Unveiling new diversity and connections in the planktonic food web using metabarcoding

Dissertation

zur Erlangung des Doktorgrades (Dr. rer. nat)

der Mathematisch-Naturwissenschaftlichen Fakultät

der Christian-Albrechts-Universität zu Kiel

vorgelegt von

Laura Käse

Kiel 2022

Erste/r Gutacher/in: Prof. Dr. Karen Wiltshire

Zweite/r Gutachter/in: PD Dr. Bernhard M. Fuchs

Tag der mündlichen Prüfung: 18.05.2022

To my parents

I Table of contents

I Table of contents i
II List of publications and author contributions iii
III List of abbreviations vii
IV Abstract ix
V Zusammenfassung x
1 Introduction 1
The marine planktonic food web 2
Seasonality of plankton in temperate environments 6
2 Objectives 9
3 Plankton monitoring 11
Study area of this study 12
Light microscopy 13
Metabarcoding 16
4 Publication I 21
5 Publication II 53
6 Publication III 95
7 Publication IV 129
8 Discussion 163
What is the specific added value of using metabarcoding to study the ecology of the eukaryotic plankton community? 164
Is the information obtained through metabarcoding comparable to that of conventional counting methods? 171
Is it feasible and scientifically reasonable to integrate metabarcoding into an existing long-term time series, and how can results of different methods be combined for the best added value in long-term time series? 176
9 Conclusion and outlook 179
10 References 181
A Appendix ..... a
Supplementary Material to Publication I ..... a
Supplementary Material to Publication II ..... d
Supplementary Material to Publication III ..... k
Supplementary Material to Publication IV ..... t
Abstracts for Publications V and VI ii
B Acknowledgements ..... kk
C Eidesstattliche Erklärung ..... mm

II List of publications and author contributions

This doctoral thesis is based on the reprints of three research papers, which have been published in internationally peer-reviewed journals (**Publication I-III**), and one manuscript, which is in preparation (**Publication IV**). All published content is unchanged and merely the formatting has been adapted to the general format of this thesis. Two internationally peer-reviewed publications (**Publication V** and **VI**) have also been written separately and published during the course of this project. They are partially related to and referenced in this thesis but are not included as separate chapters. The abstracts to these two publications can be found in the annex of this thesis.

Publication I:

Laura Käse, Alexandra C. Kraberg, Katja Metfies, Stefan Neuhaus, Pim A.A. Sprong, Bernhard M. Fuchs, Maarten Boersma, Karen H. Wiltshire. (2020) **Rapid succession drives spring community dynamics of small protists at Helgoland Roads, North Sea**. Journal of Plankton Research 42: 305–319. https://doi.org/10.1093/plankt/fbaa017.

Contribution: I conceived this topic and wrote this paper with input from all authors. Samples were provided by Bernhard Fuchs (MPI) and the MPI sampling team. DNA extraction was done in parts by Karl Peter Rücknagel and Sabine Kühn (both MPI). I conducted all other laboratory work (DNA extraction, library preparation for sequencing), with support from Pim Sprong, Swantje Rogge, and Kerstin Oetjen (both AWI). Stefan Neuhaus implemented quality control and taxonomic annotation of raw sequences. I conducted all further data analysis, visualization of the data, and wrote the original draft of the manuscript. All authors provided input and discussion resulting in the published version.

Data Availability Statement: All data to this publication is publicly available on the publisher's website. Sequence data have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession number PRJEB37135 (https://www.ebi.ac.uk/ena/data/view/PRJEB37135), using the data brokerage service of the German Federation for Biological Data (GFBio, Diepenbroek *et al*., 2014), in compliance with the Minimal Information about any (X) Sequence (MIxS) standard (Yilmaz *et al*., 2011).

Publication II:

Laura Käse, Katja Metfies, Stefan Neuhaus, Maarten Boersma, Karen Helen Wiltshire, Alexandra Claudia Kraberg. (2021) **Host-parasitoid associations in marine planktonic time series: Can metabarcoding help reveal them?**. PLoS ONE 16(1): e0244817. https://doi.org/10.1371/journal.pone.0244817.

Contribution: I conceived this topic and wrote this paper with input from all authors. Most samples were taken with the help of the Crew of the R/V Aade (AWI Helgoland). Aliquoting and filtering of the samples were done by the AWI Helgoland Long-Term Ecological Research (LTER) Team, namely Kristine Carstens, Silvia Peters, and myself. Additional samples were provided by Bernhard Fuchs (MPI) and the MPI sampling team as well as Pim Sprong (AWI). I carried out DNA extraction and all library preparation for sequencing, with the support of Swantje Rogge, Kerstin Oetjen, and Pim Sprong (all AWI). Stefan Neuhaus implemented quality control and taxonomic annotation of raw sequences. I conducted all further data analyses, visualized the data, and wrote the original draft of the manuscript. All authors provided expertise and gave input resulting in the published version.

Data Availability Statement: All data to this publication is publicly available. Sequence data have been deposited in the European Nucleotide Archive under accession number PRJEB37135 (https://www.ebi.ac.uk/ena/data/view/PRJEB37135) (see also **Publication I**).

All supplemental data are available on the publisher's website. Additionally, the metabarcoding dataset was published on PANGAEA:

Käse, Laura; Kraberg, Alexandra C; Metfies, Katja; Neuhaus, Stefan; Sprong, Pim; Fuchs, Bernhard M; Boersma, Maarten; Wiltshire, Karen Helen (2020): Eukaryotic microbial community at the LTER site Helgoland Roads from March 2016 to March 2019. PANGAEA, https://doi.org/10.1594/PANGAEA.921026.

Publication III:

Laura Käse, Katja Metfies, Alexandra C Kraberg, Stefan Neuhaus, Cédric L Meunier, Karen H Wiltshire, Maarten Boersma (2021) **Metabarcoding analysis suggests that flexible food web interactions in the eukaryotic plankton community are more common than specific predator-prey relationships at Helgoland Roads, North Sea**. ICES Journal of Marine Science, 78 (9): 3372–3386. https://doi.org/10.1093/icesjms/fsab058.

Contribution: I conceived this topic and wrote this paper with input from all authors. This publication includes the same samples and metabarcoding datasets as **Publication II and IV**. In addition to the contributions to the sampling and laboratory work as listed in the statement of **Publication II**, I conducted all data analyses needed for this publication, visualized the data, and wrote the original draft of the manuscript.

Data Availability Statement: All data to this publication is publicly available (see also **Publication II**).

Publication IV:

Laura Käse, Alexandra C. Kraberg, Stefan Neuhaus, Maarten Boersma, Karen H. Wiltshire, Katja Metfies (in prep.) **Conventional microscopy and metabarcoding of marine plankton in long-term ecological research– Are the results comparable?**. to be submitted.

Contribution: I conceived this topic and wrote this paper with input from some co-authors (Alexandra C. Kraberg and Katja Metfies). This publication includes the same samples and metabarcoding datasets as **Publication II and III**. I conducted the additional microscopic counts. Contributions to the sampling and laboratory work were listed in the statement of **Publication II**. I carried out all data analyses regarding this manuscript and the visualization of the data. I wrote the original draft of the manuscript.

Data Availability Statement: All metadata to this publication is publicly available (see also **Publication II**).

Publication V:

P.A.A. Sprong, V. Fofonova, K.H. Wiltshire, S. Neuhaus, K.U. Ludwichowski, L. Käse, A. Androsov, K. Metfies (2020) **Spatial dynamics of eukaryotic microbial communities in the German Bight**. Journal of Sea Research; 163: 101914. https://doi.org/10.1016/j.seares.2020.101914.

Contribution: A large part of the laboratory work on the Helgoland samples from spring 2016 was conducted by me (see also **Publication I** for information). The manuscript was written by Pim Sprong with input from all authors.

Publication VI:

Käse L. and Geuer J.K. (2018) **Phytoplankton Responses to Marine Climate Change – An Introduction.** In: Jungblut S., Liebich V., Bode M. (eds) YOUMARES 8 – Oceans Across Boundaries: Learning from each other. Springer, Cham. https://doi.org/10.1007/978-3-319-93284-2_5.

Contribution: This literature review was part of the conference proceedings of the YOUMARES 8 conference, which took place in Kiel, Germany, in September 2017, it was written by both authors equally.

III List of abbreviations

AIC	Akaike Information Criterion
ANOSIM	Analysis of Similarities
ANOVA	Analysis of Variance
ASV	Amplicon Sequence Variant
BLAST	Basic Local Alignment Search Tool
CalCOFI	California Cooperative Fisheries Investigations
CCA	Canonical Correspondence Analysis
clr	Centred Log-ratio
COI	Cytochrome Oxidase C Subunit I
DNA	Deoxyribonucleic Acid
eDNA	Environmental Deoxyribonucleic Acid
EMBL-EBI	European Molecular Biology Laboratory-European Bioinformatics Institute
ENA	European Nucleotide Archive
GFBio	German Federation for Biological Data
HPLC	High Performance Liquid Chromatography
ITS	Internal Transcribed Spacer
LTER	Long-Term Ecological Research
MAST	Marine Stramenopiles
MEGA X	Molecular Evolutionary Genetics Analysis X
MIxS	Minimum Information about any (x) Sequence
NGS	Next-Generation Sequencing
NMDS	Non-Metrical Multidimensional Scaling
OTU	Operational Taxonomic Unit
PC	Polycarbonate

III List of abbreviations

PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
PERMANOVA	Permutational Multivariate Analysis of Variance
PR2	Protist Ribosomal Reference Database
PVDF	Polyvinylidene Fluoride
rDNA	Ribosomal Deoxyribonucleic Acid
RNA	Ribonucleic Acid
SEM	Scanning Electron Microscopy
SIMPER	Similarity Percentage Analysis
SPIEC-EASI	SParse InversE Covariance estimation for Ecological Association and Statistical Inference
TEM	Transmission Electron Microscopy

IV Abstract

The marine pelagic environment is inhabited by various species of plankton. They can serve as a food source for the whole marine food web, and the autotrophic fraction is responsible for nearly half of global primary production. Due to the complexity of the task, methodological constraints, and a constantly changing marine environment, it has remained challenging to monitor plankton diversity and identify food web connections.

Metabarcoding techniques in marine ecological process studies as well as in long-term monitoring might help bridge these knowledge gaps. Previous research based on metabarcoding of plankton has provided unprecedented insights into biodiversity and the distribution of thus-far understudied groups such as eukaryotic picoplankton. Therefore, the main scope of this thesis was to gain information on understudied compartments of the eukaryotic plankton community at Helgoland Roads, a long-term monitoring site in the German North Sea. Using 18S metabarcoding of water samples, a three-year dataset was obtained which formed the basis for the publications in this thesis. The data and the publications based on these provide in-depth information on the planktonic structure of the pelagic system at Helgoland Roads and refine and complement our prior understanding. For example, remarkably high abundances of non-autotrophic taxa were found during spring, and spring bloom communities were not dominated by diatoms. Generally, the plankton community was highly diverse, with a large proportion of pico- and nanoplanktonic taxa. These were much more diverse and abundant than previously seen at Helgoland. They also displayed various connections in the planktonic food web. Unexpected predator-prey relationships were found in the food web, and zooplankton might be more opportunistic than previously known. A relevant finding was the high diversity and high abundance of the so-far understudied marine parasitoids at all taxonomic levels. Furthermore, I investigated the comparability of metabarcoding data to morphological counts. When implementing new methods into existing long-term monitoring, a certain level of complementarity in datasets is necessary, because correlations between new metabarcoding datasets and conventional datasets such as morphological counts might be weak depending on taxonomic levels and individual taxa, realizing the synergistic potential of these methods in combination remains a challenge. However, the promising results of metabarcoding should suffice to establish a clear role for metabarcoding in long-term monitoring. Combining different methods is the only way to obtain a more comprehensive view of the marine planktonic food web. Implementing new information about parasitoids or small-sized eukaryotic microbes in ecological models may make predicting changes in marine ecosystems easier in the future.

V Zusammenfassung

Marines Plankton dient als Nahrungsquelle für das gesamte marine Nahrungsnetz und der autotrophe Teil des Planktons ist außerdem für fast die Hälfte der globalen Primärproduktion verantwortlich. Aufgrund seiner Komplexität bleibt die Überwachung der Planktondiversität und die Identifizierung von Nahrungsnetzverbindungen eine Herausforderung, nicht nur aufgrund methodischer Einschränkungen, sondern auch aufgrund der ständigen Veränderungen in den marinen Lebensräumen.

Durch die Implementierung von Metabarcoding-Techniken in das Langzeitmonitoring von Plankton könnten Planktonkomponenten, die bisher nicht beobachtbar waren, sichtbar gemacht werden, insbesondere da frühere Forschungen, die auf Metabarcoding von Plankton basieren, vielversprechende Ergebnisse gezeigt haben. Daher war das Hauptziel dieser Arbeit, Informationen über verschiedene und meist wenig beachtete Komponenten der eukaryotischen Planktongemeinschaft an der Helgoland Reede in der deutschen Nordsee zu erhalten. Dazu wurden Wasserproben über einen Zeitraum von drei Jahren genommen und mittels 18S-Metabarcoding analysiert. Die Veröffentlichungen dieser Arbeit liefern Informationen über die planktonische Struktur in Helgoland, die den bisherigen Wissensstand ergänzen und verfeinern. Zum Beispiel wurden im Frühjahr bemerkenswert hohe Abundanzen nicht-autotropher Taxa gefunden, und Frühjahrsblütengemeinschaften wurden nicht von Kieselalgen dominiert. Im Allgemeinen war die Planktongemeinschaft vielfältiger und eng miteinander verknüpft mit einem großen Anteil an piko- und nanoplanktonischen Taxa. Es wurden flexible Räuber-Beute-Beziehungen im Nahrungsnetz gefunden und Zooplankton scheint opportunistischer zu sein, als zuvor angenommen. Ein weiteres außergewöhnliches Ergebnis war die weite Verbreitung, hohe Diversität und große Abundanz von bisher wenig erforschten Parasiten an der Station. Weiterhin wurde die Vergleichbarkeit von Metabarcoding-Daten mit konventionellen morphologischen Zählungen untersucht. Da die Korrelationen zwischen neuen Metabarcoding-Datensätzen und konventionellen Datensätzen je nach taxonomischer Ebene und einzelnen Taxa nur schwach ausgeprägt sein können, bleibt das Kombinieren von Datensätzen eine schwierige Aufgabe. Die vielversprechenden Ergebnisse, die durch Metabarcoding erzielt werden können, sollten jedoch Grund genug sein, Metabarcoding im Langzeitmonitoring zu etablieren. Die Kombination verschiedener Methoden ist die einzige Möglichkeit, einen umfassenderen Überblick über das marine planktonische Nahrungsnetz zu erhalten. Durch die Implementierung neuer Informationen über parasitoide Plankton oder kleine eukaryotische Mikroben in ökologische Modelle könnte die Vorhersage von Veränderungen im marinen Plankton in Zukunft einfacher werden.

1 Introduction

Plankton, those organisms which float in the water without being able to swim against the current (Hensen, 1887), are of particular importance in marine ecosystems as they comprise phytoplankton, important primary producers that fuel marine ecosystems. While the biomass of phytoplankton accounts for only a fraction of the biomass of terrestrial phototrophs, their contribution is nearly half of global photosynthetic net primary production or global carbon fixation (Field *et al.*, 1998; Simon *et al.*, 2009). Phytoplankton live in the euphotic zone of the ocean, where they take up inorganic carbon and convert it to organic carbon which is then available to higher trophic levels (Ducklow *et al.*, 2001). Not only do plankton serve as a food source for marine organisms throughout the food web, but they are also an important influencer of climate. For example, during photosynthesis phytoplankton take up CO_2, part of which gets sequestered, and stored in the deep ocean. Moreover, certain plankton groups can also produce climate-influencing substances such as dimethylsulfoniopropionate (Yoch, 2002; Stefels *et al.*, 2007). Thus, studies of plankton biodiversity, succession, and species interactions are essential for assessing the impact of global change on key marine ecosystem processes and functions.

