using the search terms 'protist diversity+molecular method+OTU (3 hits), protist diversity+OTU (13 hits) and protist +diversity+molecular+marine (138 hits). The initial goal had been to investigate diversity assessments that are carried out regularly as part of a monitoring time series. However, as these were rare, diversity assessments resulting from individual cruises or cruise transects were also included. As many metrics for reporting diversity from molecular surveys are in use, this survey concentrated on surveys reporting OTUs. Excluded were articles pertaining to freshwater diversity or those investigating a specific organism without reference to sample diversity. Publications pertaining exclusively to prokaryotic diversity were also excluded from the analysis as they were outside the scope of this review.

Of the hundreds of articles reviewed, 73 fulfilled the initially set criteria of being marine or marine-derived systems and after removing studies focusing on long-term marine cultures or benthic communities, 67 articles were used for analysis (Supplementary Table SI). The data from the literature review were categorized according to the questions posed in the questionnaire so that results could be compared and discussed jointly. Some studies in the literature review compared different methods. For the summary graphs of methodological approaches, these were treated separately. The molecular methods included all DNA-based detection systems. However, some studies additionally carried out microscopic (light microscopy, electron microscopy and microscopy using fluorescent DNA probes) identification comparisons.

Questionnaire

A questionnaire was devised to ascertain how methodologies were changing, what current practices were being used and current issues faced by researchers. The questionnaire was circulated in 2013 and in 2014 the responses were gathered from 14 investigators in the field of molecular assessments of marine protists. All investigators had experience in classical, microscopy-based, long-term monitoring and in molecular assessment of marine microbes. The questionnaire was not circulated to researchers involved in monitoring toxin-producing phytoplankton as part of a shellfish hygiene monitoring programme as the use of molecular methods in these programmes is clearly established (Rhodes *et al.*, 2013; Cusack *et al.*, 2016). The final questionnaire was composed of 18 clearly worded, simple and concise questions in order for respondents to fully answer the questions at a reasonable rate and to increase chances of full participation. Participation in the questionnaire was acknowledged by an offer to contribute as a co-author to this article.

The first WGPME molecular time series screen—list of questions

Question number	Question	Category
1	Have you produced molecular time-series data?	
2	What taxonomic group(s) are your primary research focus?	Target groups and target molecules
3	For what time period did you carry out your molecular survey?	Sampling and long-term storage
4	At what frequency was this molecular survey done? (e.g. monthly, weekly, etc)	Sampling and long-term storage
5	Is this from a single marine station(s) or from cruise data?	Sampling and long-term storage
6	What were your sample volumes?	Sampling and long-term storage
7	How did you prepare equipment and preserve samples?	Sampling and long-term storage
8	Have you extracted DNA, RNA or both?	Sampling and long-term storage
9	How long did you store the samples before extracting nucleic acid?	Sampling and long-term storage
10	At what temperature did you store your samples before nucleic acid extraction?	Sampling and long-term storage
11	After DNA extraction, how do you preserve your DNA/RNA?	Sampling and long-term storage
12	If you quantify your samples, what method do you use?	Target groups and target molecules
13	What assays did you use to test your samples? If sequencing, which nucleic acid markers did you chose and your primer set?	Target groups and target molecules
14	If you used next-generation sequencing, did you analyse your data using public/commercial software package or custom-made package?	Analysis approaches
15	What would you consider were the positive and negative aspects of your downstream analysis?	Analysis approaches
16	Did you need to create your own database?	Analysis approaches
17	What do you feel is the additional value of molecular data to your time series?	Analysis approaches
18	How beneficial do you feel a detailed protocol would be to you/your colleagues?	Analysis approaches

RESULTS AND DISCUSSION

The literature survey captured past methodologies that focused on short-, fixed-term assessments of diversity. These studies had different goals compared with long-term surveys but nevertheless shared procedures in DNA preservation, storage and extraction and the use of similar DNA markers. However, our review highlights a lack of consistent methodology used among studies carried out, both in the literature reviewed and those referred to in the questionnaire responses. The questionnaire response that represents more current practices revealed more consistency in methodology in some aspects. However, a wide variety of sampling methods are still applied at different sites that make comparisons difficult. This is an issue for many global microbiological efforts (Emmett Duffy *et al.*, 2013; Dubilier *et al.*, 2015) and hinders data sharing and integrated analysis across different fields of expertise. LTER

sites are subjected to different ecological factors, so using the same volume may not be a useful strategy. It may be more appropriate to develop a diversity saturation index (to obtain all diversity, including rare types) that can be experimentally evaluated at each site and a comparable variable across all sites.

The questionnaire responses showed consistency in how the respondents viewed the benefits of molecular surveys, in particular by way of increased detection of organisms that are challenging to identify by light microscopy. However, the difficulty in integrating molecular data with conventional time series was a major concern. In particular, the need to harmonize or at least agree on sample collection and storage, sample volumes and target groups and molecules, and taxonomy was highlighted.

The results of the literature survey and the questionnaire were analysed and discussed with respect to three key aspects highly relevant for the accomplishment, results and informative value of molecular surveillance of eukaryotic marine microbes: (i) sampling and long-term storage, (ii) target groups and target molecules and (iii) analyses approaches.

Sampling and long-term storage

The vast majority of literature-based studies were not a component of regular monitoring activities, but the result of biodiversity assessments during cruises for a particular project (e.g. Díez *et al.*, 2001; López-García *et al.*, 2001; López-García *et al.*, 2003; Latasa *et al.*, 2004; Lovejoy *et al.*, 2006; Marie *et al.*, 2006; Chambouvet *et al.*, 2008; Guillou *et al.*, 2008; Massana and Pedros-Alio, 2008; Not *et al.*, 2008; Viprey *et al.*, 2008; Terrado *et al.*, 2009; Gao *et al.*, 2010; Edgcomb *et al.*, 2011; Bachy *et al.*, 2012; Ishitani *et al.*, 2014; Jiang *et al.*, 2014; Kilias *et al.*, 2014a; Thiele *et al.*, 2014; Weber and Pawlowski, 2014; Wolf *et al.*, 2014). The majority of studies (82%) only sampled the same site once or twice (Fig. 1A) such as (López-García *et al.*, 2003; Lee *et al.*, 2010; Koid *et al.*, 2012; Bazin *et al.*, 2014) to carry out biodiversity studies of surface marine communities or rarely as deep ocean communities (López-García *et al.*, 2003; Countway *et al.*, 2007; López-García *et al.*, 2007; Not *et al.*, 2007a,b; Amaral-Zettler, 2012) compared with 16% that sampled a single site more than ten times (Fig. 1A) such as (Joo *et al.*, 2010; Piwosz and Pernthaler, 2011; Kim *et al.*, 2013). Most time series from cruises or monitoring stations were done within 1 year with only 19% of molecular time series extended to more than 1 year (Short and Suttle, 2003; Romari and Vaulot, 2004; Zhu *et al.*, 2005; Medlin *et al.*, 2006; Behnke *et al.*, 2010; Amacher *et al.*, 2013; Kim *et al.*, 2013). Longer studies were used to investigate protist responses to environmental factors such as particle fluxes from shallow to deep waters at LTER (Amacher *et al.*, 2013), or look at rarely characterized protists or environments such as nano- or picoeukaryotic diversity in the Arctic Ocean (Terrado *et al.*, 2009; Bachy *et al.*, 2011; Monier *et al.*, 2013, 2014), Fungal diversity (He *et al.* 2014), Polycystine radiolarian symbionts (Dolven *et al.*, 2007). Monitoring stations or regular transect routes that carried out molecular surveys include LTER, L4 (Romari and Vaulot, 2004; Taylor and Cunliffe, 2014; Stern *et al.*, 2015), Helgoland Roads Station (Medlin *et al.*, 2006), Norwegian fjords (Piquet *et al.*, 2014), Bermuda Atlantic Time Series (Amacher *et al.*, 2013) or at a Korean reservoir (Joo *et al.*, 2010). Most studies focused on newly collected samples, rather than archival sources. A large focus was on species discovery, due, in part, to new technological developments that allowed high throughput

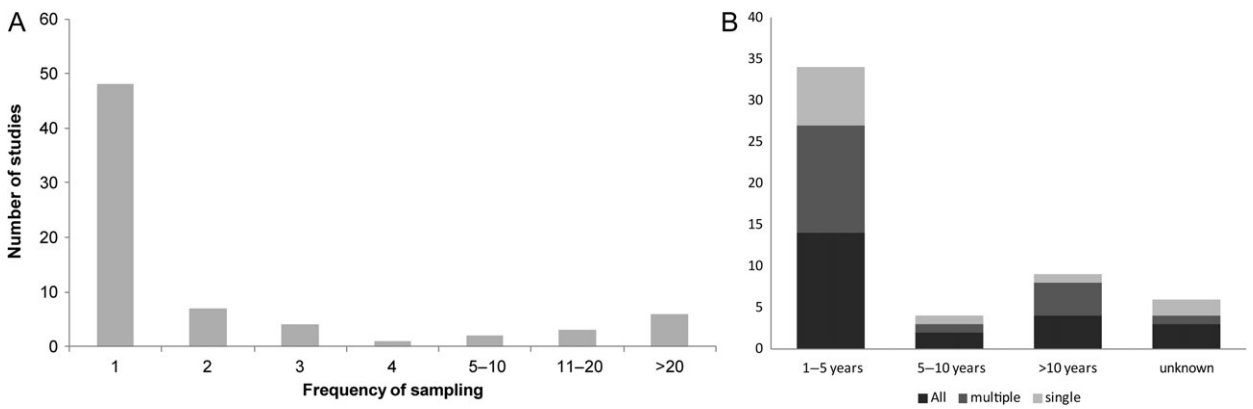


Fig. 1. Temporal sampling effort of researchers involved in molecular surveys from literature review (**A**), showing frequency of repeated sampling of a station or area and our questionnaire (**B**), showing the number surveys that have repeatedly sampled a region.

or larger scale genetic studies (Moon-Van Der Staay *et al.*, 2000; López-García *et al.*, 2001; Massana and Pedros-Alio, 2008; Heywood *et al.*, 2010; Stern *et al.*, 2010; Schnetzer *et al.*, 2011; Wylezich and Jürgens, 2011; Amaral-Zettler, 2012; Alemzadeh *et al.*, 2014). A few studies used single-cell isolation approach only with larger protists (Yuasa *et al.*, 2006; Ishitani *et al.*, 2014; Weber and Pawlowski, 2014) or using secondary sorting for small protists (Shi *et al.*, 2011) and a minority investigated protists in marine-derived cultures (Atkins *et al.*, 2000). Only seven studies sampled benthic substrates ranging from beach sand to deep sea hydrothermal vents sediments and sponge samples, with only one of these studies carried out repeated sampling (Atkins *et al.*, 2000; López-García *et al.*, 2003; López-García *et al.*, 2007; Stoeck *et al.*, 2009; Stern *et al.*, 2010; Amaral-Zettler, 2012; He *et al.* 2014). Two studies investigated ice cores (Bachy *et al.*, 2011; Kilias *et al.*, 2014b). All time series described in the questionnaire were established relatively recently during the past decade. The US microbial observatory (Kim *et al.*, 2013) was the earliest molecular survey starting in 2000 and the longest running molecular time series, ending in 2010. Responses from the questionnaire (questions 3–11) showed that sampling was highly variable from weekly time series at single locations through to those carried out over defined time intervals on annual basis that covered larger geographical areas. However, most of these were under 5 years with 10% over 10 years (Fig. 1B). Multiple surveys were often employed, about half of which are ongoing and of these, five survey multiple sites on an ongoing basis.

Sampling volumes in studies from the literature review were highly variable (Fig. 2) and ranged from <1 mL (Shi *et al.*, 2011) up to 500 L (Koid *et al.*, 2012), depending on the end-point analysis. The most frequent volume ranges were 1–10 L (43%) followed by 0–1 L (33%). A few were done through plankton nets so no volume was available (Yuasa *et al.*, 2006; Stern *et al.*, 2010; Seears *et al.*, 2012; Alemzadeh *et al.*, 2014). Most studies (52) did not use any replication using the same methods, although many reduced sampling biases through sample or PCR replication and pooling. About half compared their results to a genetic or non-genetic method. Filtration was common but methods varied. Sampling volumes among questionnaire participants were similarly variable mainly in the range 0.5–2 L, overall ranging from 0.1 to 10 L (see Fig. 2). By contrast processing of samples appears consistent in both cases. Both literature and questionnaire results showed most respondents used either polyethersulfone membrane or polycarbonate filters and either snap frozen in liquid nitrogen (more common among questionnaire participants) and stored frozen at −80°C alone or in Tris-EDTA buffer. Short-term storage was at −20°C. A quarter of studies in the literature review included preservation of the samples with lysis buffer or RNAlater prior to freezing. Reported periods of long-term storage of the samples were variable for questionnaire participants. They ranged from weeks up to 5 years (most were between 8 months and 2 years). Further complications come from size fractionation of samples either by net or filtration to measure different size components.

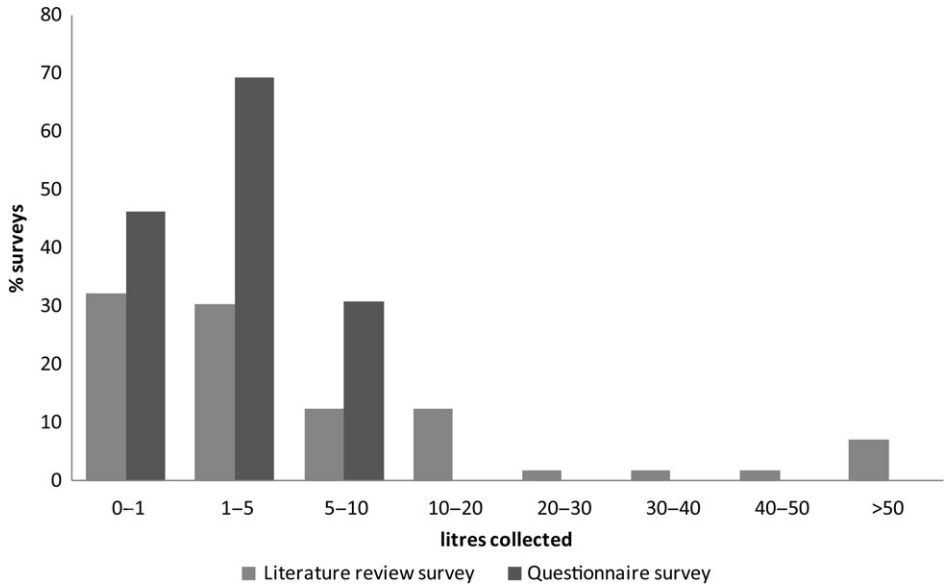


Fig. 2. Ranges of volumes collected from marine molecular surveys sourced from a literature review and questionnaire. Many respondents deployed more than one survey and open water surveys tended to collect greater volumes.

Often these are not effective and larger organisms are found in samples fractionated for smaller organisms because the methods can disrupt larger cells or from contamination of dead, fragmented cells and such methods are biased as they exclude symbiotic or parasitic relationships (De Vargas *et al.*, 2015). However, cell sorting methods can rapidly sort and photograph individual cells even up to zooplankton size (Sosik *et al.*, 2010) and linking these to HTS data may resolve physical size fractionation issues.

The answer to the question of "How much volume is enough to capture all diversity" depended on the system. Coastal systems have more nutrients and tend to support greater cell numbers and the dominance of a few species, while some open water systems have lower concentrations and more even distributions. These differences may result in non-comparable diversity assessments among sites. In some instances, the filter volume depends on the purpose of the monitoring programme and the emphasis put on "rare species". Monitoring for the impacts of anthropogenic nutrient inputs assumes a high phytoplankton biomass and thus the volume sampled in this instance may be less than a programme where the presence of "rare species" such as early warnings for toxin producing species or non-native introductions is the focus of the study. A coordinated effort to determine complete or near complete representation of diversity would allow empirical determination of recommended volumes in different regions.

Target groups and target molecules

The majority of studies in the literature review surveyed total protist diversity (Fig. 3) including a large proportion focusing on picoeukaryotes. Some studies also targeted diversity in specific taxonomic groups, e.g.

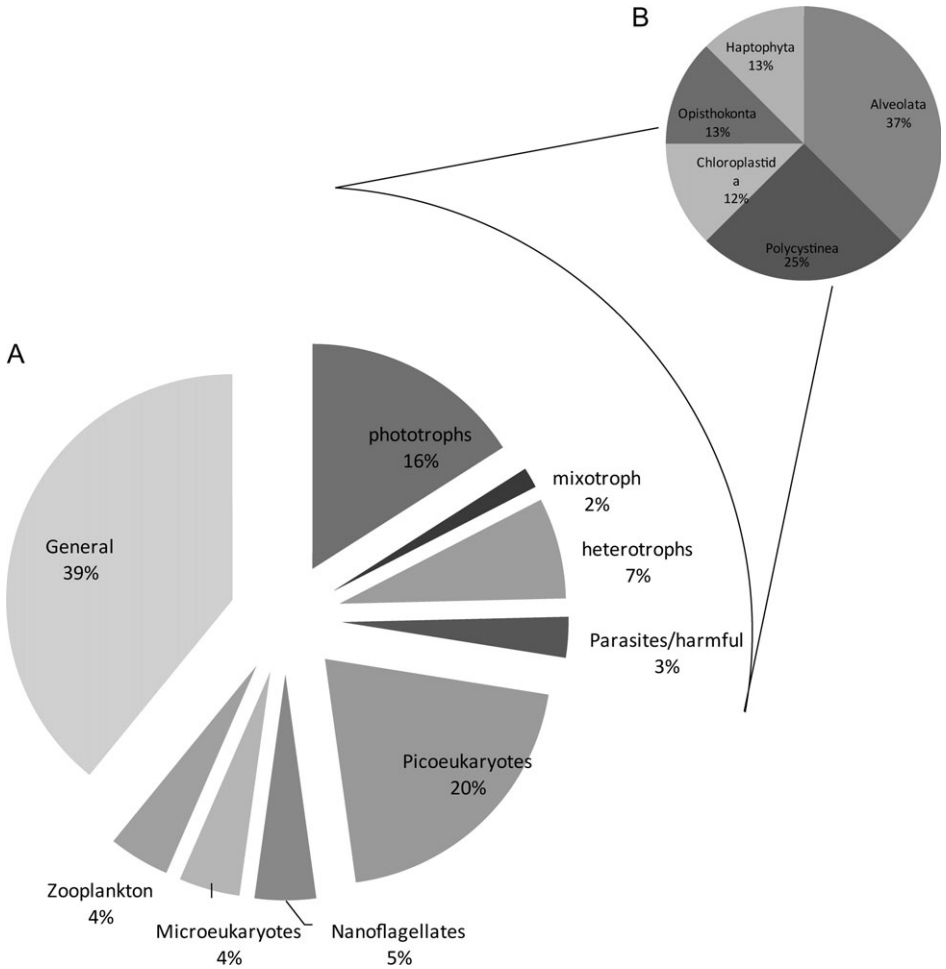


Fig. 3. Breakdown of microbial eukaryote groups studied in the literature review shown as functional groups types (**A**). Studies targeting specific taxa groups are shown in brackets and their taxonomic categories (**B**). The "General" category included author-defined terms of microbial eukaryotes, plankton, eukaryotes and marine protists. All other categories used were author-defined.

dinoflagellates, foraminifera, radiolaria or fungi as a systematic study. A total of 112 different primers were used in the published studies, reviewed here. The majority of published studies (just over 94%) used a partial 18S rDNA gene as a target taxonomic marker with a minority using the nuclear-encoded ribosomal Internal Transcribed Spacer region (ITS) where ITS1 and ITS4 primers are commonly used (White *et al.*, 1990) and/or COI markers (e.g. Stern *et al.*, 2010; Kilias *et al.*, 2014a). Some studies, additionally, used a combination of markers (18S, ITS) such as (Wolf *et al.*, 2014). The most commonly used primers for 18S marker, EukA or Euk1A (Medlin *et al.*, 1988), Euk328F (Moon-Van Der Staay *et al.*, 2001) and Euk528F (Van Hoek *et al.*, 1998) as forward primers and Euk516R (Amann *et al.*, 1990), Euk329R (Moon-Van Der Staay *et al.*, 2001) and EukB (Medlin *et al.*, 1988) targeting the entire 18S region or conserved regions proximal to the terminus of 18S marker notably the V4 or V9 region. The 18S V4 region was proposed as a primary barcode region (as COI is not suitable for many protists) by Pawlowski *et al.* (2012) primarily because of its ubiquity in all protist cells and its high usage in public databases. It is likely, this marker will continue to be used. Nested PCR amplification approaches, applying two sets of primers, were often used, resulting in a greater variety of taxa identified. Primer sets did not identify species equally and detection often was skewed towards Alveolata and Heterokontophyta. In the former case, this is likely due to lineage dependent differences in rDNA copy numbers per cell (Zhu *et al.*, 2005). Additionally, DNA from heterotrophic cells (such as MAST and MALV belonging to Heterokonta and Alveolata) can be easier to extract than autotrophs that often have tough cell walls (Vaulot *et al.*, 2002), thus more likely to be amplified by PCR Excavata and Rhizaria were the least represented using these primers, although the latter were abundant in other studies (De Vargas *et al.*, 2015).