The fact of climate change and its influence on the plankton community underpins the need for holistic examinations of plankton dynamics. However, despite continuous monitoring efforts especially over the course of long-term time series, capturing all the components of plankton communities remains challenging, despite Victor Hensen's suggestion that plankton in the ocean is generally well mixed and can be investigated on the basis of only a few samples (Hensen, 1887, 1890, 1895). While we now know that there is enormous horizontal, vertical, and temporal variation even on the smallest of scales, we still lack complete observational power to investigate every player in the planktonic community. Due to the constraints of conventional methods, such as the visually limited resolution of light microscopy, small-sized fractions of the plankton still have not been identified or enumerated properly; and as a result, their importance within the system has been greatly underestimated. Among these small fractions are small-sized microbial eukaryotes and planktonic parasites that are very likely critical to the functioning of marine ecosystems (Fogg, 1995; Lafferty *et al.*, 2006, 2008; Figueiras *et al.*, 2020; Juranek *et al.*, 2020). Concomitantly, information about connections in the food web that are based on these organisms is also missing. From microscopy and net sampling to molecular methods, new approaches and technologies have been developed with the potential to improve our observations of these organisms and the planktonic food web. The usefulness of these approaches and technologies requires further study and evaluation. Some plankton studies

have already used metabarcoding techniques, mostly in short-term or experimental contexts. Consequently, the next logical step is to consider using metabarcoding in long-term time series as an additional source of information on plankton diversity. This thesis aims to use metabarcoding to advance our understanding of the currently underrepresented eukaryotic plankton compartments and their position in the food web at the long-term monitoring site Helgoland Roads in the German North Sea. The following paragraphs introduce what is known about the marine food web and describe its current monitoring at Helgoland Roads to emphasize the plankton compartments that remain underrepresented.

The marine planktonic food web

The marine food web is a highly complex system with cross-linkages between all trophic levels (Figure 1a). Unravelling all these connections will increase the complexity of our representation of the food web. For example, depending on life stages, planktonic taxa that were previously seen as predators may also serve as prey (Roura *et al.*, 2018; Yeh *et al.*, 2020), which is especially important for small-sized unicellular organisms. Furthermore, mixotrophy is widespread throughout taxa (Flynn *et al.*, 2019) and seems to be the rule rather than the exception in eukaryotic microbes (Sanders, 1991; Mitra *et al.*, 2014; Caron, 2016). While some mixotrophic species are photosynthetic and capable of phagocytosis (engulfment of external particles), others are heterotrophic but take up phototrophic elements (plastids or symbionts) to keep in their cells (Adl *et al.*, 2019).

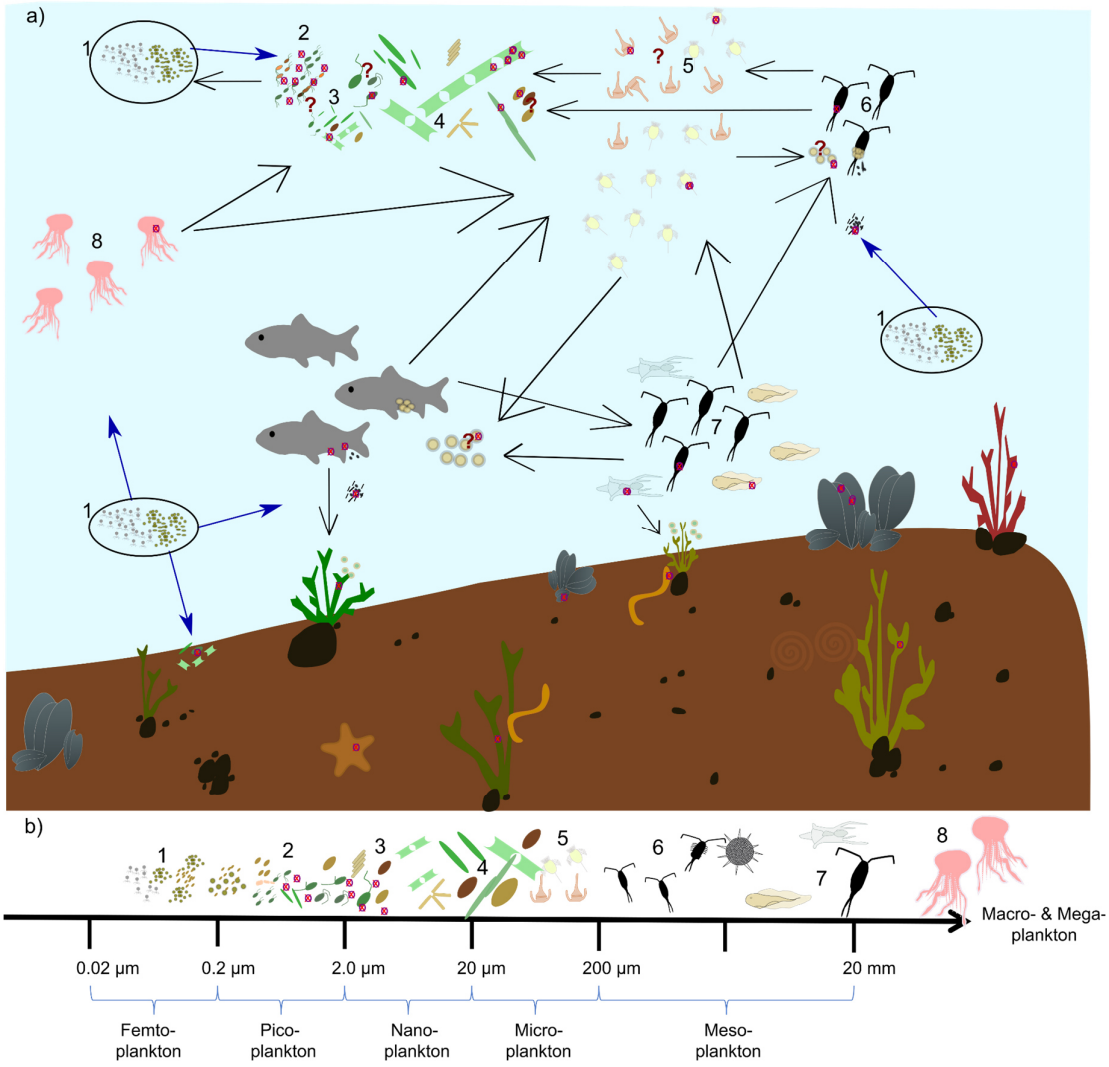


Figure 1: a) The planktonic food web in the marine environment. Feeding is indicated by black arrows, degradation by blue arrows, planktonic parasites/parasitoids indicated by violet dots with red x's; question marks indicate underrepresented components and interactions (more information in text); black circles indicate plankton components (viruses and prokaryotes) that were not surveyed in this thesis. **b) An overview of different plankton size classes using example organisms of different plankton taxa.** Size classes defined by Sieburth *et al.* (1978).

Virio- and bacterioplankton – not investigated in this thesis – can be found in the femto- and picoplankton fractions (Figure 1b). Bacteria and viruses, together with mycoplankton (fungi), are responsible for the remineralization of detritus, dead material, and dissolved organic carbon (Figure 1a**,** blue arrows, components 1 & 2). Together with other plankton components, they are part of the so-called "microbial loop" (Azam *et al.*, 1983) now accepted as an essential component of the microbial food web (Sherr and Sherr, 1988).

Picoplankton (Figure 1a, b, component 2) comprise highly diverse single-celled pro- and eukaryotes of all different trophic modes (autotrophic, heterotrophic, and mixotrophic, but

also mutualistic or parasitic), that play a key role in the remineralization of nutrients (Zhu *et al.*, 2005; Marie *et al.*, 2006; Vaulot *et al.*, 2008; Massana, 2011). Especially because of their small size and often indistinct morphological characteristics, they can be difficult to identify and as a part of the marine food web have been ignored for a long time, despite high abundance globally. They can respond and adapt to different environmental conditions much faster than larger plankton (Raven, 1998; Not *et al.*, 2004; Šlapeta *et al.*, 2006; Simon *et al.*, 2009). Eukaryotic picoplanktonic heterotrophs include for example the taxonomic group marine Stramenopiles (MAST), which consist of phagotrophic flagellates. Other components of the picoplankton are phytoplankton and therefore contribute to photosynthetic primary production while also serving as important food for diverse grazers. Picoplankton can dominate the food web seasonally or only occasionally in coastal regions; in the open ocean or mostly nutrient-depleted regions they can be the main component of the phytoplankton (Ray *et al.*, 1989; Jacquet *et al.*, 2002; Medlin *et al.*, 2006; Grob *et al.*, 2007; Not *et al.*, 2008; Knefelkamp, 2009; Metfies *et al.*, 2010; Purcell-Meyerink *et al.*, 2017; Otero-Ferrer *et al.*, 2018). While heterotrophic picoplankton feed on organic materials or plankton cells and remineralize nutrients, autotrophic picoplankton take up these nutrients (Legendre and Le Fèvre, 1995) and therefore compete with bigger phytoplankton.

Mixo- and heterotrophic nanoflagellates (Figure 1 a, b, component 3) are known to feed on bacteria and therefore play an important role in the carbon cycle (Pomeroy, 1974; Azam *et al.*, 1983; Sanders *et al.*, 2000; Yang *et al.*, 2018; Edwards, 2019). In addition to these heterotrophic nanoflagellates, nanoplankton also contain various parasitic species. Even though certain nanoplankton are known to be important within the food web and food web functions of distinct groups can vary, nanoplankton are mostly treated as one functional unit; the underlying functions of their constituents in the food web remain largely unexplored – an underrepresented fraction, especially in long-term studies (Stern *et al.*, 2018). This is mainly because of their small size, but their similar morphology is also a factor. Particularly challenging has been the identification of parasitic species (Skovgaard, 2014), which are important because they can drive phytoplankton bloom dynamics and affect phytoplankton succession due to their influence on the host population (Tillmann *et al.*, 1999; Rasconi *et al.*, 2011; Skovgaard, 2014).

Where pico- and nanoplanktonic phytoplankton dominate, for example under limitation of certain nutrients such as iron or silicon, phytoplankton blooms are less likely because grazing pressure is too high and phytoplankton population densities are typically fairly stable (Strom *et al.*, 2000; Sommer *et al.*, 2002; Calbet and Landry, 2004). By contrast, in other regions strictly autotrophic diatoms are one of the main groups of phytoplankton (Figure 1a, b, components 4 & 5). Under favourable environmental conditions in certain areas, they

form blooms and are a major food source for herbivorous zooplankton. Additionally, the grazing pressure exerted by herbivorous zooplankton can affect bloom formation and duration (Montagnes and Lessard, 1999; Sherr and Sherr, 2002; Aberle *et al.*, 2007).

Micro- and mesozooplankton (Figure 1a, b, components 5-7) may prefer to feed on particular phytoplankton, so they can have a major effect on phytoplankton composition (Riegman *et al.*, 1993; Burkill *et al.*, 1995; Edwards and Burkill, 1995; Montagnes and Lessard, 1999; Calbet and Landry, 2004; Fonda Umani *et al.*, 2005; Aberle *et al.*, 2007; Löder *et al.*, 2011). Additionally, mesozooplankton have various life stages – egg or larvae, for example (Figure 1a, b, component 7) – all with potentially differentiated feeding preferences and all serving as a food source for other planktonic predators. For example, microzooplankton include mixo- or heterotrophic dinoflagellates, heterotrophic ciliates, and other heterotrophic flagellates ranging in size from 20 to 200 µm (Figure 1a, b, component 5). Microzooplankton are the most important consumers of phytoplankton (Landry and Calbet, 2004; Sherr and Sherr, 2007); they may graze as much as 60 to 75% of daily phytoplankton production depending on the region (Calbet and Landry, 2004; Landry and Calbet, 2004; Calbet *et al.*, 2008). Unicellular microzooplankton are able to respond quickly to changes in phytoplankton (Aberle *et al.*, 2007; Löder *et al.*, 2012). Microzooplankton can feed on diatoms and a range of other plankton; dinoflagellates are even known to prey on plankton bigger than their own size (Hansen *et al.*, 1994; Hansen and Calado, 1999; Calbet, 2008; Yang *et al.*, 2020). Heterotrophic nanoflagellates are a potential food source for microzooplankton, as are copepod eggs and nauplii (Jeong, 1994). Most microzooplanktonic dinoflagellates are considered to be mixotrophs and therefore also contribute to the remineralization of nutrients in the water column (Sherr and Sherr, 2002).

Microzooplankton are an important food source for mesozooplankton, especially when phytoplankton concentrations are low (Gifford, 1991; Nejstgaard *et al.*, 2001; Calbet and Saiz, 2005; Löder *et al.*, 2011; Yang, 2014). Copepods, the most abundant type of mesozooplankton (Figure 1a, b, components 6 & 7), are able to feed selectively (Meunier *et al.*, 2016). When phytoplankton quality is low, copepods switch their diet to heterotrophic microzooplankton like ciliates or dinoflagellates (Jones and Flynn, 2005; Gentsch *et al.*, 2009; Saage *et al.*, 2009). They can also prey on fish eggs or larvae (Turner *et al.*, 1985).

The studies behind this thesis were conducted at Helgoland Roads, a location for which bigger-sized plankton components are recorded thoroughly and in great detail. However, thorough identification of nanoplankton has been mostly impossible. The plankton counts thus include mainly phytoplankton and zooplankton such as diatoms, dinoflagellates, and some ciliate species as well as copepods (Figure 1a, components 4-6), (Greve *et al.*, 2004;

Wiltshire *et al.*, 2010). Of these, diatom counts are the most reliable (Wiltshire and Dürselen, 2004). Among other things, studies using these data regularly examine long-term changes in some diatoms and the influence of environmental parameters (Gebühr *et al.*, 2009; Freund *et al.*, 2012; Schlüter *et al.*, 2012), or they estimate carrying capacity (Sarker and Wiltshire, 2017), coexisting species (Sarker *et al.*, 2018), and changes in their phenology (Scharfe and Wiltshire, 2019). Other studies have concentrated on the environmental relevance of particular taxon groups that are not regularly observed in detail, including marine prokaryotes (which this thesis omits) and eukaryotes. Some studies have focused on bacterial communities (Figure 1a, component 2) during certain seasons (Sapp *et al.*, 2007; Lucas *et al.*, 2016; Teeling *et al.*, 2016) or year round (Lucas *et al.*, 2015; Chafee *et al.*, 2018). Assessments have also included analysis of eukaryotic picoplankton (Medlin *et al.*, 2006; Knefelkamp, 2009) and specific picoplanktonic groups such as Prasinophytes (Gescher *et al.*, 2008) (Figure 1a, component 2); Cryptophyta, which are mostly found in the nanoplankton (Figure 1a, component 3) (Metfies *et al.*, 2010; Medlin *et al.*, 2017); mycoplankton (Banos *et al.*, 2020); and micro- and mesozooplankton (Figure 1a, components 5 & 6) (Löder *et al.*, 2012; Yang *et al.*, 2015, 2021). Even though several parasitoids – parasites that kill their hosts (Skovgaard, 2014) (Figure 1a, component 3) – have been found or described at Helgoland (Drebes, 1966; Thines *et al.*, 2015; Buaya *et al.*, 2017; Metfies *et al.*, 2020), their connections within and impact on the food web are not known. To have included thorough monitoring in the time series would have been impossible due to time and methodological constraints in the past. Many of these other studies have utilized molecular methods, hinting at their importance for detailed biodiversity studies; however, these studies have focussed mostly on single groups or species and used methods that specifically targeted these taxon groups. Metabarcoding might in contrast be applicable because of the overview it can provide of several size classes and taxa at once; this allows better coverage and a better understanding of these underrepresented groups.

Thus far, I have described a fairly static picture of the planktonic food web above; but there is of course extreme seasonality in plankton, especially in temperate and polar environments. Therefore, the dynamics of the planktonic food web cannot be fully understood through once-off sampling; they need to be described in light of the temporal changes in the plankton and in the environment.

Seasonality of plankton in temperate environments

Numerous abiotic factors shape the marine environment, including for example light, currents, wind, radiation, precipitation, temperature, salinity, and nutrient composition. Due to high variation on top of the somewhat predictable seasonality of these abiotic factors,

marine plankton communities are forced to be flexible and must regularly adapt their growth to changing conditions (Smayda, 1998; Chivers *et al.*, 2017). Temperature and light have an especially big impact on plankton growth and behaviour (Andersson *et al.*, 1994; Wiltshire and Manly, 2004; Wiltshire *et al.*, 2008; Lewandowska and Sommer, 2010; Winder *et al.*, 2012). Differences in nutrient availability can influence or even limit plankton food webs and their trophic levels (Sommer *et al.*, 2002). The food web is influenced by chemical and physical factors as well as by species interactions, competition for resources, grazing pressure, viral or bacterial infections, and parasitic organisms (Hutchinson, 1961; Park *et al.*, 2004; Rhodes *et al.*, 2008; Wiltshire *et al.*, 2010; Arsenieff *et al.*, 2019). Two control mechanisms have been described in this context: top-down control, which describes influences on population dynamics based on the predators in the community; and bottom-up control, in which resource availability determines the population dynamics.