All participants of the questionnaire (questions 2, 11–12) had carried out molecular time series studies targeting mostly protists but additionally also surveyed prokaryotes (bacteria and archaea). At three sites, only bacteria were investigated but as the methodology used in bacteria and protists were very similar, these were included in the analysis. The molecular methods were mainly used to analyse community composition or diversity of protists. In agreement with the literature survey, there was broad consensus on analysing rDNA taxonomic gene marker (18S) to assess protist diversity. Most of the studies were based on the analyses of DNA, rather than RNA sequences.

The use of the 18S DNA marker, although common and useful in identifying a large spectrum of organisms at the major group level, is not ideal to identify organisms to species. Adl *et al.* (2014) show a bewildering number of primer sets used to identify soil microbial eukaryotes. Different primer sets have been shown to generate a biased taxa dataset, through the exclusion of species (Hugerth *et al.*, 2014). However, both the literature review and questionnaire showed that only a few primer sets are popular, making it likely that a common set of primers could be developed to detect a broad and more representative spectrum of taxa. Primer development continues, and progress and an agreement on a set of universal eukaryotic primers might be achievable given our findings and would benefit those involved in molecular surveys. Yet, the popularity of primer combinations can be expected to wax and wane, driven by HTS developments to sequence progressively longer pieces of DNA accurately. Longer DNA metabarcodes allow for increased accuracy with which different taxa can be discriminated. Additionally, improvements in strategies to capture species representatively are continuously improving, for example through DNA-enrichment methods (reviewed by Mamanova *et al.* (2010)) while *in silico* checks can be made to detect functional regions of RNA molecules that take part in protein binding, reducing false positives detection from pseudogenes (Smyth *et al.*, 2015).

Analysis approaches

The literature review revealed, a great diversity of methods (Fig. 4). Perhaps as a result of the initial high-sequencing costs, the most common published methods were fingerprinting methods were ARISA (Automated Approach for Ribosomal Intergenic Spacer Analysis), DGGE (Denaturing Gradient Gel Electrophoresis), TGGE (Temperature Gradient Gel Electrophoresis), RFLP (Restriction Fragment Length Polymorphism), T-RFLP (Terminal Restriction Fragment Length Polymorphism), SSCP (Single=Strand Conformation Polymorphism), PFGE (Pulse-field Gel Electrophoresis), often in combination with another method such as FISH or microscopy (Diéz *et al.*, 2001; Larsen *et al.*, 2001; Savin *et al.*, 2003; Massana *et al.*, 2004, 2006; Countway *et al.*, 2005; Medlin *et al.*, 2006; Not *et al.*, 2008; Lee *et al.*, 2010; Piquet *et al.*, 2010; Balzano *et al.*, 2012; Filker *et al.*, 2013; Kilias *et al.*, 2014a), as examples. Fingerprinting methods measured diversity but not taxonomic affiliation. However, Sanger sequencing using clone libraries and high-throughput sequencing (also called 454-Titanium sequencing, pyrosequencing, massively parallel pyrotag sequencing, Illumina GAIIx sequencing, metabarcoding) methods were the most common (43/67 studies). Quantitative genetic approaches were least favoured in the survey (<5%).

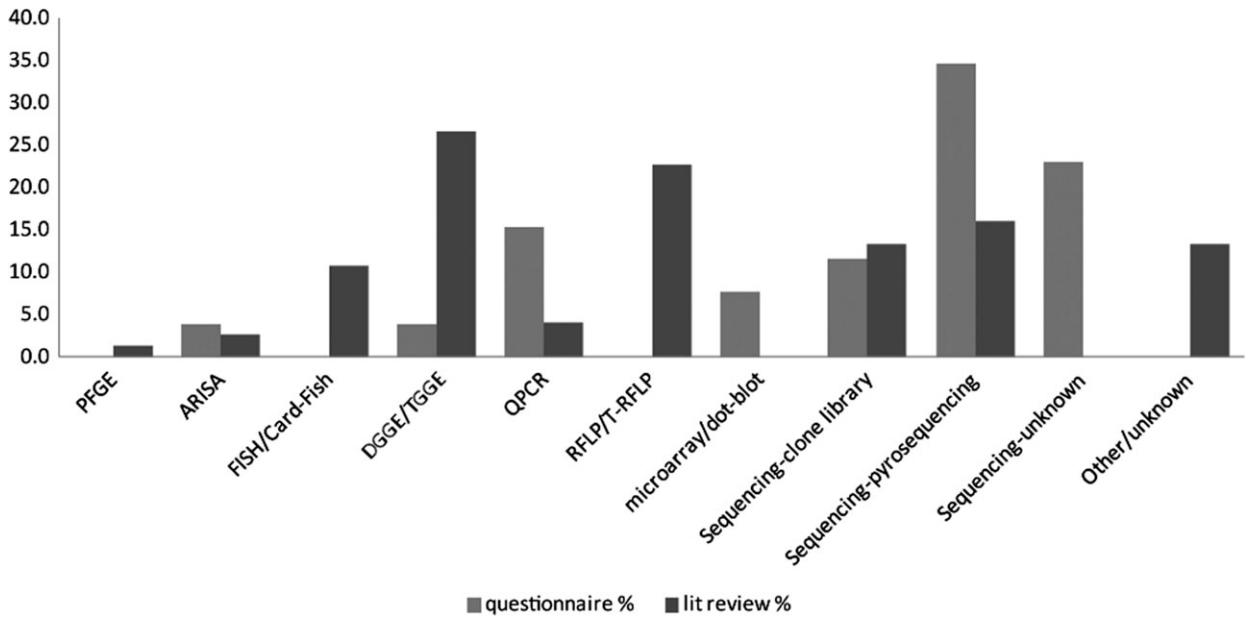


Fig. 4. Comparison of molecular methods used in environmental surveys from literature review (literature review) versus the questionnaire survey. Note how sequencing methods are more prominent in the questionnaire

By contrast, questionnaire respondents answering questions 14–18, had used Sanger sequencing or HTS for analysis (Fig. 4), showing an increased trend to use sequencing as costs have reduced, especially for HTS methods. Only three respondents specified 454 Life sciences technology, likely since it is rapidly being replaced by Illumina technology, and one respondent specified Ion PGM Torrent technology. In some time series, other approaches were used such as Quantitative PCR (qPCR) or molecular fingerprinting techniques (Fig. 4). If HTS was used, sequence libraries were analysed using a combination of publicly available software such as Mothur (Schloss *et al.*, 2009) or QIIME (Caporaso *et al.*, 2010) with tailor-made pipelines in most cases. These pipelines are well used and relatively easy to handle by a trained genetics expert but require specialist knowledge of bioinformatic models so are difficult to utilize without a molecular background to many involved in biodiversity monitoring.

It was pointed out that primer usage, limited sequence quality and ongoing technological progress are major challenges for the application of molecular methods in long-term observations. Previous studies have revealed that molecular diversity measurements can differ by orders of magnitude due to differences in bioinformatics pipelines and variable copy numbers of gene markers that are preferentially detecting some taxa over others (Medinger *et al.*, 2010; Bachy *et al.*, 2012; Egge *et al.*, 2013). Addressing how biomass can be equated to quantity has been achieved for higher trophic levels but less so for microbial eukaryotes is an equally important goal that is started by documenting biovolumes of phytoplankton (Olenina *et al.*, 2006). Automated flow cytometry sorting could be of enormous benefit to this initiative.

The majority of respondents to our questionnaire cited a lack of comprehensive and well-curated reference databases as a major impediment to record biodiversity accurately and reliably. Additionally, the literature survey revealed that variable taxonomic terms (Supplementary Table SI) that would impede cross-comparisons of studies. Most of the interviewed scientists reported that they used their own tailor-made reference databases for annotation of the sequences, set up starting from publically available databases, PR2 or SILVA (Yilmaz *et al.*, 2014). Curated databases such as the Protist Ribosomal Reference Database PR2 and Phytoref (Guillou *et al.*, 2012; Quast, 2013; Decelle, 2014) used by two questionnaire respondents have helped to provide standardized reference sequences and a good example of easily accessible reference database available online curated by experts, which could be accessible to non experts with a degree of training. The UniEuk initiative is an important initiative that could overcome these challenges. This is an expert-driven community initiative to standardize existing and new genetic observations in all biomes with the same pipeline that will link to one taxonomic framework, linked to both SILVA and PR2 (Berney *et al.*, 2017). This initiative could be used as a starting point for LTERs to process and compare their data. Zooplankton studies have used a barcoding cut-off approach helpful in defining

Table I: Genetic and morphological diversity comparison studies

Study	Ref.	Gene Marker	Environment	Target	Morphological method	Number genetic taxa	Number morphological taxa	Correspondence
1*	López-García *et al.* (2003)	18S	Deep sea	Eukaryotic microbes	EM	7	2	Coccolithaceae and Foraminifera correspond
2*	Bachy *et al.* (2012)	18S	Plankton	Tintinid ciliates	LM	9	8	Good correspondence. Tintinnidiidae not observed by LM
3	Santoferra *et al.* (2016)	18S	Plankton	Tintinid ciliates	LM	509	30	Four corresponding groups (Steenstrupiella/Amphorides, Salpingella/Amphorellopsis, Eutintinnus and other tintinnids). Additional four groups found genetically only
4	Santoferra *et al.* (2014)	18S	Plankton	Spirotrichea ciliates	LM	9	4	Good correspondence in four taxon groups
5	Stoeck *et al.* (2014)	18S	Plankton	Ciliates	LM	47	30	Variable. 36 genetically detected genera missing by LM analysis. 19 LM detected genera missing by genetic analysis
6*	Monchy *et al.* (2012)	18S	Plankton	Microplankton	LM	131	37	Ciliates and diatoms show good correspondence, dinoflagellates and phaeocystaceae show greater diversity using genetic vs LM identification
7	De Vargas *et al.* (2015)	18S	Plankton	Eukaryotic microbes	LM	37	37	Chlorachnea, Dictyochophyceae, Collodaria, Ascidiacea, Platyhelminths have more genetically identified taxa. Cephalopoda and Cubozoa have more morphologically described taxa
8	Ruggiero *et al.* (2015)	28S	Plankton	*Pseudo-nitzschia* spp.	LM	17	5	*P. fraudulenta/P. subfraudulenta, P. multistriata* good correspondence. Genetic taxa more diverse than LM-identified taxa
9	Malviya *et al.* (2016)	18S	Plankton	Diatoms	LM	20	20	Good correspondence
10	Pirreda *et al.* (2017)	18S	Plankton	Diatoms and Flagellates	LM	11	11	Good correspondence of taxa for most samples

Three studies (marked with an asterix) were part of the original literature review. LM, light microscopy; EM, electron microscopy.

and linking genetically identified taxa to those of their morphological counterparts (Bucklin *et al.*, 2016). Data reporting varied hugely in our literature review so we would recommend that all processing pipelines should be stored along with the sequences in a common repository like GitHub.

The few studies that did compare identifications molecular and morphological findings revealed variable levels of correspondence (Table I). Diatoms and certain ciliate taxa appear to correspond well (Bachy *et al.*, 2012; Monchy *et al.*, 2012; Pirreda *et al.*, 2017) and this likely relates to the number of morphological features available for identification. Taxa with less defined features, such as Phaeocystales appeared more diverse in genetic studies (Monchy *et al.*, 2012) and therefore only a few studies compare smaller or indistinct taxa using morphological and genetic studies (Collado-Fabbri *et al.*, 2011; Edgcomb *et al.*, 2011; Thiele *et al.*, 2014). Dinoflagellates appeared to be the least congruent study 6 (Table I, Monchy *et al.*, 2012), with genetic diversity outstripping morphological, which the authors attribute to lack of features in lugols-preserved cells, although conversely other dinoflagellate taxa lacked genetic resolution or the presence of indels that biased the dataset. Several comparison studies show variable correspondence of genetic with morphological taxa, in which both methods missed taxa. Studies have cited lack of genetic representation (especially rare types) and genetic resolution that leads to falsely assigned sequences as the cause of this disparity (Monchy *et al.*, 2012; Stoeck *et al.*, 2014; Malviya *et al.*, 2016). There remains a lack of integration between classical taxonomy with genetic identification, when assigning an identity to an unknown sequence. Therefore, despite the enormous diversity captured, a proportion of DNA sequences remain as "unknown eukaryotes". While morphological identity can be useful to relate to environmental factors (Santoferra *et al.*, 2016), genetic studies excel in identifying biological diversity as exemplified by the

TARA oceans global diversity study of microbial eukaryotes that found (De Vargas *et al.*, 2015) unexpectedly higher diversity (38–113 fold) for five eukaryotic classes. The traditional view of morphological diversity dominated by metazoan, larger phyto- and microzooplankton (~88%) was overturned with most taxa belonging to pico- and nanoeukaryotes (De Vargas *et al.*, 2015).

There was a consensus among the interviewed scientists that marine long-term series would benefit from an implementation of molecular methods as standard analyses tools. It was pointed out that molecular methods are well suited to provide refined and reliable information on the whole protist community, since they include the smallest protists, rare taxa, microalgae, ciliates and parasites, including information on intraspecific variability, which is nearly entirely missed in traditional approaches. With respect to this, there was also consensus on the need for a standardized protocol or protocol stages that include all levels of analyses from sampling to data processing, analyses and statistics.

The difficulty of relating cell counts from microscopic observations with genetic data remains. Quantifying cell numbers is essential for reporting toxic species, following species succession, and mapping long-term changes in species distribution. However, PCR amplification creates an initial bias that is reflected in subsequent HTS data. This means that the number of sequences from one taxon cannot be linked directly to the number of cells observed microscopically. Nevertheless, other indirect comparisons can be made such as the relative proportions of sequences within a taxonomic class (Logares *et al.*, 2014). Cell biovolume and sequence numbers can be related within one order of cell size (De Vargas *et al.*, 2015), allowing estimates of biomass and cell carbon to be determined. Quantitative PCR or DNA microarrays can determine abundance for several target microbes (Limardo *et al.*, 2017), but are not realistic for species surveys involving thousands of species. The development of a gene copy database for different species or even higher levels of taxa would facilitate genetically quantifying protists and linking these with microscopic coults of protists, but would not be capable of detecting unknown species.

CONCLUSIONS

Prospects for integrating methods

There are different extents to which one might want to integrate molecular tools into routine monitoring depending on the purpose of the programme. The most extreme, and at the same time most unlikely, scenario for the near to medium term would be to completely replace conventional morphological methods in a time series with molecular ones. This would require a long period of parallel analysis using both methods, to ensure that data from both methods are comparable, but problems with quantification and biomass estimation remain and would need to be addressed. Given the uncertainties and diversity of molecular methods and the rapid developments of these methods, and considering that even long-established methods for enumerating plankton have their own inaccuracies due to collection and counting methods, it appears unlikely that a complete move to molecular methods could work without threatening the temporal integrity of a time series. Even if they are analysed after several years, this would likely not provide sufficient data to make meaningful comparisons. Few time series have integrated molecular and microscopic protist measurements with environmental parameters. However, the collection of molecular data at LTERs has been recognized as a much needed approach, because the ecological knowledge of a site can provide an indispensable background to optimize both the planning and interpretation of molecular research (Davies *et al.*, 2014). Moreover, molecular data provide an insight in the numbers of still unknown species present at a site, fostering taxonomic explorations, which in turn can improve the routine monitoring. Once the eagle-eyed observer is made aware of a taxon's existence, it can often be spotted with ease.

Ocean Sampling Day, which is an initiative to take annual samples on the same day globally using the same methods (Kopf *et al.*, 2015) could provide a valuable reference for comparing different types of data over the long-term whilst UniEuk can provide the framework for consistent data outputs (Berney *et al.*, 2017). There is also a general trend towards integration with other technologies such flow imaging to create a large digital imaging dataset (Olson and Sosik, 2007) that can complement HTS identification. A sustained effort to record species could be advanced through such automated platforms allowing large numbers of species to be separated and then genetically analysed. Figure 5 outlines how various methods could be used to improve protist species identification by combining genetic, morphological and flow cytometric methodologies through the use of combined databases.

The benefits of HTS in detecting diversity are recognized. Advances in sequencing technology could eliminate the need for additional quantification analysis. MinION technology (Oxford nanopore technologies) generates thousands of reads and does not require a PCR amplification step that causes biases in species detection. Single-cell genomics is also becoming increasingly common (Blainey, 2013) and may change how we identify cells, for

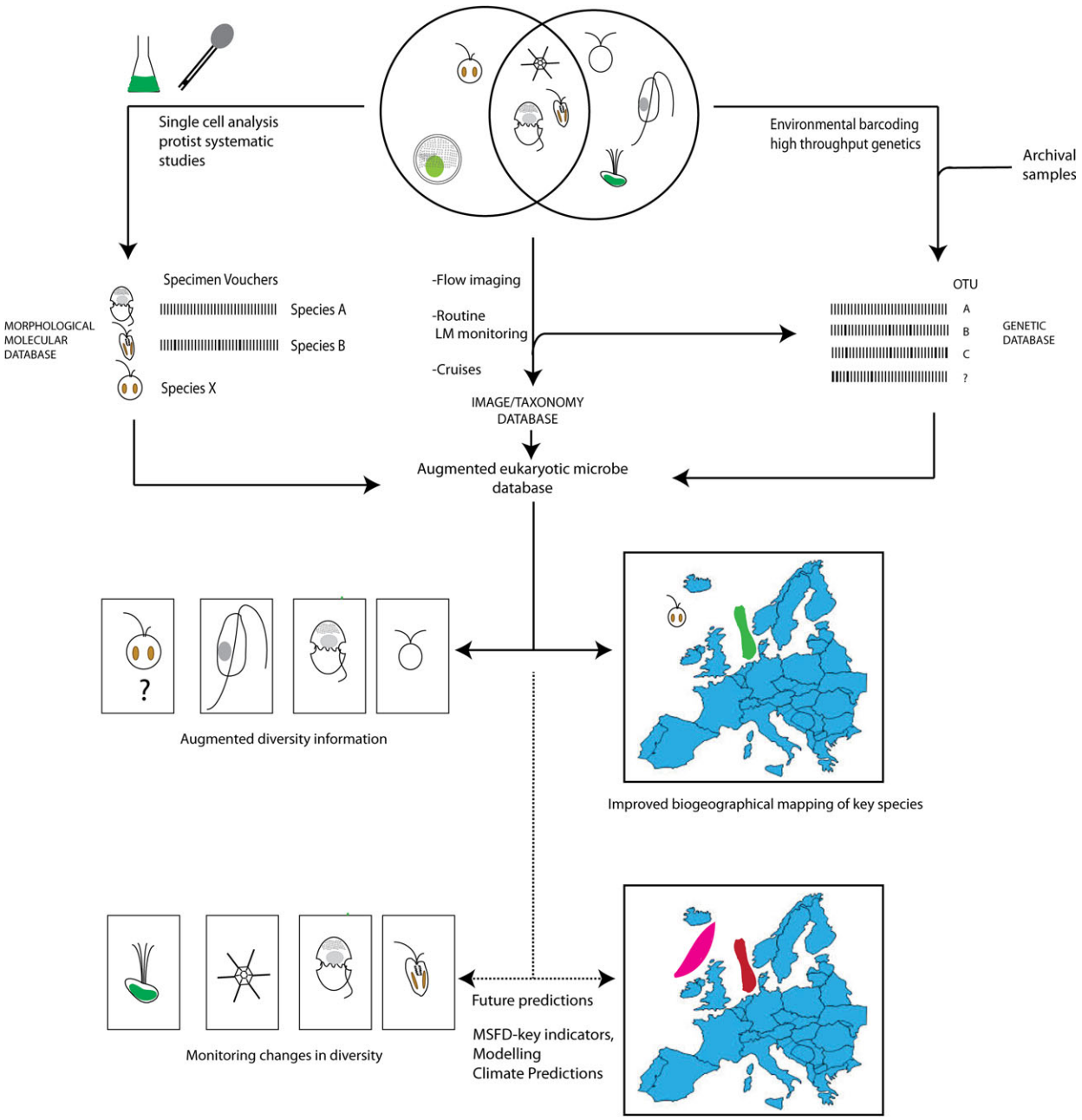


Fig. 5. Hypothetical view of how molecular information can be incorporated into routine monitoring programmes and other databases to augment plankton diversity monitoring- for more accurate long-term predictions of environmental change. New species distributions can be mapped by assigning DNA sequences to unknown cells. Over multiple years, changes in a more comprehensive species assemblage dataset can be used for predictions of biological species shifts.

example, through multiple marker identification. However, the rapid progress in this technological could add to the problem of internal consistency of long-term monitoring using molecular studies. In addition, the increased output still requires curation and the need to integrate data with other biological and oceanographic datasets. Technological advances together with automatization are required and would be a major step forward in finding a more complete estimate of protist diversity on larger scales (Logares *et al.*, 2014; Kilias *et al.*, 2014b; Metfies *et al.*, 2016). These datasets have huge potential in understanding species changes on a global level but need to be stored and curated carefully for the data to be mined usefully.