In temperate areas in general, phytoplankton or plankton diversity and biomass are fluctuating throughout the year. Instead, they are seasonal and sensitive to environmental conditions, but at the same time, the level of dynamism varies greatly depending on the regional properties. For example, conditions as well as communities differ greatly between the open ocean and shallow shelf or coastal sea regions. Looking at the North Sea alone, plankton communities differ greatly by specific region and season (Reid *et al.*, 1990; Tillmann and Rick, 2003; Leterme *et al.*, 2006; Wollschläger *et al.*, 2015). Changing environmental conditions paired with interactions in the community create a seasonal succession of plankton species (Scharfe and Wiltshire, 2019). In winter, when temperature and light availability are generally low, bacteria remineralize nutrients, making them available for phytoplankton that are growth-limited due to lack of light (Wiltshire *et al.*, 2008; Sarker, 2018). There is additional nutrient input from rivers, and water masses become well mixed due to strong currents (Brockmann *et al.*, 1990; Callies and Scharfe, 2015). In early spring, when enough light is available and nutrient supplies are high, phytoplankton can grow rapidly. At Helgoland, diatoms are the major phytoplankton component during that time (Wiltshire *et al.*, 2015), and they show distinct, massive blooming patterns (Mieruch *et al.*, 2010). However, these kinds of blooms are not limited to diatoms but also include nano- and dinoflagellates (Hickel, 1998; Knefelkamp, 2009; Löder *et al.*, 2011). At the beginning of the growing season these spring blooms are characterized by a mixture of species and the timing of succession patterns depends on the respective species (Scharfe and Wiltshire, 2019). Silicate, which diatoms require to build their cell walls, is taken up first (Wiltshire *et al.*, 2015). At the beginning of spring, potential grazers are not very abundant coming off of low food concentrations; but by late spring, top-down control heavily affects the amplitude of the bloom (Sarker, 2018). Since predators may need time to respond, zooplankton

usually increase in abundance after peak phytoplankton growth (Wiltshire *et al.*, 2015). Dinoflagellates, other heterotrophic plankton, and copepods can grow in greater abundances while diatoms are still silicate-limited (Wiltshire *et al.*, 2015; Sarker, 2018) and under greater grazing pressure than during early spring. In late autumn, with declining light and nutrients combined with heavy grazing pressure, phytoplankton are no longer able to grow in great number, and the plankton community shifts again, with pico- and nanoplankton making up a greater proportion of the plankton community in winter (Riegman *et al.*, 1993; Knefelkamp, 2009; Metfies *et al.*, 2010). This calls for further study of the influence of the plankton compartments, not only pico- and nanoplankton but also parasitoids, which have been underrepresented so far in the literature. It is at the same time important to include every compartment of the plankton in a seasonally differentiated sampling design, in order for an assessment to take place of whether bottom-up or top-down factors are the predominate drivers of the dynamics in a seasonally changing abiotic environment.

2 Objectives

Based on current knowledge and the state of the art in plankton sampling, there is clearly great need for a more integrated sampling to take every player in the system into account. This thesis therefore investigates eukaryotic plankton diversity for several years by sequencing the variable region V4 of the 18S ribosomal DNA (rDNA) from water samples taken at the Helgoland Roads sampling point. In particular, I investigated the influence of small plankton and parasitoids on the plankton food web as these cannot be identified by light microscopy, so conventional time series have largely ignored them. Overall, this thesis addresses the following research objectives:

1. What is the specific added value of using metabarcoding to study the ecology of the eukaryotic plankton community?
    a. How diverse is the (small-sized) eukaryotic plankton community (chapters 4 to 7)?
    b. How are plankton communities structured throughout the year and can they be linked to environmental changes (chapters 4 to 7)?
    c. Does metabarcoding capture hitherto largely ignored compartments of the planktonic community such as parasitoids, and if so, how do these components influence the community structure throughout the year? How can the parasitoids be linked to potential hosts (chapter 5)?
    d. How are the different plankton components linked throughout the food web (chapter 6)?
2. Is the information obtained through metabarcoding comparable to that of conventional counting methods?
    a. Do the respective methods show matching results (chapters 4 and 7)?
    b. How do individual methods differ and what are the advantages of each (chapter 7)?

Based on findings in response to the above questions, I then aimed to address the following questions: Is it both feasible and scientifically reasonable to integrate metabarcoding into an existing long-term time series, and how can results of different methods be combined for the best added value in long-term time series?

Chapter 3 of this thesis describes the study area and the long-term monitoring program Helgoland Roads; light microscopy as conducted in the time series and in parts of this thesis; and metabarcoding as the main method used in this thesis. Chapters 4 to 7 include reprints of **Publication I to IV**, which are then followed by discussion (chapter 8) and a conclusion

and outlook (chapter 9). **Publication I** identifies eukaryotic microbes including the spring pico- and nanoplankton fraction, as spring is known as an important season in temperate waters. **Publication II** identifies the extent of parasitoid plankton diversity in the environment. **Publications III and IV** focus on identifying all the other plankton size fractions. **Publications I and IV** compare the diversity identified through metabarcoding with the results of conventional counts in the long-term series. **Publications I to IV** investigate potential links between environmental conditions and plankton communities, both on the whole as well as with respect to certain planktonic fractions. **Publication II** is aimed at using the new knowledge about the occurrence and diversity of parasitoids to detect host–parasitoid associations. To this end, it examines known host–parasitoid systems and investigates potential links to the metabarcoding data. **Publication III** uses a similar approach to identify potential predator-prey pairings and conducts a network analysis to achieve an overview of associations throughout the food web. Besides comparing the diversity shown by metabarcoding with that of conventional light microscopy, **Publication IV** discusses in detail the advantages and disadvantages of conventional methods versus metabarcoding and gives initial insights into the potential of using metabarcoding data for long-term research. This thesis provides information on the extent to which metabarcoding can offer insights into plankton communities and succession which are valuable, distinct, and complementary to conventionally obtained data. Finally, I recommend ways in which metabarcoding may be integrated into existing time series.

3 Plankton monitoring

Countless approaches have been developed in order to study and monitor plankton. Each has supported the pursuit of various research questions, addressing whole communities, specific groups, or a single species of plankton. Investigating planktonic and other marine biota still requires time, expert personnel, and microscopic identification of plankton, and light microscopy still remains the most common and important identification method. However, these methods also have limitations (Zingone *et al.*, 2015; Stern *et al.*, 2018), and most studies do not focus on complex community dynamics. So new approaches still need to be tested, proved, and potentially incorporated into existing monitoring programmes. New automatic and remote systems such as satellite sensors, data loggers, ferry boxes, floats, gliders, and moorings have been implemented (Gower *et al.*, 2008; Henson, 2014; Petersen, 2014; Clayton *et al.*, 2022). While some time series have already implemented pigment analysis via sensors or high performance liquid chromatography (HPLC) (Letelier *et al.*, 1993; Durand *et al.*, 2001; Karl *et al.*, 2001), these can provide only limited information on plankton (Irigoien *et al.*, 2004). Evolving methods like metabarcoding show enormous potential for implementation into time series (Stern *et al.*, 2018) because they can provide very detailed taxon information.

Time series – regular observations over a long period of time – are essential to understanding long-term ecosystem dynamics. Only through continuous monitoring is it possible to observe ecosystem functional and structural dynamics and assess any consequences on ecosystem services. Long-term experiments and ongoing time series help us understand the ecology of the entire earth system, disclosing information on ecological processes and helping to improve models that should predict future ocean dynamics (Ducklow *et al.*, 2009; Henson, 2014), such as changes in food webs or shifts in environmental processes. Long-term observations are generally needed to distinguish long-term environmental changes from the noise of typical fluctuations or even of rare events in an ecosystem that short-term studies cannot identify (Parr *et al.*, 2002; Sukhotin and Berger, 2013; Henson, 2014). Such fluctuations can depend on seasonal cycles, and they can be very specific on a local or global scale; for example, ecological shifts related to climate change or regime changes in a specific region can only be detected if at least a few decades' worth of data is available (Rebstock, 2002; Walther *et al.*, 2002).

It has been acknowledged that time series are important for understanding ecology, especially in a time of global change (Pugnetti *et al.*, 2013; Hughes *et al.*, 2017; Kuebbing *et al.*, 2018); but at the same time, several limitations have forced some time series to be discontinued or measurements to be reduced to save time and money (Wolfe *et al.*, 1987;

Duarte *et al.*, 1992; Hughes *et al.*, 2017). These limitations are related to funding and to the time-consuming nature of collecting a significant volume of data while producing less research than short-term studies can (Sukhotin and Berger, 2013; Henson, 2014; Kuebbing *et al.*, 2018). Sampling consistency is one of the most important factors in long-term monitoring (see Wiltshire and Dürselen, 2004). The sampling frequency, which can range from yearly to daily to hourly measurements, determines both how much information on the biological processes in the plankton communities is collected as well as how the resulting datasets can be interpreted. Existing time series need to be maintained and expanded to provide information on plankton communities in times of global change.

One common monitoring effort has been to apply marine (planktonic) time series. Immense efforts have been undertaken all over the world to run time series in the marine environment (Hays *et al.*, 2005), from open ocean time series like the Hawaii Ocean Time-series (Landry *et al.*, 2001; Karl and Church, 2014) and the Bermuda Atlantic Time-series Study (Letelier *et al.*, 1993; Durand *et al.*, 2001); the Continuous Plankton Recorder survey in the North Atlantic (Reid *et al.*, 2003; Richardson *et al.*, 2004; McQuatters-Gollop *et al.*, 2015); the HAUSGARTEN observatory in the Fram Strait (Soltwedel *et al.*, 2005, 2016); to time series whose foci tend toward the coastal, like the California Cooperative Fisheries Investigations (CalCOFI) programme in the California Current along the Pacific coast (Rebstock, 2002; Bograd *et al.*, 2003); the Long-Term Ecological Research (LTER) MareChiara at Naples on the Mediterranean Sea (Zingone *et al.*, 2019); the Helsinki Commission surveys in the Baltic Sea (Wasmund *et al.*, 2011); the Helgoland Roads time series in the German Bight (North Sea) (Greve *et al.*, 2004; Wiltshire *et al.*, 2010); or Station L4 in the western English Channel (John *et al.*, 2001; Harris, 2010; Widdicombe *et al.*, 2010).

Study area of this study

The North Sea is a temperate and mostly shallow sea adjacent to the Atlantic Ocean in northern Europe (Sündermann and Pohlmann, 2011). It is bordered by the coastlines of several European countries, including Norway, Sweden, Denmark, Germany, the Netherlands, Belgium, and the United Kingdom Marine water masses are transported from the North Atlantic or to a lesser extent through the English Channel; and another connection exists to the adjacent brackish Baltic Sea via the Skagerrak and Kattegat. Freshwater input into the North Sea is high due to outflow from several rivers such as the Thames, Seine, Rhine, and Elbe as well as meltwater input from the Scandinavian coasts (Ducrotoy *et al.*, 2000).

The shelf sea system is used intensely by shipping, tourism, wind parks and other human interests (Halpern *et al.*, 2008; Emeis *et al.*, 2015). Helgoland consists of two islands located

about fifty kilometres from the coast in the German Bight, the south-eastern part of the North Sea. Hydrographic conditions here are highly dynamic and influenced by both coastal water and the open North Sea (Callies *et al.*, 2017). The water column is generally well mixed with depths fluctuating from six to ten metres depending on the tide and wind directions (Callies and Scharfe, 2015). As a result, water samples taken at the surface are representative of the whole water column (Hickel, 1998; Wiltshire *et al.*, 2015).

In 1962, a long-term monitoring site was established between Helgoland's two islands – a mainland and a small, sandy outcrop called Düne – at Helgoland Roads (54°11.3'N, 7°54.0'E). The site is maintained by the Biologische Anstalt Helgoland / Alfred-Wegener-Institut (Wiltshire and Dürselen, 2004; Wiltshire *et al.*, 2010). Because samples are taken five times a week, the Helgoland Roads time series is an unparalleled and very detailed source of information on phytoplankton composition (mainly diatoms and dinoflagellates), water temperature, salinity, nutrient concentrations, turbidity (Secchi depth), and other parameters (Wiltshire *et al.*, 2010, 2015). Secchi depth and water temperature are measured directly at the station. Sampling is conducted with a bucket, and the haul is sub-sampled into a bottle for all other analyses. An aliquot of the water sample is filtered and used for nutrient measurements (silicate, phosphate, ammonium, nitrate, and nitrite) based on the methods of Grasshoff (1976). Another subsample (100 mL) of the water is filled into brown glass bottles and preserved with neutral Lugol's iodine (0.1%) and later used for phytoplankton identification (Hoppenrath, 2004; Wiltshire and Dürselen, 2004; Wiltshire *et al.*, 2010; Kraberg *et al.*, 2019). Phytoplankton is identified by light microscopy based on the Utermöhl method (Lund *et al.*, 1958; Utermöhl, 1958). An additional zooplankton time series was established in 1974 (Greve *et al.*, 2004), in which samples are taken three times a week. The zooplankton sampling uses two different nets: a 150 µm Nansen net, which is used for oblique hauls, and a 500 µm CalCOFI net, which is towed behind the research vessel (Wiltshire *et al.*, 2010). The long-term dataset has been reviewed and undergone quality control (Wiltshire and Dürselen, 2004; Raabe and Wiltshire, 2009), and it is used widely to answer ecological questions and to show long-term changes in the environment and plankton communities at Helgoland Roads (Wiltshire and Manly, 2004; Wiltshire *et al.*, 2008, 2010, 2015; Boersma *et al.*, 2015).

Light microscopy

Conventional light microscopy is still the only means of checking species based on the optical taxonomic markers used in the literature. It is frequently applied to identify and count planktonic organisms (Soares *et al.*, 2011). In general, analysing plankton samples involves making either live or preserved (fixed) counts. Two major advantages of live cell counting

compared to preserved counts are that the cells keep their natural colour and their motility patterns remain observable. The inverse counting method developed by Utermöhl (Utermöhl, 1958) is preferred mainly for quantitative phytoplankton analysis. The LTER at Helgoland implements it, and so does this thesis (see Figure 2 for an overview of the workflow behind this thesis). The method is cheap (except for the labour needed to do the actual counting) and easy. However, reliable species identification requires the analyst to have detailed taxonomic knowledge. Only a small sample volume of up to 100 mL is analysed, which is why this method does not quantitatively cover rare taxa (Utermöhl, 1958). The sample is mixed with a fixative and stored in glass bottles if samples need to be stored at length. The next step is to pour a sample into a settling chamber, of which there are different sizes (up to 100 mL). After several hours of sedimentation – the precise interval depends on chamber size – the sample is counted using inverted microscopy. The long settling time is one of the main disadvantages, and methods have been developed to shorten the procedure (Paxinos and Mitchell, 2000). How long the analysis takes also depends on the concentration of plankton in the sample (Edler and Elbrächter, 2010). Common fixatives are formaldehyde- or glutaraldehyde-based solutions and Lugol's solutions. The latter are made up of aqueous potassium iodide plus iodine and available in alkaline, neutral, or acid varieties (Andersen and Throndsen, 1995). In Lugol's-fixed samples, cells are stained a brownish colour, which can be removed using sodium thiosulfate if necessary (Edler and Elbrächter, 2010). Compared with formaldehyde, Lugol's solution increases the settling velocity of cells (Andersen and Throndsen, 1995). Cell loss and the difficulty of identifying phytoplankton depend on the type of preservation and the time between preservation and analysis (Utermöhl, 1958; Stoecker *et al.*, 1994; Williams *et al.*, 2016). Chlorophyll autofluorescence as it is found in photo- and mixotrophic organisms cannot be measured to identify these organisms when stained with Lugol's (Lund *et al.*, 1958; Utermöhl, 1958; Cermeño *et al.*, 2014). In addition to impeding clear identification because of the change in colour, Lugol's can cause cell shrinkage and eventually dissolve various groups of phytoplankton (Choi and Stoecker, 1989; Stoecker *et al.*, 1994; Menden-Deuer *et al.*, 2001; Zarauz and Irigoien, 2008). Silicates, for example, are affected by a too-high pH whereas a low pH causes calcified cells such as coccolithophorids to dissolve (Hällfors *et al.*, 1979). It has been reported that the abundance and formation of aggregates of nanoplankton decrease when samples are fixed with Lugol's solution (Zarauz and Irigoien, 2008).