Interpreting combined datasets

In some instances, molecular information can provide information that can revolutionize the understanding of the dynamics of key phytoplankton species in an area. An example of this is the detection of both the paralytic shellfish toxin producing *Alexandrium catenella* (previosusly *A. tamarense* Group I, Lilly *et al.*, 2007) and the non-toxin producing *Alexandrium tamarense* (previously *A. tamarense* Group III, Lilly *et al.*, 2007) in Scottish waters. Prior to this study, it was believed the distribution of *A. tamarense* in the UK was restricted to England and *A. catenella* was restricted to Scottish waters (Collins *et al.*, 2009) and information gained from this molecular study has fed into the advice that is being given to the aquaculture industry.

Interpretation of long-term datasets collected using microscopy and molecular data needs careful consideration and in some instances, the data may not be compatible. Additional issues come with rapidly evolving genetic technologies. We recommend at least 1 year/seasonal cycle of simultaneous microscopy and sample collection but ideally 2–3 years or seasonal cycles one set to be held in archive for later re-analysis if necessary. The community needs to develop guidelines for experimental procedures to extract genetic material so cross-comparisons can be made. It is advisable to keep the primary sequence data along with the metadata (including original taxonomic assignments) for reprocessing using either different bioinformatic methods. Additionally, improvements to databases both in terms of more accurate taxonomic annotations and novel sequences will make retrospective analysis/correction easier.

One approach that has been successfully used in assessing the state of plankton from a biodiversity perspective is the plankton index approach that has been used to assess the status of plankton in Portugal, France, UK and Hong Kong (Tett *et al.*, 2008; Brito *et al.*, 2015; Gowen *et al.*, 2015; Whyte *et al.*, 2017; Lei *et al.*, 2018). This approach identifies plankton life forms based on ecological function (e.g. diatoms vs dinoflagellates) and using an annual plankton index to identify the change in life form pairs. This approach was developed to overcome sampling and identification disparities to allow comparisons between different laboratories and methods and is currently being used in the assessment of "Good Environmental Status" for the Marine Strategy Framework Directive in Europe. It is possible using a relative abundance score to incorporate molecular methods into this approach and overcome the differences in species level identitifications between the molecular and microscopic methodologies. Molecular methods will increase the life form pairings that can be included in this approach and allow the smaller and microbial components to be investigated, thereby increasing understanding of the ecological functioning of a marine monitoring site.

OUTLOOK

In this review, we looked at current and past practices applied to molecular surveys of microbial eukaryotes. Participants in our questionnaire agreed that molecular surveys are unparalleled in uncovering vast range of undetected diversity for use in ecological models of marine systems. However, we found many areas of ambiguity both in methodological and analytical practices that are preventing researchers and the greater public from gaining the full benefits of these advancements. As such, it would be useful to have a set of best-practice guidelines to integrate molecular surveys into routine morphological plankton surveys.

A dream to be realized for the coming decade would be to miniaturize and automatize HTS methods and imaging methods, link them *in situ*, and broadcast the processed information about species diversity and metadata to the shore in near real-time. The integrated toolkit can be loaded into buoys, gliders, floats and whatever devices still in development, and these automated monitoring tools can then go about their business by the hundreds, providing a detailed picture of the planktonic biodiversity wherever and whenever the observer would like to have a close look.

SUPPLEMENTARY DATA

Supplementary data can be found online at *Journal of Plankton Research* online.

ACKNOWLEDGEMENTS

We wish to thank all participants in this survey, the ICES Working Group for Phytoplankton and Microbial Ecology and a special thanks to Jule Carstens who helped gather data together.

REFERENCES

Adl, S., Habura, A. and Eglit, Y. (2014) Amplification primers of SSU rDNA for soil protists. *Soil Biol. Biochem.*, **69**, 328–342.

Alemzadeh, E., Haddad, R. and Ahmadi, A.-R. (2014) Phytoplanktons and DNA barcoding: characterization and molecular analysis of phytoplanktons on the Persian Gulf. *Iran J Microbiol*, **6**, 296–302.

Amacher, J., Neuer, S. and Lomas, M. (2013) DNA-based molecular fingerprinting of eukaryotic protists and cyanobacteria contributing to sinking particle flux at the Bermuda Atlantic time-series study. *Deep Sea Res. Part II Top. Stud. Oceanogr.*, **93**, 71–83.

Amann, R. I., Binder, B. J., Olson, R. J., Chisholm, S. W., Devereux, R. and Stahl, D. A. (1990) Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. *Appl. Environ. Microbiol.*, **56**, 1919–1925.

Amaral-Zettler, L. A. (2012) Eukaryotic diversity at pH extremes. *Front. Microbiol.*, **3**, 441.

Amaral-Zettler, L., Artigas, L. F., Baross, J., Bharathi, L., Boetius, A., Chandramohan, D., Herndl, G., Kogure, K. *et al.* (2010) A global census of marine microbes. In Mcintyre (ed.), *Life in the World's Oceans: Diversity, Distribution and Abundance*. Blackwell Publishing Ltd, Oxford, pp. 223–245.

Amaral-Zettler, L. A., Mccliment, E. A., Ducklow, H. W. and Huse, S. M. (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.

Atkins, M. S., Teske, A. P. and Anderson, R. O. (2000) A survey of flagellate diversity at four deep-sea hydrothermal vents in the Eastern Pacific Ocean using structural and molecular approaches. *J. Eukaryot. Microbiol.*, **47**, 400–411.

Bachy, C., Dolan, J. R., Lopez-Garcia, P., Deschamps, P. and Moreira, D. (2012) Accuracy of protist diversity assessments: morphology compared with cloning and direct pyrosequencing of 18S rRNA genes and ITS regions using the conspicuous tintinnid ciliates as a case study. *ISME J.*, **7**, 244–255.

Bachy, C., Lopez-Garcia, P., Vereshchaka, A. and Moreira, D. (2011) Diversity and vertical distribution of microbial eukaryotes in the snow, sea ice and seawater near the north pole at the end of the polar night. *Front. Microbiol.*, **2**, 106.

Balzano, S., Marie, D., Gourvil, P. and Vaulot, D. (2012) Composition of the summer photosynthetic pico and nanoplankton communities in the Beaufort Sea assessed by T-RFLP and sequences of the 18S rRNA gene from flow cytometry sorted samples. *ISME J.*, **6**, 1480–1498.

Bazin, P., Jouenne, F., Friedl, T., Deton-Cabanillas, A.-F., Le Roy, B. and Véron, B. (2014) Phytoplankton diversity and community composition along the estuarine gradient of a temperate macrotidal ecosystem: combined morphological and molecular approaches. *PLoS One*, **9**, e94110.

Beaugrand, G., Luczak, C. and Edwards, M. (2009) Rapid biogeographical plankton shifts in the North Atlantic Ocean. *Glob. Change Biol.*, **15**, 1790–1803.

Behnke, A., Barger, K. J., Bunge, J. and Stoeck, T. (2010) Spatio-temporal variations in protistan communities along an O_2/H_2S gradient in the anoxic Framvaren Fjord (Norway). *FEMS Microbiol. Ecol.*, **72**, 89–102.

Berney, C., Ciuprina, A., Bender, S., Brodie, J., Edgcomb, V., Kim, E., Rajan, J., Parfrey, L. W. *et al.* (2017) UniEuk: time to speak a common language in protistology! *J. Eukaryot. Microbiol.*, **64**, 407–411.

Blainey, P. C. (2013) The future is now: single-cell genomics of bacteria and archaea. *FEMS Microbiol. Rev.*, **37**, 1574–6976.

Bowers, H. A., Brutemark, A., Carvalho, W. F. and Graneli, E. (2010) Combining flow cytometry and real-time PCR methodology to demonstrate consumption by Prymnesium parvum. *J. Am. Water Resour. Assoc.*, **46**, 133–143.

Bresnan, E., Cook, K. B., Hindson, J., Hughes, S., Lacaze, J-P., Walsham, P., Webster, L. and Turrell, B. (2016) *The Scottish Coastal Observatory 1997–2013: Part 1—Executive Summary Scottish Marine and Freshwater Science*, **Vol. 7**. Marine Scotland, Scotland.

Brito, A. C., Moita, T., Gameiro, C., Silva, T., Anselmo, T. and Brotas, V. (2015) Changes in the Phytoplankton Composition in a Temperate Estuarine System (1960 to 2010). *Estuaries Coasts*, **38**, 1678–1691.

Bucklin, A., Lindeque, P. K., Rodriguez-Ezpeleta, N., Albaina, A. and Lehtiniemi, M. (2016) Metabarcoding of marine zooplankton: prospects, progress and pitfalls. *J. Plankton Res.*, **38**, 393–400.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Gonzalez Pena, A. *et al.* (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods*, **7**, 335–336.

Chambouvet, A., Morin, P., Marie, D. and Guillou, L. (2008) Control of toxic marine dinoflagellate blooms by serial parasitic attack. *Science*, **322**, 1254–1257.