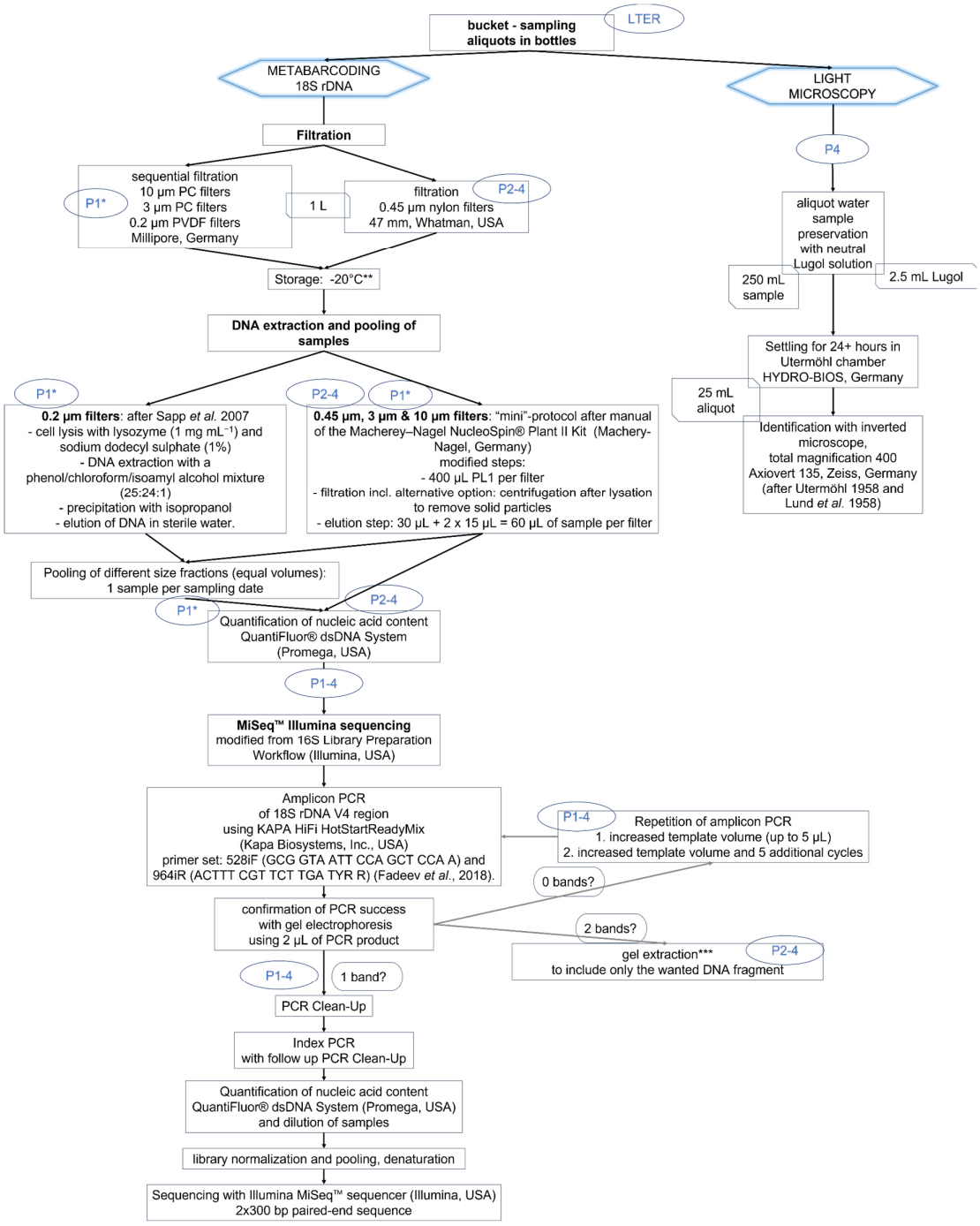


Figure 2: Workflow of the methods (metabarcoding, light microscopy) used in this thesis. P (in blue) = Publication; PC = Polycarbonate; PVDF = polyvinylidene fluoride; grey lines indicate steps not part of the "main workflow"; *Samples from this publication were also included in P2–4. **After filtration, samples were always stored at −20 °C. ***Less than ten samples.

Metabarcoding

Molecular methods such as in situ hybridization, molecular sensors, and sequencing methods like metabarcoding all have advantages compared with other plankton identification methods. They are relatively fast, and morphological expertise is not required (Ebenezer *et al.*, 2012). Most molecular methods rely on polymerase chain reaction (PCR) to amplify the genetic material. All biodiversity levels can be investigated with small sample volumes, if databases are reliable. For most applications, their sensitivity is high, and it is not necessary to cultivate or isolate species to obtain specific data. It is thus possible to analyse dead or non-culturable cells (Medlin and Kooistra, 2010).

Like any method, molecular approaches also have disadvantages. The biggest disadvantages are harming or destroying cells and biases and errors in the PCR (Suzuki and Giovannoni, 1996; Becker *et al.*, 2000; Acinas *et al.*, 2005) that creep in through PCR cyclers, enzymes, or primer choice as well as through PCR conditions such as DNA and primer concentrations, temperatures, and the number of cycles, all of which can complicate the comparability of workflows. Furthermore, the organisms have a different number of copies of genes of interest (LaJeunesse *et al.*, 2005; Connolly *et al.*, 2008; Hong *et al.*, 2016), leading to variation in amplification and impeding cell quantification.

Metabarcoding is the simultaneous identification of different taxa of a community via DNA or RNA gene sequences, that is, barcodes. As metabarcoding is being used more and more frequently in plankton research, the amount of information on biodiversity in marine microbial communities is increasing significantly. Various sequencing techniques have the potential to reveal the trophic modes not only of new groups such as parasitoids but also of heterotroph or phototroph unicellular microbes in the marine realm (Massana *et al.*, 2004, 2014, 2015; Guillou *et al.*, 2008; Logares *et al.*, 2012; de Vargas *et al.*, 2015). It is furthermore possible to detect rare species (Sogin *et al.*, 2006; Medinger *et al.*, 2010; Logares *et al.*, 2014) and to uncover interactions between different plankton components (Lima-Mendez *et al.*, 2015; Millette *et al.*, 2018).

While new developments in and approaches to metabarcoding never cease, they can vary greatly depending on the research objectives, and no standardized procedure exists. At each of the steps involved, ranging from study design and sampling, to preparing and sequencing the samples, to the bioinformatic and statistical analyses, there are numerous possibilities for optimization as well as numerous sources of errors or bias (Table 1) (Santoferrara, 2019). This great variability makes a comparison between different metabarcoding studies more complicated, especially if the metadata are not made available.

Table 1: Organismal (O) and methodological (M) drawbacks and biases by of metabarcoding, their impact on analysis and plankton identification, as well as a list of and potential solutions to minimize their effects.

Drawback/Bias Type	Explanation / impact	Potential solution
Life stages (O)	Cannot be distinguished	Additional morphological studies, identification of genes specific to certain life stages
Species recognition (O)	Not always possible due to high similarity of sequences in certain regions	Development of specific primers
Community diversity (O)	Depends on size; bigger taxa might be over-amplified	Use of mock samples and correction factors
Copy numbers (O)	Depend on size and life stage but differ between taxa	Use of mock samples and correction factors
Sampling (M)	Sampling technique and volume influence community composition; fragile taxa might be lost; for net samples, rare taxa might be under-sampled depending on cell size	Standardized sampling, replicates, and variation of sampling volumes depending on study design
Primer and marker choice (M)	Differing taxonomic precision, favoured/ignored taxa	Use of specific primers depending on focus, Use of different markers
PCR, extraction, amplification, library preparation, sequencing (M)	Extraction quality may differ depending on morphology; only certain taxa may be amplified and some favoured; sequencing errors and bias of sequencing platforms	Optimizing of PCR conditions, use of mock samples, replicates, and applicable thresholds in the bioinformatic pipeline
Bioinformatic procedure (M)	Thresholds influence dataset quality and limit detection of rare taxa	Varying thresholds and quality control of pipelines
Database (M)	Incomplete or wrong annotations, insufficient resolution for species-level discernment of taxa	Thorough curation of databases; use of different databases depending on study focus

In order to conduct metabarcoding studies on (environmental) plankton samples (see Figure 2 for an overview of the workflow used in this thesis), the water sample needs to be filtered to retain the organisms or environmental DNA (eDNA) from the sample. Alternatively, one can use whole organisms sampled by nets. If the scope of the research includes identifying different size fractions, the procedure also includes conducting fractionated filtration.

Conventionally, the next steps are DNA or RNA extraction followed by purification to keep the sample clean. The nucleic acids are then amplified by PCRs, with different marker genes

targeted depending on the focus of the metabarcoding study. Most primers are constructed to amplify the regions of the ribosomal DNA (eukaryotes: 18S or 28S rDNA, prokaryotes: 16S), the internal transcribed spacer (ITS) region, the mitochondrial cytochrome oxidase c subunit I (COI) region or the large subunit of ribulose bisphosphate carboxylase (rbcL) (Santoferrara, 2019). Use of regions V4 and V9 is common for the 18S rDNA regions; for ITS, primers are mostly constructed for ITS1 and ITS4 (Stern *et al.*, 2018; Santoferrara, 2019).

Samples are then set up into libraries for the sequencing procedure; the steps to be taken varying depending on the respective platform, usually systems based on 454 sequencing (Margulies *et al.*, 2005; Rothberg and Leamon, 2008) or Illumina (Ebenezer *et al.*, 2012; Mahé *et al.*, 2015a). Subsequently, the reads (raw sequencing data) are bioinformatically processed. Since these bioinformatic pipelines are individually developed and continuously optimized (see Figure 3 for an overview of an example of the pipeline used in this thesis), usage depends on a given study's objectives and design. Numerous pipelines are available (e.g. Brandt *et al.* (2021); QIIME 1 & 2 (Caporaso *et al.*, 2010; Bolyen *et al.*, 2019; Estaki *et al.*, 2020); PEMA (Zafeiropoulos *et al.*, 2020); MARES (Arranz *et al.*, 2020)); and there are analysis tools, software, and clustering strategies. These software, tools and strategies include the applications mothur (Schloss *et al.*, 2009) and DADA2 (Callahan *et al.*, 2016), for example, or the tools PEAR (Zhang *et al.*, 2014), Cutadapt (Martin, 2011), and VSEARCH (Rognes *et al.*, 2016),or the clustering algorithms CD-HIT (Fu *et al.*, 2012), UCLUST (Edgar, 2010), and Swarm & Swarm v2 (Mahé *et al.*, 2014, 2015b). During processing of the raw reads, sequences are commonly clustered into operational taxonomic units (OTUs) or amplicon sequence variants (ASVs) (Callahan *et al.*, 2017) whose respective sequences have to be aligned with reference sequences for taxa identification. The common reference databases do not cover in equal depth all that is known about sequences of (marine) biota, and so the database must be chosen carefully depending on wanted taxa. Prominent databases include SILVA (Pruesse *et al.*, 2007), the Protist Ribosomal Reference database (PR2) (Guillou *et al.*, 2013), and GenBank (Benson *et al.*, 2013).

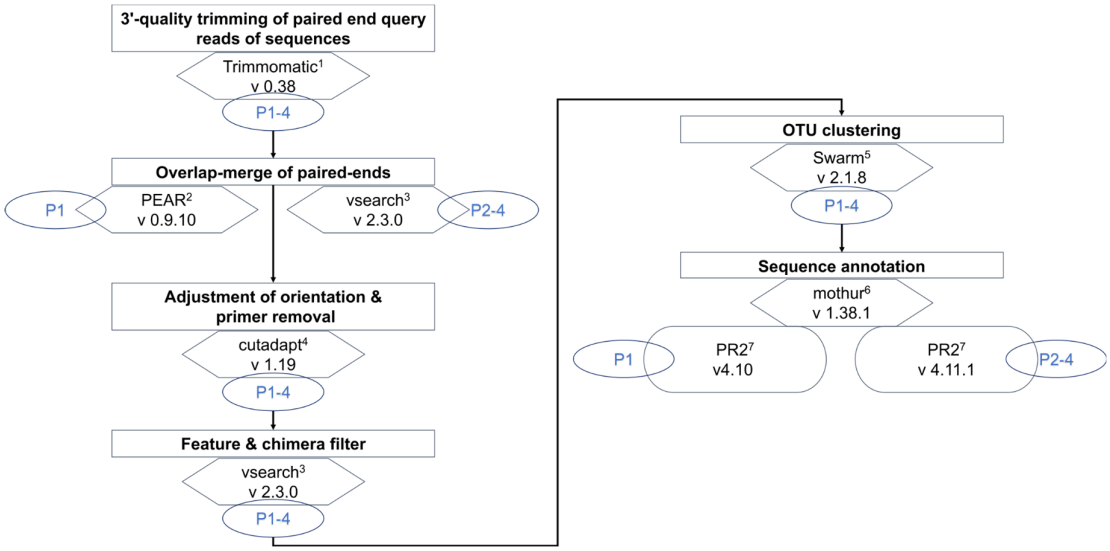


Figure 3: Exemplary overview of bioinformatic pipeline used in this thesis to process metabarcoding data. See https://github.com/PyoneerO/qzip for detailed information, supplemental material from **Publications I** and **II** for exact settings, and supporting information from **Publication II** for comparison with different settings, [1] Bolger *et al*. (2014), [2] Zhang *et al.* (2014), [3] Rognes *et al.* (2016), [4] Martin (2011), [5] Mahé *et al.* (2014, 2015b), [6] Schloss *et al.* (2009), [7] Guillou *et al.*(2013).

In addition to methodological biases of the platforms, pipelines, and reference databases, metabarcoding of planktonic communities faces the same obstacles as other molecular methods. These may include a methodical bias in PCR procedures or organismal biases related to the investigated organisms themselves (Table 1). One issue is the number of gene copies, which not only varies between different taxa but also depends on the size and life stage of the respective organisms (LaJeunesse *et al.*, 2005; Zhu *et al.*, 2005; Connolly *et al.*, 2008; Hong *et al.*, 2016). Despite all these biases, metabarcoding is a potentially useful tool for all kinds of plankton research, in particular for diversity and community analyses (Santoferrara, 2019). It has widely been suggested for use in monitoring plankton dynamics as well as in long-term ecological monitoring (Stern *et al.*, 2018; Ruppert *et al.*, 2019; Compson *et al.*, 2020). This thesis follows in this vein with metabarcoding studies investigating the plankton community at Helgoland Roads. The various studies are presented in the following chapters.

4 Publication I

Publication I, namely: “Rapid succession drives spring community dynamics of small protists at Helgoland Roads, North Sea”, aimed to investigate the community structure and dynamics of microbial eukaryotes during spring 2016. By using 18S metabarcoding, we showed the diversity of nano- and picoplankton during a spring bloom (aim 1a of this thesis). Little was known about these fractions previously because of the difficulties of identifying them by microscope. Instead of autotrophic diatoms, we found mixo- and heterotrophic taxa in higher relative abundances (of the Illumina reads per sample) with dinoflagellates being the biggest contributor in the bigger size classes. However, several other eukaryotic microbes belonging to the pico- and nanoplankton were also found in high relative abundances. Besides the autotrophic *Phaeocystis* and *Emiliania,* parasitoid Syndiniales (which are included in the Dinoflagellata), heterotrophic Picozoa, MAST, and Choanoflagellida were highly abundant during spring 2016. The plankton community included several taxa that were present at all times. The presence and abundance of other species shifted in relation to abiotic dynamics in the water column over the course of the spring bloom (aim 1b of this thesis). We also studied whether the typical succession of diatoms and dinoflagellates as observed by microscopy was also detectable through metabarcoding (aim 2a of this thesis). We found large discrepancies between microscopy and the metabarcoding data as metabarcoding did not detect the diatom bloom, as was registered by the microscopic counts.

Rapid succession drives spring community dynamics of small protists at Helgoland Roads, North Sea

Laura Käse [1,*], Alexandra C. Kraberg[2], Katja Metfies[2,3], Stefan Neuhaus[2], Pim A.A. Sprong[2], Bernhard M. Fuchs[4], Maarten Boersma[1,5] and Karen H. Wiltshire[1,6]

[1]Alfred-Wegener-Institut, Helmholtz-Zentrum für Polar- und Meeresforschung, Biologische Anstalt Helgoland, 27498 Helgoland, Germany

[2]Alfred-Wegener-Institut, Helmholtz-Zentrum für Polar- und Meeresforschung, 27570 Bremerhaven, Germany

[3]Helmholtz-Institut für Funktionelle Marine Biodiversität, 26129 Oldenburg, Germany

[4]Department of Molecular Ecology, Max Planck Institute for Marine Microbiology, 28359 Bremen, Germany

[5]University of Bremen, 28359 Bremen, Germany

[6]Alfred-Wegener-Institut, Helmholtz-Zentrum für Polar- und Meeresforschung, Wadden Sea Station, 25992 List auf Sylt, Germany

*corresponding author: laura.kaese@awi.de

Received October 14, 2019; editorial decision March 13, 2020; accepted March 13, 2020

Corresponding editor: John Dolan

Published in Journal of Plankton Research,
Volume 42, Issue 3, May/June 2020, Pages 305–319,
https://doi.org/10.1093/plankt/fbaa017, first published online May 14

Sequence data for this study have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession number PRJEB37135. All supplementary data can be found at Journal of Plankton Research online, supplementary figures and text can also be found in the appendix of this thesis.