Collado-Fabbri, S., Vaulot, D. and Ulloa, O. (2011) Structure and seasonal dynamics of the eukaryotic picophytoplankton community in a wind-driven coastal upwelling ecosystem. *Limnol. Oceanogr.*, **56**, 2334–2346.

Collins, C., Graham, J., Brown, L., Bresnan, E., Lacaze, J. P. and Turrell, E. A. (2009) Identification and toxicity of Alexandrium tamarense (Dinophyceae) in Scottish waters. *J. Phycol.*, **45**, 692–703.

Countway, P. D., Gast, R. J., Dennett, M. R., Savai, P., Rose, J. M. and Caron, D. A. (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.

Countway, P. D., Gast, R. J., Savai, P. and Caron, D. A. (2005) Protistan diversity estimates based on 18S rDNA from seawater incubations in the Western North Atlantic1. *J. Eukaryot. Microbiol.*, **52**, 95–106.

Culverhouse, P. F. (2015) Biological oceanography needs new tools to automate sample analysis. *J. Mar. Biol. Aquacult.*, **1**, 1–2.

Cusack, C., Dabrowski, T., Lyons, K., Berry, A., Westbrook, G., Salas, R., Duffy, C., Nolan, G. and Silke, J. (2016) Harmful algal bloom forecast system for SW Ireland. Part II: Are operational oceanographic models useful in a HAB warning system? *Harmful Algae*, **53**, 86–101.

Davies, N., Field, D., Amaral-Zettler, L., Clark, M. S., Deck, J. and Drummond, A. (2014) The founding charter of the genomic observatories network. *Gigascience*, **3**, 2.

Decelle, J., Romac, S., Sasaki, E., Not, F. and Mahé, F. (2014) Intracellular diversity of the V4 and V9 regions of the 18S rRNA in marine protists (radiolarians) assessed by high-throughput sequencing. *PLoS One*, **9**, e104297.

De Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., Lara, E., Berney, C. *et al.* (2015) Eukaryotic plankton diversity in the sunlit ocean. *Science*, **348**, 1261605_1_10.

Diercks, S., Metfies, K. and Medlin, L. K. (2008) Development and adaptation of a multiprobe biosensor for the use in a semi-automated device for the detection of toxic algae. *Biosens. Bioelectron.*, **23**, 1527–1533.

Díez, B., Pedrós-Alió, C., Marsh, T. L. and Massana, R. (2001) Application of denaturing gradient gel electrophoresis (DGGE) to study the diversity of marine picoeukaryotic assemblages and

comparison of DGGE with other molecular techniques. *Appl. Environ. Microbiol.*, **67**, 2942–2951.

Diéz, B., Pedrós-Alió, C. and 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.

Dolven, J. K., Lindqvist, C., Albert, V. A., Bjørklund, K. R., Yuasa, T., Takahashi, O. and Mayama, S. (2007) Molecular Diversity of Alveolates Associated with Neritic North Atlantic radiolarians. *Protist*, **158**, 65–76.

Dubilier, N., Mcfall-Ngai, M. and Zhao, L. (2015) Microbiology: create a global microbiome effort. *Nature*, **526**, 631–634.

Edgcomb, V., Orsi, W., Bunge, J., Jeon, S. and Christen, R. E. A. (2011) Protistan microbial observatory in the Cariaco Basin, Caribbean. I. Pyrosequencing vs Sanger insights into species richness. *ISME J.*, **5**, 1344–1356.

Egge, E., Bittner, L., Andersen, T., Audic, S., De Vargas, C. and Edvardsen, B. (2013) 454 pyrosequencing to describe microbial eukaryotic community composition, diversity and relative abundance: a test for marine haptophytes. *PLoS One*, **8**, e74371.

Emmett Duffy, J., Amaral-Zettler, L., Fautin, D. G., Pauley, G., Rynearson, T. A., Sosik, H. M. and Stachowitz, J. J. (2013) Envisioning a marine biodiversity observation network. *Bioscience*, **63**, 350–361.

Epstein, S. and Lopez-Garcia, P. (2008) 'Missing' protists: a molecular perspective. *Biodivers. Conserv.*, **17**, 261–276.

Filker, S., Stock, A., Breiner, H.-W., Edgcomb, V., Orsi, W., Yakimov, M. M. and Stoeck, T. (2013) Environmental selection of protistan plankton communities in hypersaline anoxic deep-sea basins, Eastern Mediterranean Sea. *Microbiologyopen*, **2**, 54–63.

Gao, Z., Johnson, Z. I. and Wang, G. (2010) Molecular characterisation of the spatial diversity and novel lineages of mycoplankton in Hawaiian coastal waters. *ISME J.*, **4**, 111–120.

Gowen, R. J., Collos, Y., Tett, P., Scherer, C., Bec, B., Abadie, E., Allen, M. and O'brien, T. (2015) Response of diatom and dinoflagellate lifeforms to reduced phosphorus loading: a case study in the Thau lagoon, France. *Estua. Coast. Shelf Sci.*, **162**, 45–52.

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

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., Boutte, C., Burgaud, G. *et al.* (2012) The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. *Nucl. Acids Res.*, **41**, D597–D604.

Hays, G. C., Richardson, A. J. and Robinson, C. (2005) Climate change and marine plankton. *Trends Ecol. Evol.*, **20**, 337–344.

He, L., Liu, F., Karuppiah, V., Ren, Y. and Li, Z. (2014) Comparisons of the fungal and protistan communities among different marine sponge holobionts by pyrosequencing. *Microb. Ecol.*, **67**, 951–961.

Heywood, J. L., Sieracki, M. E., Bellows, W., Poulton, N. J. and Stepanauskas, R. (2010) Capturing diversity of marine heterotrophic protists: one cell at a time. *ISME J.*, **5**, 674.

Hugerth, L. W., Muller, E. E. L., Hu, Y. O. O., Lebrun, L. A. M., Roume, H., Lundin, D., Wilmes, P. and Andersson, A. F. (2014) Systematic design of 18S rRNA gene primers for determining eukaryotic diversity in microbial consortia. *PLoS One*, **9**, e95567.

Ishitani, Y., Ujiié, Y. and Takishita, K. (2014) Uncovering sibling species in radiolaria: evidence for ecological partitioning in a marine planktonic protist. *Mol. Phylogenet. Evol.*, **78**, 215–222.

Jeuck, A., Nitsche, F., Wylezich, C., Wirth, O., Bergfeld, T., Brutscher, F., Hennemann, M., Monir, S. *et al.* (2017) A comparison of methods to analyze aquatic heterotrophic flagellates of different taxonomic groups. *Protist*, **168**, 375–391.

Jiang, Z.-Y., Wang, Y.-S. and Sun, F.-L. (2014) Spatial structure of eukaryotic ultraplankton community in the northern South China Sea. *Biologia*, **69**, 557–565.

Joo, S., Lee, S.-R. and Park, S. (2010) Monitoring of phytoplankton community structure using terminal restriction fragment length polymorphism (T-RFLP). *J. Microbiol. Methods*, **81**, 61–68.

Kilias, E., Kattner, G., Wolf, C., Frickenhaus, S. and Metfies, K. (2014a) A molecular survey of protist diversity through the central Arctic Ocean. *Polar Biol.*, **37**, 1271–1287.

Kilias, E. S., Peeken, I. and Metfies, K. (2014b) Insight into protist diversity in Arctic sea ice and melt-pond aggregate obtained by pyrosequencing. *Polar Res.*, **33**, 23466.

Kim, D. Y., Countway, P. D., Jones, A. C., Schnetzer, A., Yamashita, W., Tung, C. and Caron, D. A. (2013) Monthly to interannual variability of microbial eukaryote assemblages at four depths in the eastern North Pacific. *ISME J.*, **8**, 515.

Koid, A., Nelson, W. C., Mraz, A. and Heidelberg, K. B. (2012) Comparative analysis of eukaryotic marine microbial assemblages from 18S rRNA gene and gene transcript clone libraries by using different methods of extraction. *Appl. Environ. Microbiol.*, **78**, 3958–3965.

Kopf, A., Bicak, M., Kottmann, R., Schnetzer, J., Kostadinov, I., Lehmann, K., Fernandez-Guerra, A., Jeanthon, C. *et al.* (2015) The ocean sampling day consortium. *GigaScience*, **4**, 27.

Kraberg, A. C., Rodriguez, N. and Salewski, C. R. (2015) Historical phytoplankton data from Helgoland Roads: Can they be linked to modern time series data? *J. Sea Res.*, **101**, 51–58.

Larsen, A., Castberg, T., Sandaa, R. A., Brussaard, C. P. D., Egge, J., Heldal, M., Paulino, A., Thyrhaug, R. *et al.* (2001) Population dynamics and diversity of phytoplankton, bacteria and viruses in a seawater enclosure. *Mar. Ecol. Prog. Ser.*, **221**, 47–57.

Latasa, M., Scharek, R., Le Gall, F. and Guillou, L. (2004) Pigment suites and taxonomic groups in Prasinophyceae. *J. Phycol.*, **40**, 1149–1155.

Lee, S.-R., Oak, J. H., Chung, I. K. and Lee, J. A. (2010) Effective molecular examination of eukaryotic plankton species diversity in environmental seawater using environmental PCR, PCR-RFLP, and sequencing. *J Appl. Phycol.*, **22**, 699–707.

Lei, Y., Whyte, C., Davidson, K., Tett, P. and Yin, K. (2018) A change in phytoplankton community index with water quality improvement in Tolo Harbour, Hong Kong. Mar Pollut Bull, in press.

Lilly, E. L., Halanych, K. M. and Anderson, D. M. (2007) Species boundaries and global biogeography of the *Alexandrium tamarense* complex (Dinophyceae). *J. Phycol.*, **43**, 1329–1338.

Limardo, A. J., Sudek, S., Choi, C. J., Poirier, C., Rii, Y. M., Blum, M., Roth, R., Goodenough, U., Church, M. J. and Worden, A. Z. (2017) Quantitative biogeography of picoprasinophytes establishes

ecotype distributions and significant contributions to marine phytoplankton. *Environ. Microbiol.*, **19**, 3219–3234.

Logares, R., Audic, S., Bass, D., Bittner, L., Boutte, C., Christen, R., Claverie, J.-M., Decelle, J. *et al.* (2014) Patterns of rare and abundant marine microbial eukaryotes. *Curr. Biol.*, **24**, 813–821.

López-García, P., Philippe, H., Gail, F. and Moreira, D. (2003) Autochthonous eukaryotic diversity in hydrothermal sediment and experimental microcolonizers at the Mid-Atlantic Ridge. *Proc. Natl Acad. Sci.*, **100**, 697–702.

López-García, P., Rodríguez-Valera, F., Pedrós-Alió, C. and Moreira, D. (2001) Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. *Nature*, **409**, 603–607.

López-García, P., Vereshchaka, A. and Moreira, D. (2007) Eukaryotic diversity associated with carbonates and fluid–seawater interface in Lost City hydrothermal field. *Environ. Microbiol.*, **9**, 546–554.

Lovejoy, C., Massana, R. and Pedrós-Alió, C. (2006) Diversity and distribution of marine microbial eukaryotes in the Arctic Ocean and adjacent seas. *Appl. Environ. Microbiol.*, **72**, 3085–3095.

Lund, J. W. G., Kipling, C. and Le Cren, E. D. (1958) The inverted microscope method of estimating algal numbers and the statistical basis of estimations by counting. *Hydrobiologia*, **11**, 143–170.

Malviya, S., Scalco, E., Audic, S., Vincent, F., Veluchamy, A., Poulain, J., Wincker, P., Iudicone, D. *et al.* (2016) Insights into global diatom distribution and diversity in the world's ocean. *Proc. Natl Acad. Sci. USA*, **113**, E1516–E1525.

Mamanova, L., Coffey, A., Scott, C., Kozarewa, I. and Turner, E. (2010) Target-enrichmnet strategies for next-generation sequencing. *Nat. Methods*, **7**, 111–118.

Marie, D., Zhu, F., Balague, V., Ras, J. and Vaulot, D. (2006) Eukaryotic picoplankton communities of the Mediterranean Sea in summer assessed by molecular approaches (DGGE, TTGE, QPCR). *FEMS Microbiol. Ecol.*, **55**, 403–415.

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. and Pedros-Alio, C. (2008) Unveiling new microbial eukaryotes in the surface ocean. *Curr. Opin. Microbiol.*, **11**, 213–218.

Massana, R., Terrado, R., Forn, I., Lovejoy, C. and Pedrós-Alió, C. (2006) Distribution and abundance of uncultured heterotrophic flagellates in the world oceans. *Environ. Microbiol.*, **8**, 1515–1522.

Medinger, R., Nolte, V., Pandey, R. V., Jost, S., Ottenwälder, B., Schlötterer, C. and Boenigk, J. (2010) Diversity in a hidden world: potential and limitation of next generation sequencing for surveys of molecular diversity of eukaryotic organisms. *Mol. Ecol.*, **19**, 32–40.

Medlin, L., Elwood, H. J., Stickel, S. and Sogin, M. L. (1988) The characterization of enzymatically amplified eukaryotic 16S-like rRNA-coding regions. *Gene*, **71**, 491–499.

Medlin, L. K., Metfies, K., Mehl, H., Wiltshire, K. and Valentin, K. (2006) Picoeukaryotic Plankton Diversity at the Helgoland Time Series Site as Assessed by Three Molecular Methods. *Microb. Ecol.*, **52**, 53–71.

Metfies, K., Schroeder, F., Hessel, J., Wollschläger, J., Micheller, S., Wolf, C., Kilias, E., Sprong, P. *et al.* (2016) High-resolution monitoring of marine protists based on an observation strategy integrating automated on-board filtration and molecular analyses. *Ocean Sci.*, **12**, 1237–1247.

Moffat, C., Aish, A., Hawkridge, J. M., Miles, H., Mitchell, I., Mcquatters-Gollop, A., Frost, M., Greenstreet, S. *et al.* (2011) Advice on United Kingdom biodiversity indicators and targets for the Marine Strategy Framework Directive. *Report to the Department for Environment, Food and Rural Affairs. H. a. B. D. S. E. Group, London.* pp. 207.

Monchy, S., Grattepanche, J.-D., Breton, E., Meloni, D. and Sanciu, G. E. A. (2012) Microplanktonic community structure in a coastal system relative to a phaeocystis bloom inferred from morphological and tag pyrosequencing methods. *PLoS One*, **7**, e39924.

Monier, A., Comte, J., Babin, M., Forest, A., Matsuoka, A. and Lovejoy, C. (2014) Oceanographic structure drives the assembly processes of microbial eukaryotic communities. *ISME J.*, **9**, 990.

Monier, A., Sudek, S., Fast, N. M. and Worden, A. (2013) Gene invasion in distant eukaryotic lineages: discovery of mutually exclusive genetic elements reveals marine biodiversityI. *ISME J.*, **7**, 1764–1774.

Moon-Van Der Staay, S. Y., De Wachter, R. and Vaulot, D. (2001) Oceanic 18S rDNA sequences from picoplankton reveal unsuspected eukaryotic diversity. *Nature*, **409**, 607–610.

Moon-Van Der Staay, S. Y., Van Der Staay, G. W., Guillou, L., Vaulot, D., Claustre, H. and Medlin, L. K. (2000) Abundance and diversity of prymnesiophytes in the picoplankton community from the equatorial Pacific Ocean inferred from 18S rDNA sequences. *Limnol. Oceanogr.*, **45**, 98–109.

Motes, M. L., Depaola, A., Cook, D. W., Veazey, J. E., Hunsucker, J. C., Garthright, W. E., Blodgett, R. J. and Chirtel, S. J. (1998) Influence of water temperature and salinity on Vibrio vulnificus in Northern Gulf and Atlantic Coast oysters (*Crassostrea virginica*). *Appl. Environ. Microbiol.*, **64**, 1459–1465.

Muhling, B. A., Jacobs, J., Stock, C. A., Gaitan, C. F. and Saba, V. S. (2017) Projections of the future occurrence, distribution, and seasonality of three Vibrio species in the Chesapeake Bay under a high-emission climate change scenario. *GeoHealth*, **1**, 278–296.

Not, F., Gausling, R., Azam, F., Heidelberg, J. F. and Worden, A. Z. (2007a) Vertical distribution of picoeukaryotic diversity in the Sargasso Sea. *Environ. Microbiol.*, **9**, 1233–1252.

Not, F., Latasa, M. and Scharek, R. E. A. (2008) Protistan assemblages across the Indian Ocean, with a specific emphasis on the picoeukaryotes. *Deep Sea Res. Part II Top. Stud. Oceanogr.*, **55**, 1456–1473.

Not, F., Valentin, K., Romari, K., Lovejoy, C., Massana, R., Toebe, K., Vaulot, D. and Medlin, L. K. (2007b) Picobiliphytes: a marine picoplanktonic algal group with unknown affinities to other eukaryotes. *Science*, **315**, 253–255.

Olenina, I., Hajdu, S., Edler, L., Andersson, A., Wasmund, N., Busch, S., Göbel, J., Gromisz, S. *et al.* (2006) Biovolumes and size-classes of phytoplankton in the Baltic Sea. *HELCOM Balt. Sea Environ. Proc.*, **106**, 144.

Olson, R. J. and Sosik, H. M. (2007) A submersible imaging-in-flow instrument to analyze nano and mciroplankton: Imaging FlowCytobot. *Limnol. Oceanogr. Meth.*, **5**, 195–203.

Pawlowski, J., Audic, S., Adl, S., Bass, D., Belbahr, I. L., Berney, C., Bowser, S. S., Cepicka, I. *et al.* (2012) CBOL protist working group: barcoding eukaryotic richness beyond the animal, plant, and fungal kingdoms. *PLoS Biol.*, **10**, e1001419.

Piquet, A. M.-T., Scheepens, J. F., Bolhuis, H., Wiencke, C. and Buma, A. G. J. (2010) Variability of protistan and bacterial communities in two Arctic fjords (Spitsbergen). *Polar Biol.*, **33**, 1521–1536.

Piquet, A. M. T., Van De Poll, W. H., Visser, R. J. W., Wiencke, C., Bolhuis, H. and Buma, A. G. J. (2014) Springtime phytoplankton

dynamics in Arctic Krossfjorden and Kongsfjorden (Spitsbergen) as a function of glacier proximity. *Biogeosciences*, **11**, 2263–2279.

Pirreda, R., Tomasino, M. P., D'erchia, A. M., Manzari, C., Pesole, G., Montresor, M., Kooistra, W. H., Sarno, D. *et al.* (2017) Diversity and temporal patterns of planktonic protist assemblages at a Mediterranean Long Term Ecological Research site. *FEMS Microbiol. Ecol.*, **93**, 1–14.

Piwosz, K. and Pernthaler, J. (2011) Enrichment of omnivorous cercozoan nanoflagellates from coastal Baltic Sea waters. *PLoS One*, **6**, e24415.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J. and Glöckner, F.O. (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucl. Acids Res.*, **41**, D590–D596.

Reid, P. C., Colebrook, J. M., Matthews, J. B. L. and Aiken, J., Continuous Plankton Recorder Team (2003) The Continuous Plankton Recorder: concepts and history, from Plankton Indicator to undulating recorders. *Prog. Oceanogr.*, **58**, 117–173.

Reid, P. C., Edwards, M., Hunt, H. G. and Warner, A. J. (1998) Phytoplankton change in the North Atlantic. *Nature*, **391**, 546.

Rhodes, L., Smith, K. and Moisan, C. (2013) Shifts and stasis in marine HAB monitoring in New Zealand. *Environ. Sci. Pollut. Res. Int.*, **20**, 6872–6877.

Romari, K. and Vaulot, 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.

Ruggiero, M. V., Sarno, D., Barra, L., Kooistra, W. H. C. F., Montresor, M. and Zingone, A. (2015) Diversity and temporal pattern of *Pseudo-nitzschia* species (Bacillariophyceae) through the molecular lens. *Harmful Algae*, **42**, 15–24.

Santoferra, L. F., Grattepanche, J. D., Katz, L. A. and Mcmanus, G. B. (2014) Pyrosequencing for assessing diversity of eukaryotic microbes: analysis of data on marine planktonic ciliates and comparison with traditional methods. *Environ. Microbiol.*, **16**, 2752–2763.

Santoferra, L. F., Grattepanche, J. D., Katz, L. A. and Mcmanus, G. B. (2016) Patterns and processes in microbial biogeography: do molecules and morphologies give the same answers? *ISME J.*, **10**, 1779.