Abstract

The dynamics of diatoms and dinoflagellates have been monitored for many decades at the Helgoland Roads Long-Term Ecological Research site and are relatively well understood. In contrast, small-sized eukaryotic microbes and their community changes are still much more elusive, mainly due to their small size and uniform morphology, which makes them difficult to identify microscopically. By using next-generation sequencing, we wanted to shed light on the Helgoland planktonic community dynamics, including nano- and picoplankton, during a spring bloom. We took samples from March to May 2016 and sequenced the V4 region of the 18S rDNA. Our results showed that mixotrophic and heterotrophic taxa were more abundant than autotrophic diatoms. Dinoflagellates dominated the sequence assemblage, and several small-sized eukaryotic microbes like Haptophyta, Choanoflagellata, Marine Stramenopiles and Syndiniales were identified. A diverse background community including taxa from all size classes was present during the whole sampling period. Five phases with several communities were distinguished. The fastest changes in community composition took place in phase 3, while the communities from phases 1 to 5 were more similar to each other despite contrasting environmental conditions. Synergy effects of next-generation sequencing and traditional methods may be exploited in future long-term observations.

KEYWORDS: phytoplankton; diversity; German bight; Illumina MiSeq sequencing; Long-Term Ecological Research (LTER)

Introduction

Planktonic eukaryotic microbes as encompassed by the term “phytoplankton” represent a diverse array of plankton groups of all size classes including pico- and nanoplankton. They comprise the most frequent autotrophic groups such as diatoms, coccolithophores, green algae and cyanobacteria, but also dinoflagellates, which contain autotrophs, as well as heterotrophs and mixotrophs (Sournia *et al*., 1991; Simon *et al*., 2009). Photoautotrophic phytoplankton is responsible for half of the global primary production (Field *et al*., 1998). Primary producers are important not only as a food source for their consumers but also for bacterial plankton, as bacteria can feed on their excretory products or internal storage compounds after cell death in the form of dissolved or particulate organic matter (Sherr and Sherr, 2002). Microbial mixotrophic and heterotrophic consumers (e.g. choanoflagellates, cryptophytes, dinoflagellates) can feed on the heterotrophic bacterioplankton (bacterivorous protists) or on phytoplankton (herbivorous protists) and are themselves food for the higher trophic zooplankton. Thus, planktonic eukaryotic microbes play an important role in the so-

called microbial loop (Azam *et al*., 1983; Sherr and Sherr, 2002; Caron and Hu, 2019). All size classes, including nano- and picoplankton, are present at different trophic levels of the planktonic community. However, thus far, these are barely identifiable to species level by traditional microscopic methods because of their miniscule size and uniform morphology.

On a global scale, phytoplankton growth periods vary depending on the climate zone; while long growth periods with low biomass occur mostly in tropical and subtropical regions, short growing periods with high biomass have been recorded for high-latitude regions (Racault *et al*., 2012). How different components of the eukaryotic microbial community are present throughout the year in the North Sea and at Helgoland is governed by many abiotic and biotic factors (Reid *et al*., 1990; Wiltshire *et al*., 2015), and species often show distinct seasonal succession patterns (Scharfe and Wiltshire, 2019). During winter there is not much light available, and temperature is low in temperate regions; however, towards the end of winter, nutrient concentrations are high due to remineralization. This leaves optimal conditions specifically for autotrophic organisms like diatoms to bloom once temperature and, more importantly, light availability increase. These spring blooms neither manifest with the same species, nor are they dominated by one species throughout the bloom. Instead they are often a mixture of species at the beginning and show a distinct succession on different timescales (Lewandowska *et al*., 2015; Scharfe and Wiltshire, 2019). Traditionally in spring blooms in temperate regions, e.g. at the Long-Term Ecological Research (LTER) site Helgoland Roads in the German Bight (Wiltshire *et al.*, 2010), diatoms are considered as the major phytoplankton bloom components, showing distinct and massive blooming patterns (Mieruch *et al*., 2010). The bloom of autotrophic phytoplankton is then typically followed by dinoflagellates, heterotrophic plankton or larger zooplankton such as copepods (Lewandowska *et al*., 2015; Wiltshire *et al*., 2015).

The knowledge of spring bloom dynamics in specific regions can be validated and extended by implemented time series like the Continuous Plankton Recorder Survey (Reid *et al*., 2003; McQuatters-Gollop *et al*., 2015), the L4 coastal time-series station (Harris, 2010) or Helgoland Roads LTER (Wiltshire and Manly, 2004; Wiltshire *et al*., 2015; Scharfe and Wiltshire, 2019). One potential problem, however, is that traditional time series currently rely on microscopy techniques, such as the Utermöhl method, which is time-consuming and limited by the size of organisms (Stern *et al.*, 2018). This means that the smallest organisms cannot be assigned to taxonomic level accurately (Culverhouse, 2015). Therefore, especially small protists are barely investigated due to the resolution limit of the identification methods used in traditional long-term observations.

New molecular methods, and especially next-generation sequencing (NGS), could have a high potential for very detailed monitoring (Ebenezer *et al*., 2012; Stern *et al*., 2018) as these molecular methods reliably capture the entire phytoplankton community including nano- and picoplanktonic components. For example, seasonality patterns could be found for at the Adventfjorden time-series station using 454 sequencing (Marquardt *et al*., 2016) and in the Mediterranean Sea using Illumina sequencing (Giner *et al*., 2019). Seasonal patterns as well as diel shifts in activity could be found using the V4 region of RNA and DNA in Illumina sequencing in the North Pacific (Hu *et al*., 2016, 2018). Other studies have been conducted, focusing on different European coastal waters like the L4 time-series station in the Western English Channel (Taylor and Cunliffe, 2014), several stations along the European coast within the BioMarKs project (Logares *et al*., 2014; Massana *et al*., 2014, 2015) or estuaries, e.g. in the eastern English Channel (Bazin *et al*., 2014). Some studies only focus on certain taxa, e.g. uncultured marine heterotrophic flagellates (Logares *et al*., 2012) or Chlorophyta (Tragin *et al*., 2018). With regard to prokaryotic monitoring, several NGS studies were conducted at Helgoland (Lucas *et al*., 2015, 2016; Teeling *et al*., 2016; Chafee *et al*., 2018). While several studies have been conducted in the general North Atlantic at large, only a few studies focusing on specific groups of small-sized eukaryotic protists have been done using other molecular methods, which are focused specifically on Helgoland (Medlin *et al*., 2006, 2017; Gescher *et al*., 2008; Knefelkamp, 2009; Metfies *et al*., 2010).

This study aims to (1) understand the community structure and dynamics of eukaryotic protists including the pico- and nanoplankton fraction during spring and (2) discover if the typical spring bloom succession of diatoms and dinoflagellates can be detected using NGS data at Helgoland Roads from 15 March to 31 May 2016. (3) We also aim to relate abiotic dynamics in the water column to taxonomic group shifts in the community during the spring bloom based on a much more detailed assessment of phytoplankton biodiversity.

Materials and Methods

In total, we took 50 plankton samples during spring 2016, analysed these samples using next-generation sequencing (18S) and investigated successional patterns.

Study site and sampling

Sampling was conducted at the Helgoland Roads LTER sampling site at the station “Kabeltonne” (54°11.03’ N, 7°54.00’ E, Germany) (Wiltshire and Dürselen, 2004). The sampling site is situated between the main island and the dune island of Helgoland. The generally well-mixed water column fluctuates between 6 and 10 m depth, depending on the

tides (Callies and Scharfe, 2015). Samples were taken from 1 m depth between 15 March and 31 May 2016. Sampling frequency was work-daily, according to the LTER sampling. About 1 L of seawater was sequentially filtered using 10 µm polycarbonate filters (PC), 3 µm PC filters and 0.2 µm polyvinylidene fluoride filters (Millipore, Schwalbach, Germany) to obtain the whole prokaryotic and eukaryotic plankton community (Teeling *et al*., 2016). Secchi depth and temperature were measured directly in the water at the sampling site. Other parameters, including salinity and nutrients such as silicate, phosphate and inorganic nitrogen using the methods of Grasshoff (1976), were measured in the laboratory according to the LTER protocols (Hickel *et al*., 1993; Wiltshire *et al*., 2008, 2010). Daily observations of sunshine duration in hours were downloaded from the Deutscher Wetterdienst, Climate Data Centre (2019). To check whether the spring of 2016 showed a typical phytoplankton community succession of diatoms followed by dinoflagellates as observed in the LTER, we used total diatom and total dinoflagellate counts and chlorophyll *a* measured by HPLC modified after Zapata *et al*. (2000) from 1st March to 31st May.

DNA extraction and pooling of samples

The DNA extraction from 0.2 µm filters was conducted as described previously at the Max Planck Institute for Marine Microbiology (Bremen, Germany) (Sapp *et al*., 2007). In short, lysozyme (1 mg mL^{-1}) and sodium dodecyl sulphate (1%) were used for cell lysis; DNA was extracted with a phenol/chloroform/isoamyl alcohol mixture (25:24:1) and precipitated with isopropanol, before the DNA extracts were eluted in sterile water. This fraction, which was previously used for 16S analysis, was then added to the other fractions to include all potential eukaryotes in all size ranges. The DNA from the 10 and 3 µm filters was extracted following the manual of the Macherey–Nagel NucleoSpin® Plant II Kit, and all extracts were stored at −20°C. To include the whole eukaryotic plankton community from all size classes, equal volumes of the DNA extracts of the smallest size fraction (0.2 µm pore-size filters) were then pooled with the DNA extracts of the remaining size fractions (3 and 10 µm) to obtain one sample per sampling date. Measurement of nucleic acid content of the pools was conducted with a fluorometer (QuantiFluor® dsDNA System, Promega, USA).

MiSeq™ Illumina sequencing

After pooling, the samples were prepared for MiSeq™ Illumina sequencing following the Nextera XT DNA Library Preparation protocol (Illumina, USA) with the following modifications: a fragment (V4 region) of the 18S ribosomal (r) DNA was amplified using KAPA HiFi HotStartReadyMix (Kapa Biosystems, Inc., USA) and the following primer set: 528iF (GCG GTA ATT CCA GCT CCA A) and 964iR (ACTTT CGT TCT TGA TYR R) (Fadeev *et al*., 2018). The success of this amplicon PCR was confirmed with gel

electrophoresis using 2 µL of the PCR product. If no bands were detected, the amplicon PCR was repeated with an increased template volume (up to 5 µL). If this still was not sufficient to detect the respective band, five additional cycles were added to the original program (eight samples). Before library normalization and pooling, the DNA concentration was once again measured using a Quantus Fluorometer (Promega, USA) and diluted accordingly. Amplicon sequencing was then performed on an Illumina MiSeq™ sequencer (Illumina, USA), and about 6.3 million 2×300 bp paired-end reads were produced in total.

Bioinformatics processing

Sequence processing, operational taxonomic unit (OTU) clustering and annotation were done with an internally developed pipeline at the Alfred Wegener Institute as described below (detailed description as Supplemental Material), wrapping common bioinformatics tools and “GNU parallel” (Tange, 2011) for fast and massive parallel workflow execution. The low-quality 3’-ends of the reads were trimmed by Trimmomatic, version 0.38 (Bolger *et al*., 2014), and the paired-ends were merged by PEAR, version 0.9.10 (Zhang *et al*., 2014). Cutadapt, version 1.17 (Martin, 2011), was used to adjust the sequence orientation and remove the forward and reverse primer matching sequence segments. Sequences were only kept if both primer matching segments could be detected. The remaining sequences were filtered by VSEARCH, version 2.3.0 (Rognes *et al*., 2016), and sequences were discarded, (i) if they were 50 bp longer or shorter than the median length of the targeted amplicon (376 bp), (ii) if they carried any ambiguity or (iii) if the expected base error (sum of all base error probabilities) of a sequence was above 0.5. Chimeric sequences were sample-wise predicted by VSEARCH, version 2.3.0, in de novo mode with default settings and removed from the sample files. Only samples with at least 10 000 sequences after filtering were considered for further analyses (49 out of 50 samples). The remaining 4.3 million sequences were clustered into OTUs by the tool swarm, version 2.1.8 (Mahé *et al*., 2014, 2015), with default settings. For each OTU the most abundant amplicon was selected as representative and taxonomically annotated with the default classifier implemented in mothur, version 1.38.1 (Schloss *et al*., 2009). As reference the Protist Ribosomal Reference database (PR2), version 4.10 (Guillou *et al*., 2013), was chosen and the confidence cut-off was set to a value of 90.

A conservative threshold of 0.005% (of total reads) after Bokulich *et al*. (2013) was applied to the remaining 37 608 OTUs, leaving 694 OTUs present in the 49 samples. After removal of Metazoa alignments, 587 OTUs were used for further analysis to determine the protist community. Identification up to genus level was accepted as species annotations were generally poor. Higher taxonomic levels included family, class, order, phylum and kingdom

level. For taxa that could not be further identified, the previous higher taxonomic level was adopted and additions to the name were attached (e.g. unclassified) and counted as a different taxon on the respective taxonomic level.

Statistical analysis

All statistical analyses were conducted in R, version 3.5.0 (R Core Team, 2018). The following packages were used for visualization: ggplot2 (Wickham, 2016), dendextend (Galili, 2015), ampvis2 (Andersen *et al*., 2018), RColor- Brewer (Neuwirth, 2014), gplots (Warnes *et al*., 2019) and gridExtra (Auguie and Antonov, 2017). For significance tests, the significance level was set at $P < 0.05$.

For identification of significant abiotic correlations to our OTU abundance table, which was normalized to the total number of reads per sample, temperature, salinity, Secchi depth, tide and sunshine duration as well as silicate, nitrate, ammonium and phosphate concentrations were added to a “Constrained Ordination Model”. This model was based on an ANOVA-like permutation test for canonical correspondence analysis (CCA) to assess the significance of the constraining factors, by testing for single term additions (Oksanen *et al*., 2019). Single variables were chosen by their significance and added to the next step in the model, before the next significant variable was added in the next step. If several variables were given as significant, the variable with the lowest Akaike information criterion (AIC) value was chosen first to minimize the information loss (Akaike, 1974).

After calculation of the alpha diversity of the different taxonomic levels, the proportion of unclassified taxa – taxa that could not be determined and assigned by the PR2 database – were summarized and compared. Non-metrical multidimensional scaling (NMDS) plots were created in vegan with Bray–Curtis dissimilarities to compare the community composition of the samples on different taxonomic levels (Oksanen *et al*., 2019). Hereby, the data were converted to presence–absence data at genus and at phylum level. Beta diversity was calculated on genus level using the betadiver function vegan (Oksanen *et al*., 2019) and Whittaker index (Whittaker, 1960). To visualize the matrix, it was converted into a cluster with the hclust function. The phases that were chosen after comparing the NMDS plot with the beta diversity clusters were then tested for significance with an Analysis of Similarities (ANOSIM). A distance matrix of the phases defined by the beta diversity analysis was compared to the significant environmental parameters separately as they were defined by the CCA using a Mantel test from the ade4 package (Dray and Dufour, 2007; Bougeard and Dray, 2018). For the dissimilarity matrices of the determined phases and of environmental parameters, an Euclidean distance metric was used. To determine the most abundant genera, further analysis was based on the relative abundance of the Illumina

reads per sample. To calculate the relative abundance, the dataset was normalized to the total number of reads per sample. Here, the most abundant genera had a relative sequence abundance of more than 5% in at least one sample during the whole period.

To define OTUs of interest, we conducted a Similarity Percentage analysis (SIMPER). The SIMPER analysis helps to identify those OTUs that contribute the most to the variation between the different phases. Using the phases that were defined based on the beta diversity calculation in the simper.pretty function (Steinberger, 2018), the OTUs with the biggest contributions to the similarity between two phases were identified. Hereby, OTUs that contributed less than 1% were removed. Afterwards the kruskal.pretty function (Steinberger, 2018) was used to find significant differences across the phases that were defined by the beta diversity calculation. The significant OTUs were then assigned to their respective genera and visualized as a heatmap. To find clusters of OTUs on presence–absence level, we used hierarchical cluster analysis with multiscale bootstrap in the parallel parPvclust function using the package pvclust (Suzuki and Shimodaira, 2006; Ryota Suzuki and Shimodaira, 2015). By development of a dendrogram with additional bootstrapping procedures, it is possible to calculate the significance of each cluster in the dendrogram. The number of bootstraps was elevated to 20 000 to minimize the standard error of the resulting clusters of OTUs. The calculation of distances for the hierarchical cluster was based on the asymmetric binary method, because the data are based on presence–absence level. For agglomeration, the complete linkage method (farthest neighbour clustering) was set. The pvpick function was used to find clusters with significant *P*-values. Support of data for these clusters was validated by manual estimation and comparison of the confidence interval to the respective *P*-values.