Savin, M. C., Martin, J. L., Legresley, M., Giewat, M. and Rooney-Varga, J. (2003) Plankton diversity in the Bay of Fundy as measured by morphological and molecular methods. *Microb. Ecol.*, **48**, 51–65.

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R. and Hartmann, M. (2009) Introducing mothur: open source, patform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.*, **75**, 7537–7541.

Schnetzer, A., Moorthi, S. D., Countway, P. D., Gast, R. J., Gilg, I. C. and Caron, D. A. (2011) Depth matters: microbial eukaryote diversity and community structure in the eastern North Pacific revealed through environmental gene libraries. *Deep Sea Res. Pt I Oceanogr. Res. Paper*, **58**, 16–26.

Scholz, M. B., Lo, C. C. and Chain, P. S. G. (2012) Next generation sequencing and bioinformatic bottleneck: the current state of metagenomic data analysis. *Curr. Opin. Biotechnol.*, **23**, 9–15.

Seears, H. A., Darling, K. F. and Wade, C. M. (2012) Ecological partitioning and diversity in tropical planktonic foraminifera. *BMC Evol. Biol.*, **12**, 54.

Shi, X. L., Lepère, C., Scanlan, D. J. and Vaulot, D. (2011) Plastid 16S rRNA gene diversity among eukaryotic picophytoplankton sorted by flow cytometry from the South Pacific Ocean. *PLoS One*, **6**, e18979.

Short, S. M. and Suttle, C. A. (2003) Temporal dynamics of natural communities of marine algal viruses and eukaryotes. *Aquat. Microb. Ecol.*, **32**, 107–119.

Smyth, R. P., Despons, L., Huili, G., Bernacchi, S., Hijnen, M., Mak, J., Jossinet, F., Weixi, L. *et al.* (2015) Mutational interference mapping experiment (MIME) for studying RNA structure and function. *Nat. Methods*, **12**, 866–872.

Sorte, C. J. B., Williams, S. L. and Carlton, J. T. (2010) Marine range shifts and species introductions: comparative spread rates and community impacts. *Glob. Ecol. Biogeogr.*, **19**, 303–316.

Sosik, H. M., Olson, R. J. and Ambrust, E. V. (2010) Flow cytometry in Phytoplankton research. In Suggett, D. J., Prášil, O. and Borowitzka, M. (eds), *Chlorophyll a Fluorescence in Aquatic Sciences: Methods and Applications, Developments in Applied Phycology*, Vol. 4. Springer, Dordrecht, pp. 171–185.

Sournia, A. (1978) *Phytoplankton Manual*. **Vol. 6**. UNESCO, Paris, p. 337.

Southward, A. J., Langmead, O., Hardman-Mountford, N. J., Aiken, J., Boalch, G. T., Dando, P. R., Genner, M. J., Joint, I. *et al.* (2004) Long-term oceanographic and ecological research in the western English Channel. *Adv. Mar. Biol.*, **47**, 1–105.

Stern, R., Horak, A., Andrew, R., Andersen, R. A., Kuepper, F., Coffroth, M. A., Jameson, I., Hoppenrath, M. *et al.* (2010) Environmental barcoding reveals massive dinoflagellate diversity in marine environments. *PLoS One*, **5**, e13991.

Stern, R. F., Picard, K., Hamilton, K. M., Walne, A., Tarran, G. A., Mills, D., Mcquatters-Gollop, A. and Edwards, M. (2015) An automated water sampler from Ships of Opportunity detects new boundaries of marine microbial biodiversity. *Prog. Oceanogr.*, **137**, 409–420.

Stoeck, T., Behnke, A., Christen, R., Amaral-Zettler, L., Rodriguez-Mora, M. J., Chistoserdov, A., Orsi, W. and Edgcomb, V. P. (2009) Massively parallel tag sequencing reveals the complexity of anaerobic marine protistan communities. *BMC Biol.*, **7**, 72.

Stoeck, T., Briener, H.-W., Filker, S., Ostermaier, V., Kammerlander, B. and Sonntag, B. (2014) A morphogenetic survey on ciliate plankton from a mountain lake pinpoints the necessity of lineage-specific barcode markers in microbial ecology. *Environ. Microbiol.*, **16**, 430–444.

Sunagawa, S., Coelho, L. P., Chaffron, S., Kultima, J. R., Labadie, K., Salazar, G., Djahanschiri, B., Zeller, G. *et al.* (2015) Structure and function of the global ocean microbiome. *Science*, **348**, 1261359.

Taylor, J. and Cunliffe, M. (2014) High throughput sequencing reveals neustonic and planktonic microbial eukaryotic diversity in coastal waters. *J. Phycol.*, **50**, 960–965.

Terrado, R., Vincent, W. F. and Lovejoy, C. (2009) Mesopelagic protists: diversity and succession in a coastal Arctic ecosystem. *Aquat. Microbial. Ecol.*, **56**, 25–39.

Tett, P., Carreira, C., Mills, D. K., Van Leeuwen, S., Foden, J., Bresnan, E. and Gowen, R. J. (2008) Use of a Phytoplankton Community Index to assess the health of coastal waters. *ICES J. Mar. Sci.*, **65**, 1475–1482.

Thiele, S., Wolf, C., Schulz, I. K., Assmy, P., Metfies, K. and Fuchs, B. M. (2014) Stable composition of the nano- and picoplankton community during the ocean iron fertilization experiment LOHAFEX. *PLoS One*, **9**, e113244.

Toebe, K., Alpermann, T. J., Tillmann, U., Krock, B., Cembella, A. and John, U. (2013) Molecular determination of toxic and non-toxic Alexandrium species (Dinophyta) in natural phytoplankton assemblages from the Scottish Coast of the North Sea. *Eur. J. Phycol.*, **48**, 12–26.

Ussler, B., Preston, C., Tavormina, P., Pargett, D., Jensen, S., Roman, B., Marin, R., Sunita, R. *et al.* (2013) Autonomous application of quantitative PCR in the deep sea: in situ surveys of aerobic methanotrophs using the deep-sea environmental sample processor. *Environ. Sci. Technol.*, **47**, 9339–9346.

Utermöhl, H. (1931) Neue Wege in der quantitativen Erfassung des Planktons (mit besonderer Berücksichtigung des Ultraplanktons). *Verh. Int. Verein. Theor. Angew. Limnol.*, **5**, 567–596.

Van Hoek, A. H., Van Alen, T. A., Sprakel, V. S., Hackstein, J. H. and Vogels, G. D. (1998) Evolution of anaerobic ciliates from the gastrointestinal tract: phylogenetic analysis of the ribosomal repeat from *Nyctotherus ovalis* and its relatives. *Mol. Biol. Evol.*, **15**, 1195–1206.

Vaulot, D., Romari, K. and Not, F. (2002) Are autotrophs less diverse than heterotrophs in marine picoplankton? *Trends Microbiol.*, **10**, 266–267.

Vezzulli, L., Grande, C., Reid, P. C., Helaouet, P., Edwards, M., Höfle, M. G., Brettar, I., Colwell, R. R. *et al.* (2016) Climate influence on Vibrio an associated human diseases during the past half-century in the coastal Atlantic. *Proc. Natl Acad. Sci. USA*, **113**, E5062–E5071.

Viprey, M., Guillou, L., Ferréol, M. and Vaulot, D. (2008) Wide genetic diversity of picoplanktonic green algae (Chloroplastida) in the Mediterranean Sea uncovered by a phylum-biased PCR approach. *Environ. Microbiol.*, **10**, 1804–1822.

Weber, A. A.-T. and Pawlowski, J. (2014) Wide occurrence of SSU rDNA intragenomic polymorphism in foraminifera and its implications for molecular species identification. *Protist*, **165**, 645–661.

Wells, M. L., Trainer, V. L., Smayda, T. J., Karlson, B. S. O., Trick, C. G., Kudela, R. M., Ishikawa, A., Bernard, S. *et al.* (2015) Harmful algal blooms and climate change: learning from the past and present to forecast the future. *Harmful Algae*, **49**, 68–93.

White, T. J., Bruns, T., Lee, S. and Taylor, J. (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In Innis, M. A., Gelfand, D. H., Sninsky, J. J. and White, T. J. (eds), *PCR Protocols: A Guide to Methods and Applications*. Academic Press, New York, U.S.A., pp. 315–322.

Whyte, C., Davidson, K., Gilpin, L., Mitchell, E., Moschonas, G., Mcneill, S. and Tett, P. (2017) Tracking changes to a microplankton community in a North Atlantic sea loch using the microplankton index PI (mp). *ICES J. Mar. Sci*, **74**, 311–325.

Woese, C. R. (1987) Bacterial evolution. *Microbiol. Rev.*, **51**, 221–271.

Wolf, C., Frickenhaus, S., Kilias, E. S., Peeken, I. and Metfies, K. (2014) Protist community composition in the Pacific sector of the Southern Ocean during austral summer 2010. *Polar Biol.*, **37**, 375–389.

Wylezich, C. and Jürgens, K. (2011) Protist diversity in suboxic and sulfidic waters of the Black Sea. *Environ. Microbiol.*, **13**, 2939–2956.

Yilmaz, P., Parfrey, L. W., Yarza, P., Gerken, J., Pruesse, E., Quast, C., Schweer, T., Peplies, J. *et al.* (2014) The SILVA and "all-species Living Tree Project (LTP)" taxonomic frameworks. *Nucleic Acids Res.*, **42**, D643–D648.

Yuasa, T., Takahashi, O., Dolven, J. K., Mayama, S., Matsuoka, A., Honda, D. and Bjørklund, K. R. (2006) Phylogenetic position of the small solitary phaeodarians (Radiolaria) based on 18S rDNA sequences by single cell PCR analysis. *Mar. Micropaleontol.*, **59**, 104–114.

Zhu, F., Massana, R., Not, F., Marie, D. and Vaulot, D. (2005) Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. *FEMS Microbiol. Ecol.*, **52**, 79–92.

Zingone, A. LTER Marechiara—ItalyDynamic Ecological Information Management System—Site and Dataset Registry. Vol. 2018. NERC CEH and DEIMS core.

Zingone, A., Harrison, P. J., Kraberg, A. C., Lehtinen, S., Mcquatters-Gollop, A., O'brien, T., Sun, J. and Jakobsen, H. H. (2015) Increasing the quality, comprability and accessibility of phytoplankton species composition time-series data. *Estuar. Coast. Shelf Sci.*, **162**, 151–160.