Results

Environmental parameters and spring bloom succession as observed in the LTER

The water temperature at Helgoland Roads was 5.9°C on 15 March and gradually increased to 13°C until the end of May (Fig. 1, Supplementary Table SI). Salinity ranged from 31.5 to 33.7, showing fluctuations throughout the period. Silicate concentrations rose from 5.3 µM to reach a maximum on 21st March with 13.4 µM. At the end of March, concentrations declined and remained below 3 µM. Secchi depth varied throughout the sampling period between 1.8 and 7.0 m with several fluctuations. Daily sunshine duration varied greatly from day to day and ranged from 0 h of sunshine on 5th, 13th and 29th April and from 23rd May to 26th May up to 14.4 h of sunshine (12th May).

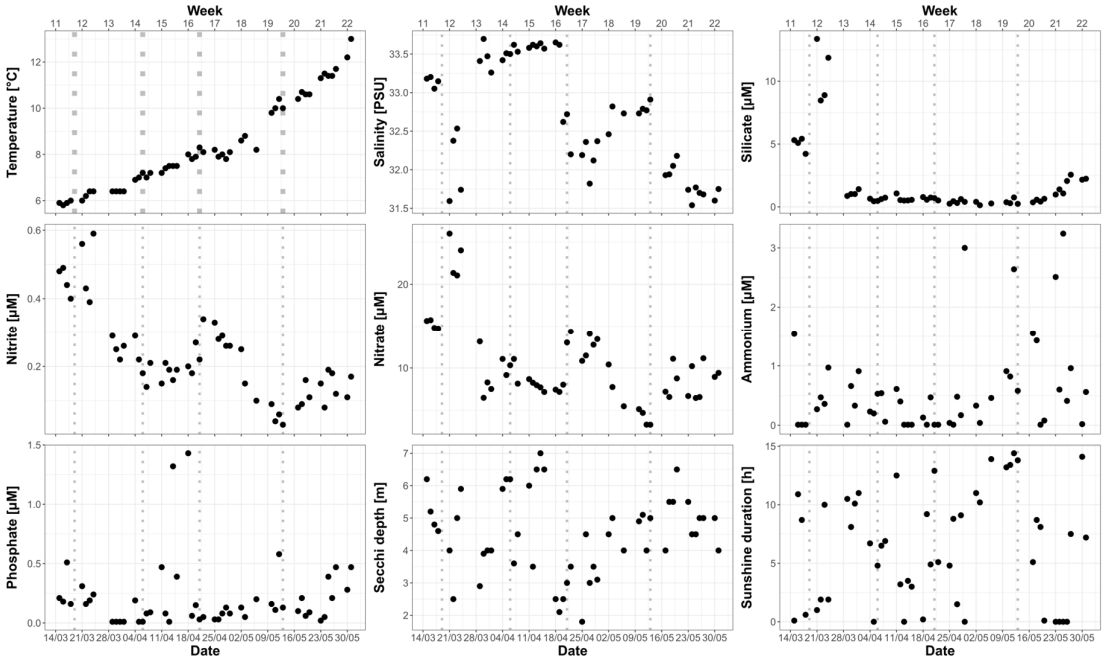


Fig.1. Profiles of temperature [°C], salinity, silicate [µM], nitrite [µM], nitrate [µM], ammonium [µM], phosphate [µM], Secchi depth [m] and sunshine duration [h] at Helgoland Roads LTER sampling station during spring 2016; vertical dotted lines indicate the different phases as defined by beta diversity analysis.

The LTER microscopic counts revealed a pattern, which resembled a typical spring phytoplankton succession with high diatom abundances, followed by a peak in dinoflagellates (Fig. 2a). Diatoms showed highest abundances (3116∗103 cells L^{-1}) from week 14 to 16 (April) as well as during week 19 (May) (2795∗103 cells L^{-1}). Dinoflagellate total counts revealed a maximum abundance at the end of May (week 21) (111∗103 cells L^{-1}). In the beginning of March, HPLC chlorophyll *a* (Fig. 2a) was below 1.00 µg L^{-1} and increased to reach a first peak on 17th March (3.97 µg L^{-1}). In contrast to the diatom maximum peak, chlorophyll *a* reached its peak on 29th March (week 13) with 6.77 µg L^{-1}. Afterwards the concentration gradually declined with two maxima interrupting this trend on 19th April and 10th May, at 2.61 and 2.39 µg L^{-1}, respectively.

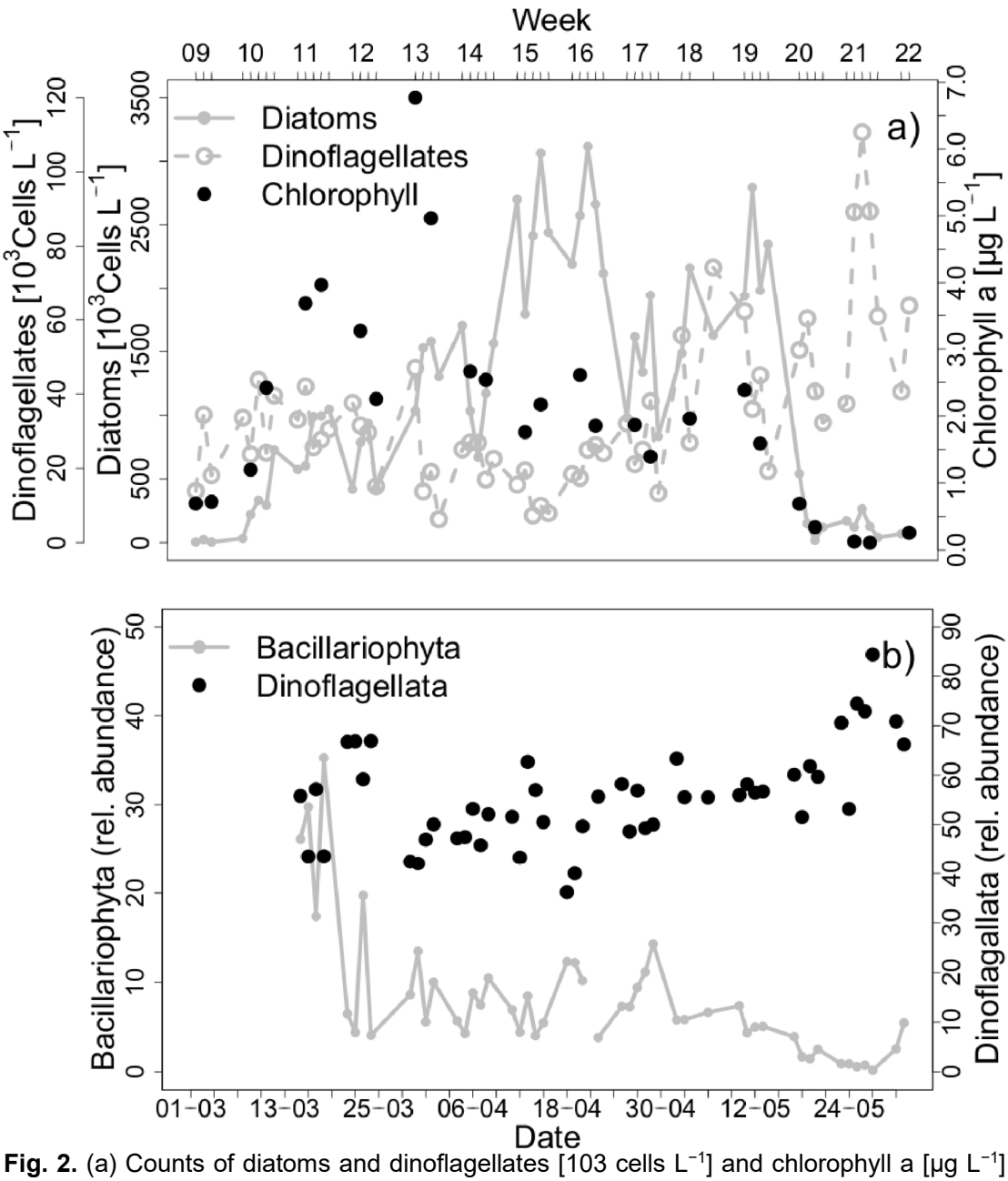


Fig. 2. (a) Counts of diatoms and dinoflagellates [103 cells L^{-1}] and chlorophyll a [µg L^{-1}] measured with HPLC at Helgoland Roads LTER station from March 1 to 31 May 2016; (b) relative abundance [%] of Bacillariophyta and Dinoflagellata from 15 March to 31 May 2016.

General description of the sequencing dataset

After quality control, 587 OTUs were assigned to 21 phyla. Identification was conducted up to genus level (Fig. 3). Based on the total number of OTUs that was analysed, approximately 96% could be assigned at kingdom level. Assignment at phylum and class level was possible with more than 90% of the OTUs. At order and family level, 76 and 65% of all OTUs could be assigned, respectively. Most genera were represented by several OTUs. Examples are the dinoflagellate Gyrodinium, which was represented by nine OTUs, or the diatom Chaetoceros, which was assigned to seven OTUs. Overall, reliable identification at

genus level was possible for only 29.3% of OTUs (83 genera), which indicates that the biggest information gap regarding taxonomic assignments occurs between family and genus level.

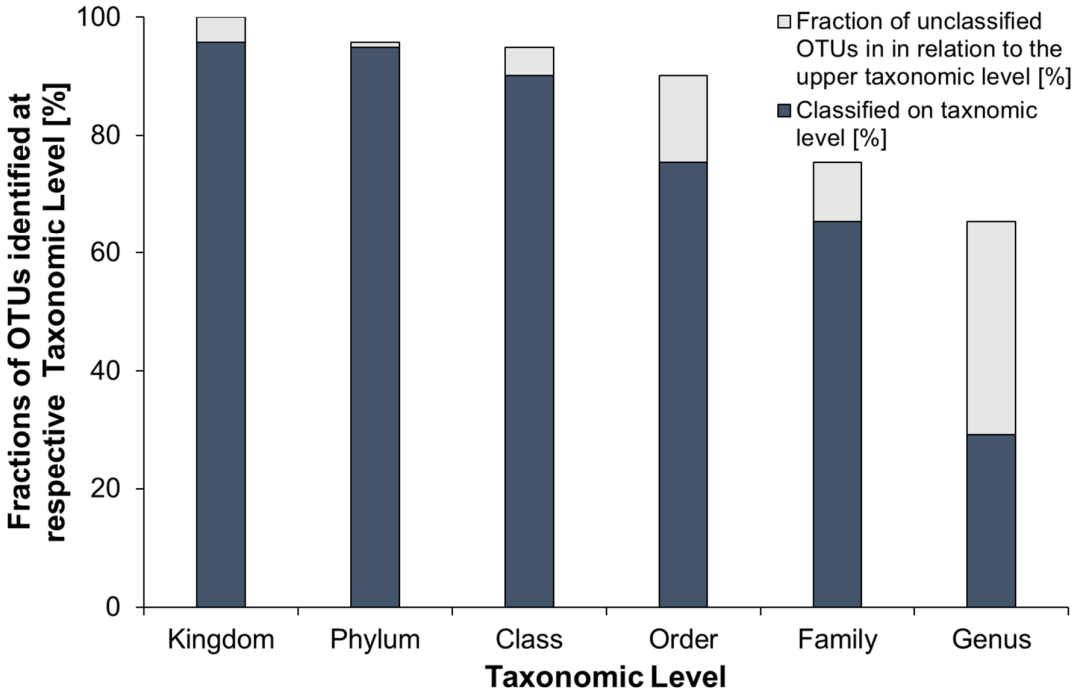


Fig. 3. Fractions of OTUs identified at respective taxonomic level: dark grey indicates that identification on the respective taxonomic level was successful, light grey indicates that identification information did not go beyond the previous level; it includes all unclassified taxa (marked with a suffix_unclassified) and taxa where monophyly could not be insured (marked with a suffix_X according to the database).

Temporal dynamics in the community

As shown in the 2D NMDS plots of community dissimilarities at presence–absence level (See online supplementary Fig. S1 for a colour version of this figure), a temporal pattern was found at genus level. However, the different communities are not visible at phylum level, since all phyla are represented by several genera that are always present. In general, beta diversity revealed a maximum species turnover of ~25% (Fig. 4). Five different phases could be identified during the spring bloom: phase 1 during week 11, phase 2 from week 12 to week 14, phase 3 from week 14 to week 16, phase 4 from week 16 to week 19 and phase 5 from week 19 to week 22 (see also Supplementary Table SI). The ANOSIM confirmed the significance of these clusters (R = 0.7, significance = 0.001).

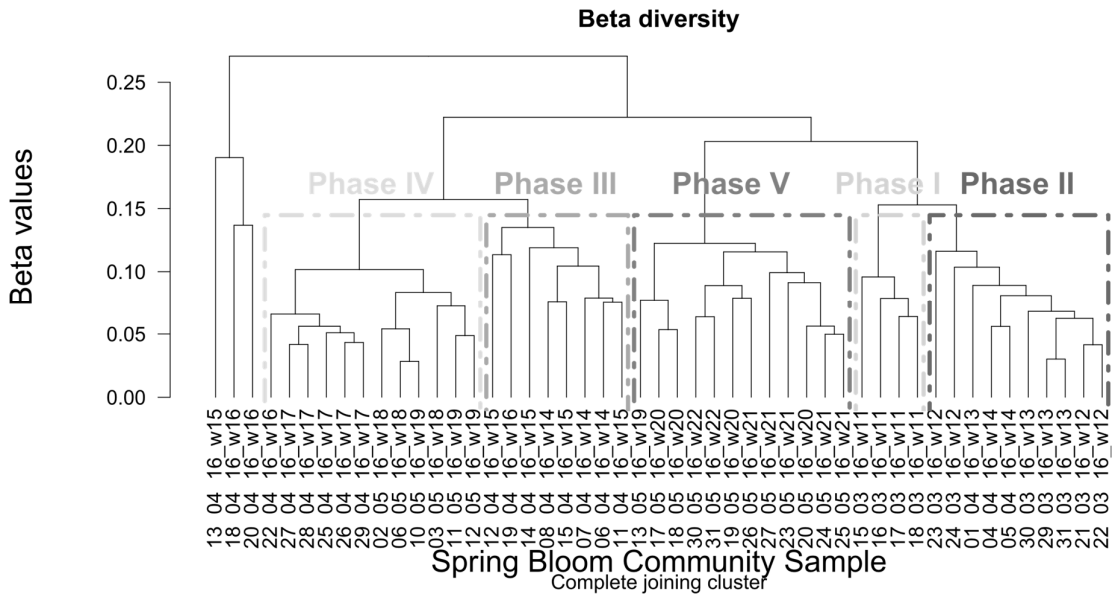


Fig. 4. Beta diversity of the different samples during spring 2016. It was calculated using the betadiver function (vegan package) and Whittaker index; visualization of the matrix was done with the hclust function.

For this spring bloom period, temperature (AIC = 251.73, *P* = 0.005) was found to be the most important environmental parameter based on the CCA model, followed by silicate (AIC = 247.26, *P* = 0.005), salinity (AIC = 245.95, *P* = 0.005), sunshine duration (AIC = 245.75, *P* = 0.005) and tide (AIC = 245.66, *P* = 0.005). Other parameters tested in the model that were not significant were nitrate, phosphate, ammonium and Secchi depth. The CCA plot (Fig. 5) indicated that at the beginning of the study period, the community was mostly correlated with silicate concentration. During April, this correlation shifted towards salinity which increased in April. Especially samples from the end of April and beginning of May were correlated to sunshine duration and low tide (information on tides can be found in Supplementary Table SI). The strongest correlation for the May community was with higher temperature.

The follow-up Mantel test revealed that the environmental factors temperature (*r* = 0.5738, *P* = 0.001) and salinity (*r* = 0.3483, *P* = 0.001) were significantly correlated to the beta diversity patterns, while silicate, sunshine duration and tide were not. Especially in phase 3, high variations in the community assemblage could be observed. Compared to the other phases, there were greater daily fluctuations in the community composition during phase 3. Three samples (13th, 18th, 20th April), which were taken during the same period, where the community of phase 3 was identified, showed higher variations in community composition and therefore could not be assigned to any phase.

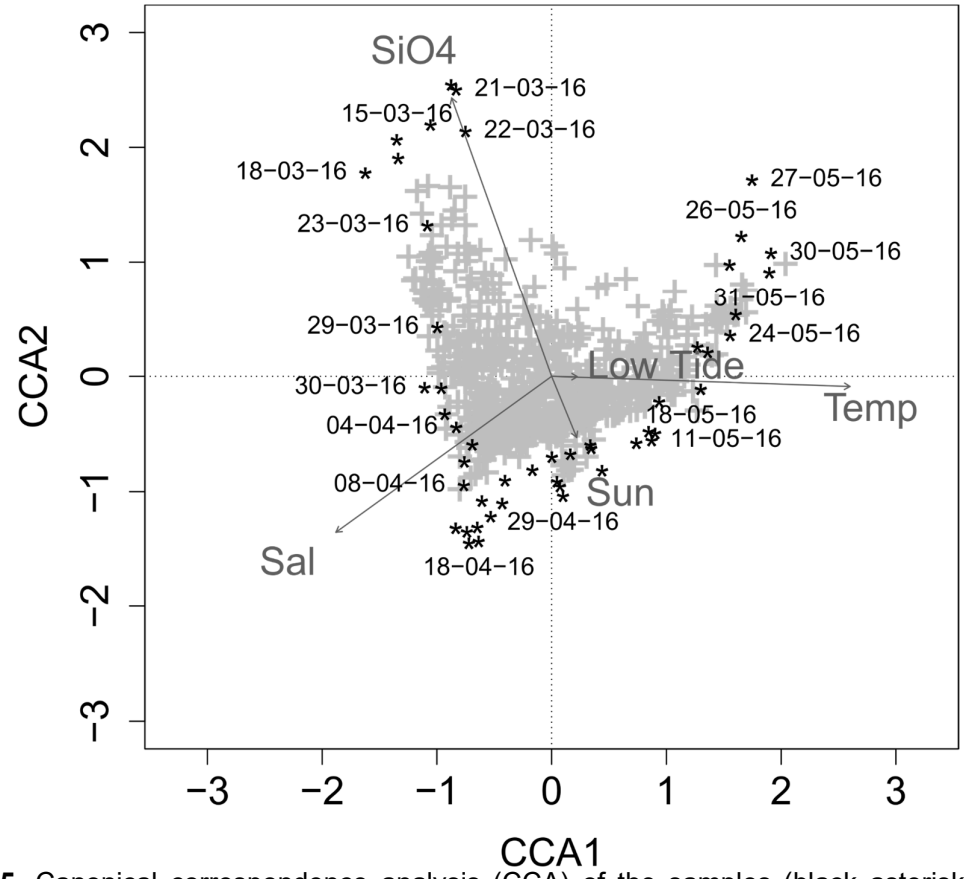


Fig. 5. Canonical correspondence analysis (CCA) of the samples (black asterisks with sampling date) including abiotic factors in dark grey: temperature (Temp), salinity (Sal), silicate (SiO4), sunshine duration (Sun) and tide (Low Tide); OTUs in a light grey plus symbol, 37.9% of total inertia, could be explained by all variables, CCA1 explained 17.5% of the variance and CCA2 explained 10.5%.

SIMPER analysis showed that 53 OTUs explained at least 1% each of the variation between the five phases. With a Kruskal–Wallis test, 37 of these OTUs were found to be significantly different (Supplementary Table SII). These OTUs were from six different phyla and 28 genera, respectively. The number of contributing OTUs was increasing with later phases, and the highest number was found for phase 5 (15 OTUs). In total, 21 genera of 8 different phyla were found as most abundant (Fig. 6). Out of these 21 genera, 10 could not be assigned at genus level and 10 genera belonged to dinoflagellates. Unclassified Gymnodiniales and unclassified Dinophyceae OTUs contributed the most to the communities during all phases. The Ochrophyta genus *Ditylum,* followed by *Pseudo-nitzschia*, also contributed to the change in the overall community in all compared phases (Supplementary Table SII). *Ditylum* had the highest relative abundance during phase 1 and was declining at a fast rate during phase 2 and absent beyond phase 3 (Fig. 6). When comparing phases 2 and 3, unclassified Dinophyceae were identified as the biggest contributor to changes in the community, followed by the dinoflagellate *Heterocapsa* and heterotrophic Marine Stramenopiles-1A (MAST-1A). Regarding phase 3, 12 out of 16 OTUs,

which contributed to the variation when compared to phase 4, belonged to dinoflagellate genera. At the end of phase 3 to the beginning of phase 4, a peak of *Phaeocystis* sp. abundances could be observed. Abundances of the nanoplanktonic coccolithophore *Emiliania* sp. rose to a peak during phase 4, with approximately 30% in relative abundance on 18 May 2016. This single *Emiliania* OTU had the biggest influence on the changes in community from phase 4 to phase 5. Over the complete sampling period, the same *Emiliania* OTU was found to have the largest influence when comparing phases 1 and 5, as abundance was increasing during the sampling period.

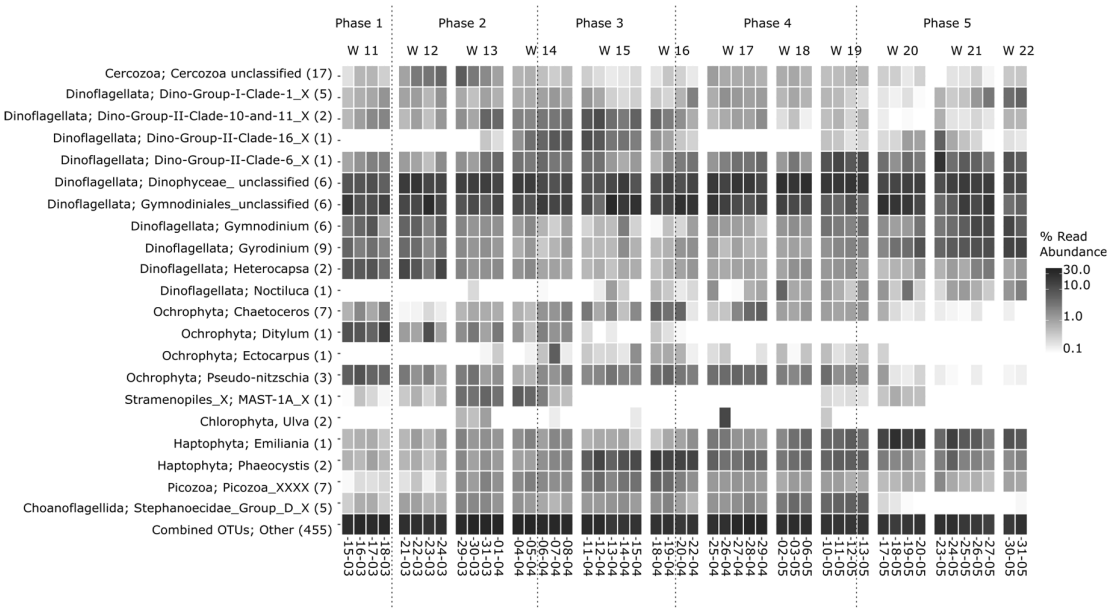


Fig. 6. Most abundant genera from 15 March to 31 May 2016. Number of OTUs per genus indicated in parentheses. Shown are all genera with a relative sequence abundance of more than 5% in at least one sample during the whole timeframe.

Community structure and diversity of pico- and nanoplankton

Based on all 587 OTUs, the first phase consisted of the highest proportion of autotrophs and mixotrophs (on average ~60% in total) and 40% of heterotrophs (see Supplementary Table SIII for summarized suggested trophic modes; trophic modes were defined based on the taxonomy and known information from literature as Gomez (2012), Kubiszyn *et al*. (2014) and the Tara Oceans Database W4 from the Companion Website of the article of de Vargas *et al*. (2015); if the last identified taxon was on a higher taxonomic level, we assumed the likelier/more frequent trophic mode when suitable, otherwise no trophic mode was assigned.). For all other phases, heterotroph OTUs contributed the most with over 50%. Ochrophyta (See online supplementary Fig. S2 for a colour version of this figure), which were mostly represented by autotroph Bacillariophyta (diatoms, see online supplementary Fig. S3 for a colour version of this figure), were most abundant in phase 1. Single genera

like *Chaetoceros* or *Pseudo-nitzschia* were also abundant during phases 3 and 4 (Fig. 6). During phase 5, diatom abundances were always low (<10%).

In accordance with the most heterotrophic phases 2-5, Dinoflagellata (See online supplementary Fig. S2 for a colour version of this figure) had consistently the highest relative abundances during the whole period, with relative abundances ranging from 36.2 up to 84.4%. Highest abundances were reached in phase 5. Three different classes of dinoflagellates (Dinophyceae, Noctilucophyceae and Syndiniales) could be identified (See online supplementary Fig. S3 for a colour version of this figure). Whereas Dinophyceae and Noctilucophyceae consist of mostly bigger sized dinoflagellates, Syndiniales consist of mostly picoplanktonic parasites. Dinophyceae were the biggest contributors to the community for most days, followed by Syndiniales, which were the biggest contributor in phases 3 and 5. The high contribution of dinoflagellate taxa was also visible in the number of taxa that were most abundant during this timeframe.

The next most important phylum was Haptophyta (See online supplementary Fig. S2 for a colour version of this figure). This pico- and nanoplanktonic phylum increased steadily in abundance during phases 1 and 2. High abundances with a maximum of 32.2% started from phase 3 onwards until the end of phase 5. Other phyla (See online supplementary Fig. S2 for a colour version of this figure) included the heterotroph Cercozoa (mostly unclassified), which showed high abundances during phase 2 (maximum 19.9%), but were generally low (<10%) before and after this period. Heterotrophic Stramenopiles represented by the pico- and nanoplanktonic MAST (See online supplementary Fig. S3 for a colour version of this figure) were mostly present during phases 2, 3 and 5 (24 different genera, in total 37 OTUs). During phase 2, a maximum abundance of 15.7% was reached, whereas during phase 5 the highest abundance was below 8%. Nine out of 21 phyla always had relative abundances below 1% (See online supplementary Fig. S2 for a colour version of this figure).

Based on presence–absence data, 33 significant clusters of OTUs were detected with hierarchical cluster analysis using 20 000 bootstraps (Supplementary Table SIV). The 33 clusters, representing a community of correlated OTUs, included 229 OTUs (39% of OTUs). Twenty-one of these clusters could be validated with the *P* value being inside the confidence interval. Only four clusters had a relative abundance above 5% (Supplementary Table SV). The biggest cluster (cluster 4) included 79 OTUs. Except for four OTUs, the cluster consisted only of OTUs that were present in every sample and therefore during all phases. Fifty-two genera, belonging to 12 different phyla, could be found in this cluster. The biggest contributors were unclassified Dinophyceae and several pico- and nanoplanktonic MAST

groups (14 OTUs). Another big and diverse cluster was cluster 31 with 11 OTUs, which were mainly found during phases 1 and 2. It included OTUs identified as Cercozoa, Ochrophyta (3 OTUs each), Stramenopiles_X (2 OTUs), Choanoflagellida, Ciliophora and Dinoflagellata (1 OTU each). The 12 significant clusters, where the confidence interval did not support the existence of the clusters, included between two and seven OTUs each. For example, six OTUs were part of cluster 22. Herein, two OTUs were unclassified Eukaryotes and four OTUs belonged to Hacrobia. Three nanoplanktonic cryptophytes (*Falcomonas* sp., *Teleaulax* sp., *Plagioselmis* sp.) and *Leucocryptos* sp. clustered together. Reads for these OTUs were available in phase 1, partially phase 4 and in phase 5.

Discussion

In this work, we could gather new information on several small-sized eukaryotic microbes. We identified nano- and picoplankton such as several Syndiniales (Dino-Groups) and MAST groups, *Phaeocystis* sp. and *Emiliania* sp., which contributed to the communities with high abundances. Additionally, we observed that our sequence assemblage was dominated by dinoflagellates, in contrast to the microscopic count data, and a peak of diatoms was not observed in the dataset.

Environmental parameters and spring bloom succession as observed in the LTER

Our environmental conditions were mostly in accordance with the general pattern described by Wiltshire *et al*. (2015). While temperature and sunshine duration increased during our sampling period, salinity showed abrupt short-term changes. Higher salinity during our sampling indicates either a decreasing influence of riverine inputs or a bigger influence of Atlantic-driven waters during this time. As salinity was falling gradually, the increase in freshwater sources appears more likely. Wiltshire *et al*. (2015) stated that salinity reduction in spring happens mainly due to riverine input in late winter. As a result of incoming water masses, high concentrations of nutrients can be advected into the Helgoland Roads sampling site (Callies and Scharfe, 2015). In addition to biological nutrient cycling, change of water masses therefore can cause shifts in nutrient concentrations.

Comparison of spring bloom conditions regarding diatom and dinoflagellate occurrence

According to the LTER total counts, diatoms were much more abundant than dinoflagellates, and dinoflagellates reached their highest abundances after diatom abundances declined. This phenomenon is in accordance with previous literature from Helgoland Roads as well as other coastal European and North American regions (Lewandowska *et al.*, 2015; Wiltshire *et al.*, 2015; Carstensen *et al.*, 2015). For the Western English Channel the spring

diatom bloom is mostly followed by *Phaeocystis*, coccolithophorids and dinoflagellates (Widdicombe *et al*., 2010). However, for single regions the disappearance of a typical diatom spring bloom has been reported (Nixon *et al*., 2009).

The early peak in chlorophyll *a* measured by HPLC might be caused by picoplanktonic autotrophs (Knefelkamp, 2009) or the simultaneous high abundance of unclassified Cercozoa. Here, heterotrophic Cercozoa could have ingested chlorophyll-containing cells, or the Cercozoa were represented by chlorarachniophytes, which contain chloroplasts (Ishida *et al*., 1999). Also, it has to be noted that chlorophyll *a* sampling frequency was lower (two times a week), compared to the LTER counting data (five times a week).

If we compare the sequencing abundances regarding diatom and dinoflagellate abundances to the LTER total counts, we do not find a good match, even though the sampling frequency was similar and the high sampling frequency minimizes the chance that we missed individual abundance peaks that were seen in the microscopic counts. In addition, the typical decline in silicate concentration supports the presence of diatoms in high abundances. For example, as *Chaetoceros socialis* is known as a colonizing and mucous forming species; potential aggregation of cells needs to be taken into account (Riebesell, 1993). It is unclear to what extent aggregation potential of single species can influence the match in peak abundances for both methods, since aggregates in either sample might lead to overestimation.

With respect to diatoms in general (Ochrophyta), the highest abundances in our sequencing dataset were found in phase 1 at the beginning of the sampling period (week 11), with single genera also abundant during later phases. In total they did not show a distinct peak, but most genera found were in accordance with typical diatoms occurring in the area in spring (Hoppenrath, 2004; Wiltshire and Dürselen, 2004; Kraberg *et al*., 2015; Wollschläger *et al*., 2015). However, important species such as *Guinardia delicatula*, *Thalassionema nitzschioides* and *Odontella aurita*, which are known to have growth periods fitting to our sampling period, could not be found in high abundances. It has been shown that shifts in blooming periods and widening of occurrences of single species occurred in the past (Wiltshire *et al*., 2010; Schlüter *et al*., 2012), which could explain the absence of these species in our sequence assemblage. Comparison to the regular long-term microscopic counts revealed that *O. aurita* and *T. nitzschioides* were only reported for four and two times, respectively, during this timeframe. For *G. delicatula*, counts revealed that the species was mostly present from March to April (data not shown), which is in accordance with the sequencing results.

As the primer set used was engineered to better match contributions of diatoms and *Phaeocystis* sp. to the community, a sequencing bias should be unlikely. However, instead of a diatom-dominated community, our sequence assemblage was dominated by several dinoflagellate taxa. These included a wide diversity of large-sized species, but also potential parasites from different Syndiniales groups. The constant high abundance of dinoflagellates does not correlate with the LTER counts, where abundances steadily grew throughout the sampling period. Both datasets, however, showed the highest abundances in week 21.

So, what drives this conflicting information between microscopic counts and sequencing results? First, it has to be taken into account that the high abundances might be influenced by different dinoflagellate gene copy numbers. The generally high abundance of dinoflagellate genera was similar in previous studies. For instance, Massana *et al*. (2015) and Massana (2011) found mostly dinoflagellates including several parasitic Syndiniales in European coastal waters. Similar high abundances of Syndiniales and Gymnodiniales were found by Taylor and Cunliffe (2014) at the L4 coastal LTER station (Western English Channel). One issue is the use of relative abundances for comparison of communities that is influenced by gene copy numbers per cell, which differ greatly in between species. Several studies have emphasized the different rDNA copies among protist taxa like diatoms (Connolly *et al*., 2008) or dinoflagellates (LaJeunesse *et al*., 2005; Hong *et al*., 2016). Therefore, an approach based only on relative abundances is difficult to interpret. Several analyses in our study such as NMDS and OTU clustering were conducted at presence-absence level to avoid this phenomenon. However, one problem in using this approach was that most genera in this analysis were present at any time during the sampling period. In addition, since a comprehensive and reliable resolution at species level is not possible so far, it is necessary to include the relative abundances as well, if we want to see changes and relationships in the community. Furthermore, species that might be abundant at Helgoland and visible in the traditional long-term series when using microscopy might not be available in our dataset. Reasons for this could be the threshold we used, or a bias in DNA extraction, PCR and sequencing procedures. At the same time, it is possible that dinoflagellate occurrence in the environment is underestimated in microscopic studies, since several small-sized taxa cannot be identified.

Several factors might influence the reliability of sequence identification. Considerable difficulties and possible sources of biases include the use of target molecules (e.g. RNA, rDNA), regions (e.g. V4, V9) and databases like PR2 (Guillou *et al*., 2013) or SILVA (Pruesse *et al*., 2007). These databases are not of equal detail for different taxon groups. For example, identification on genus level for both databases was poor, and a direct comparison between PR2 and SILVA sometimes revealed contradictory results. In our

dataset, barely any OTU could be differentiated down to species level, and a major proportion of OTUs could not be named at genus or higher levels either, indicating a considerable degree of hidden diversity in our dataset. For example, a high amount of big-sized dinoflagellate taxa could not be identified further, but might be identifiable using microscopy. However, for microscopy, too, it has to be noted that resolution at species level is mostly depending on taxonomic expertise, although resolution limits might not be as important for some easily identifiable taxa (Zingone *et al*., 2015). Moreover, the choice of target molecules and different regions influence the quality of the database alignment, since the genetic diversity of the target region might not be specific enough for identification at species level.

Connections to environmental parameters and community dynamics

The CCA explained 37.9% of total inertia, which indicates that one or several additional factors, not yet taken into account, influenced the community at Helgoland Roads significantly. For example, Callies and Scharfe (2015) found hydrodynamic transport in regard to currents to be the most influential forcing parameter during spring, which was not considered in this study. The interplay between freshwater introducing influence by river discharge and marine water could only be discussed in regard to the rapid changes in salinity. In addition to the high influence of hydrodynamic transport and weather conditions, internal influences due to species interactions and grazing by zooplankton need to be taken into account as well in the future.

We observed five distinct phases in the spring bloom of 2016. As different analyses like the hierarchical clustering and NMDS showed, the community in phase 3 was having rapid changes compared to other phases. In addition, three samples that were taken in between samples from phase 3 could not be assigned to any phase, since they were more diverse. This indicates that additional communities might undergo rapid changes and would not be visible with a lower sampling frequency. It is noteworthy that this timeframe coincides with the maximum of the total counts of diatoms at the LTER site.

The results regarding community composition showed that phases 1 and 5 were more similar to each other than the communities during phases 3 and 4. Comparing the beta diversity matrix with the environmental parameters, a significant correlation to temperature and salinity was shown. This result suggests that the contrasting environmental conditions like temperature differences did not inhibit the development of similar communities, which decreases the influence of temperature on community succession.

The most abundant genera were found in the OTUs with the greatest contributions by our SIMPER analyses. Since most abundant genera were available in our dataset during the

whole sampling period, we can assume that these influence the community the most. Especially *Phaeocystis* sp. and *Emiliania* sp. could be identified as important blooming small-sized eukaryotic microbes. It has to be noted that our study is the first study using Illumina sequencing in this temporal resolution at Helgoland Roads. Therefore, it is not possible to compare our findings with sequencing data from previous years. However, several campaigns and efforts have been made to sample certain taxa or neighbouring areas. In the following paragraphs, we try to compare these findings by several different molecular methods with our results regarding the different taxa and small-sized eukaryotes.

Diversity of nano- and picoplankton taxa

A considerable amount of new information about the spring bloom community on nano- and picoplankton composition was gained through this study, providing new insights into heterotrophic and possible parasitic components of the microbial loop communities.

Nano- and picoplankton taxa such as Syndiniales (Dino-Groups), *Emiliania* sp., *Phaeocystis* sp. and Choanoflagellida groups (*Stephanoecidae* Group D) were found in relatively high abundances and showed a distinctive blooming pattern during spring. Out of these taxa, only *Phaeocystis* sp. is counted at Helgoland Roads currently, while coccolithophorids like *Emiliania* and choanoflagellates cannot be identified on genus level. For other regions of the North Sea especially *Phaeocystis* and coccolithophorids are already known to be important compartments of the spring bloom community (Widdicombe *et al*., 2010). Despite their small cell size, *Phaeocystis* sp. are resistant against grazing by small-sized copepods due to their forming of gelatinous colonies and production of deterring chemicals, while different microzooplankton such as ciliates and heterotrophic dinoflagellates are known to feed on single *Phaeocystis* cells and on colonies (Hamm, 2000; Stelfox-Widdicombe *et al*., 2004; Schoemann *et al*., 2005). A shift from diatom blooms to *Phaeocystis*-dominated blooms therefore would influence the grazing success of the known copepods such as *Acartia* spp. and *Temora* spp. and change the whole food web dynamics at Helgoland.

If we look at other heterotroph small-sized eukaryotic microbes like the MAST groups, which are also not included in the LTER, we found a high amount of OTUs, of which several were present in all samples and clustered within the biggest cluster (cluster 4). This cluster included most OTUs that were present during the whole sampling period and represents a diverse background community. Accordingly, Logares *et al*. (2012) and Massana *et al*. (2014), who used data from several stations from European coasts, found the biggest contributions of different MAST groups in the pico- but also in the nanoplankton fraction.

Furthermore, cluster 22 stood out with mostly Hacrobia OTUs. In general, the OTUs in this cluster appear to play a role in early and late spring (phases 1 and 5), hinting that they might

be suppressed by blooming plankton fractions, such as other Hacrobia like *Phaeocystis* and *Emiliania*. The cryptophytes fromthis cluster coincide well with findings fromearlier studies (Metfies *et al*., 2010). Further analysis by Medlin *et al*. (2017) identified *Teleaulax*, *Plagioselmis* and *Geminigera* spp. as possible important cryptophytes during the spring bloom. In accordance with our results, these genera were abundant during the early and late phases of our spring bloom, but did not significantly contribute to the statistical similarities.

Diverse communities, such as represented by cluster 31, included taxa, which belonged to diatoms, heterotrophic flagellates or ciliate taxa. For example, choanoflagellates, as part of the heterotrophic nanoplankton, are a big contributor to carbon cycling in marine food webs, since they are grazing on bacteria and detritus but are themselves food for larger predators (King, 2005). The present fungi or fungi-like organisms act as decomposers of organic matter but can also be parasites of autotrophic primary producers and control their growth (Jobard *et al*., 2010). For Helgoland, it has been found that selective grazing by microzooplankton is important for phytoplankton spring bloom development and the occurrence of ciliates is dependent on specific preys (Löder *et al*., 2011). As these taxa cluster occurred during the early phases, where we observed the highest diatom abundances, a similar relationship can be suggested for our study.

Conclusion

In order to achieve new insights to the Helgoland Roads eukaryotic microbial community during spring, we analysed the sequence assemblage and identified main abiotic correlations to the community dynamics. We obtained several unexpected results, which should be addressed in future observations. Most prominently, we observed a low occurrence of diatoms in our molecular dataset, despite the high sampling frequency, which we expect to be mainly caused by methodological constraints. Instead, our assemblage was mainly dominated by dinoflagellate OTUs. We could identify several taxa that occur at Helgoland during the whole period. At the same time, a rapid phytoplankton succession was observed, with some taxa only making occasional appearances. In accordance with our aim, we could identify many small-sized eukaryotic microbes, which showed a distinctive blooming pattern such as *Emiliania* and *Phaeocystis*. Pico- and nanoplankton are part of a core community, vary in bloom timing and form community clusters. Taking into account the abiotic factors used in our analysis, temperature and salinity were the abiotic parameters with the biggest correlations to the microbial communities present during our sampling period. However, it needs to be mentioned that contrasting conditions in these parameters did not prevent similar communities to evolve. Also, there are still unknown variables, which

also influence the community structure that have not been taken into account. Since previous knowledge relies on microscopy, such as the known diatom spring bloom peak, which could not be identified in our dataset, there is a need to compare methods in more detail to overcome this issue and identify gaps and possibilities of synergy effects of the different datasets.

Data Archiving

Sequence data for this study have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession number PRJEB37135 (https://www.ebi.ac.uk/ena/data/view/PRJEB37135), using the data brokerage service of the German Federation for Biological Data (GFBio, Diepenbroek *et al.*, 2014), in compliance with the Minimal Information about any (X) Sequence (MIxS) standard (Yilmaz *et al.*, 2011).

Supplementary Data

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

Acknowledgements

We wish to thank the crew of the research vessel “Aade”, our technicians Kristine Carstens and Silvia Peters for collecting all LTER data and the MPI sampling team: David Probandt, Greta Giljan, Karen Krüger, Laura Zeugner, Lennart Kappelmann and Meghan Chafee. Additionally, we want to thank Karl Peter Rücknagel and Sabine Kühn for DNA extraction of the 0.2 µm filter samples as well as Swantje Rogge and Kerstin Oetjen for their support on the sequencing procedure. We would also like to thank Mirco Scharfe and Cédric Meunier for their support and suggestions on the manuscript.

Conflict of interest

The authors declare that there is no conflict of interest.

Funding

This work was supported by the Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung. We further acknowledge support by the Open Access Publication Funds of Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung.

References

Akaike,H. (1974) A new look at the statistical model identification. *IEEE Trans. Automat. Contr.*, **19**, 716–723.

Andersen, S. K., Kirkegaard, R. H., Karst, S. M. and Albertsen, M. (2018) ampvis2: an R package to analyse and visualise 16S rRNA amplicon data. *bioRxiv* [Preprint], available at https://doi.org/10.1101/299537.

Auguie, B. and Antonov, A. (2017) gridExtra: Miscellaneous Functions for grid Graphics. R package version 2.3, available at https://CRAN.R-project.org/package=gridExtra.

Azam, F., Fenchel, T., Field, J. G., Gray, J. S., Meyer-Reil, L. A. and Thingstad, F. (1983) The ecological role of water-column microbes in the sea. *Mar. Ecol. Prog. Ser.*, **10**, 257–263.

Bazin, P., Jouenne, F., Friedl, T., Deton-Cabanillas, A. F., Le Roy, B. and Veron, 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.

Bokulich, N. A., Subramanian, S., Faith, J. J., Gevers, D., Gordon, J. I., Knight, R., Mills, D. A. and Caporaso, J. G. (2013) Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. *Nat. Methods*, **10**, 57–59.

Bolger, A.M., Lohse, M. and Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 2114–2120.

Bougeard, S. and Dray, S. (2018) Supervised multiblock analysis in R with the ade4 package. *J. Stat. Softw.*, **86**, 1–17.

Callies, U. and Scharfe, M. (2015) Mean spring conditions at Helgoland roads, North Sea: graphical modeling of the influence of hydroclimatic forcing and Elbe River discharge. *J. Sea Res.*, **101**, 1–11.

Caron, D. A. and Hu, S. K. (2019) Are we overestimating protistan diversity in nature? *Trends Microbiol.*, **27**, 197–205.

Carstensen, J., Klais, R. and Cloern, J. E. (2015) Phytoplankton blooms in estuarine and coastal waters: seasonal patterns and key species. *Estuar. Coast. Shelf Sci.*, **162**, 98–109.

Chafee, M., Fernàndez-Guerra, A., Buttigieg, P. L., Gerdts, G., Eren, A. M., Teeling, H. and Amann, R. I. (2018) Recurrent patterns of microdiversity in a temperate coastal marine environment. *ISME J.*, **12**, 237–252.

Connolly, J. A., Oliver, M. J., Beaulieu, J. M., Knight, C. A., Tomanek, L. andMoline, M. A. (2008) Correlated evolution of genome size and cell volume in diatoms (Bacillariophyceae). *J. Phycol.*, **44**, 124–131.

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

Diepenbroek, M., Glöckner, F., Grobe, P., Güntsch, A., Huber, R., König-Ries, B., Kostadinov, I., Nieschulze, J. *et al.* (2014, 2014) Towards an integrated biodiversity and ecological research data management and archiving platform: The German Federation for the curation of biological data (GFBio). In Plödereder, E., Grunske, L., Schneider, E. and Ull, D. (eds.), *Informatik 2014 – Big Data Komplexität meistern. GI-Edition: Lecture Notes in Informatics (LNI) – Proceedings*, Vol. **232**, Köllen Verlag, Bonn, pp. 1711–1724.

Dray, S. and Dufour, A. B. (2007) The ade4 package: implementing the duality diagram for ecologists. *J. Stat. Softw.*, **22**, 1–20.

DWD Climate Data Center (CDC) (2019) *Daily Station Observations of Sunshine Duration in Hours for GERMANY*, version v19.3, last accessed 24 August 2019.

Ebenezer, V., Medlin, L. K. and Ki, J. S. (2012) Molecular detection, quantification, and diversity evaluation of microalgae. *Marine Biotechnol.*, **14**, 129–142.

Fadeev, E., Salter, I., Schourup-Kristensen, V., Nöthig, E.-M., Metfies, K., Engel, A., Piontek, J., Boetius, A. *et al.* (2018) Microbial communities in the east and west Fram Strait during sea ice melting season. *Front. Mar. Sci.*, **5**, 1–21.

Field, C. B., Behrenfeld, M. J., Randerson, J. T. and Falkowski, P. (1998) Primary production of the biosphere: integrating terrestrial and oceanic components. *Science (80-.)*, **281**, 237–240.

Galili, T. (2015) Dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering. *Bioinformatics*, **31**, 3718–3720.

Gescher, C., Metfies, K., Frickenhaus, S., Knefelkamp, B., Wiltshire, K.H. and Medlin, L. K. (2008) Feasibility of assessing the community composition of prasinophytes at the Helgoland roads sampling site with a DNA microarray. *Appl. Environ. Microbiol.*, **74**, 5305–5316.

Giner, C. R., Balagué, V., Krabberød, A. K., Ferrera, I., Reñé, A., Garcés, E., Gasol, J. M., Logares, R. *et al.* (2019) Quantifying long-term recurrence in planktonic microbial eukaryotes. *Mol. Ecol.*, **28**, 923–935.

Gómez, F. (2012) A quantitative review of the lifestyle, habitat and trophic diversity of dinoflagellates (Dinoflagellata, Alveolata). *Syst. Biodivers.*, **10**, 267–275.

Grasshoff, K. (1976) *Methods of Seawater Analysis*, Verlag Chemie, Weinheim and New York.

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., Boutte, C. *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.

Hamm, C. E. (2000) Architecture, ecology and biogeochemistry of Phaeocystis colonies. *J. Sea Res.*, **43**, 307–315.

Harris, R. (2010) The L4 time-series: the first 20 years. *J. Plankton Res.*, **32**, 577–583.

Hickel, W., Mangelsdorf , P. and Berg, J. (1993) The human impact in the German bight: eutrophication during three decades (1962-1991). *Helgoländer Meeresuntersuchungen*, **47**, 243–263.

Hong, H.H., Lee, H.G., Jo, J., Kim, H.M., Kim, S.M., Park, J. Y., Jeon, C. B., Kang, H. S. *et al.* (2016) The exceptionally large genome of the harmful red tide dinoflagellate cochlodinium polykrikoides Margalef (Dinophyceae): determination by flow cytometry. *Algae*, **31**, 373–378.

Hoppenrath, M. (2004) A revised checklist of planktonic diatoms and dinoflagellates from Helgoland (North Sea, German bight). *Helgol. Mar. Res.*, **58**, 243–251.

Hu, S. K., Campbell, V., Connell, P., Gellene, 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.*, **92**, fiw050.

Hu, S. K., Connell, P. E.,Mesrop, L. Y. and Caron, D. A. (2018) A hard Day's night: Diel shifts in microbial eukaryotic activity in the North Pacific subtropical gyre. *Front. Mar. Sci.*, **5**, 1–17.

Ishida, K., Green, B. R. and Cavalier-Smith, T. (1999) Diversification of a Chimaeric algal group, the Chlorarachniophytes: phylogeny of nuclear and Nucleomorph small-subunit rRNA genes. *Mol. Biol. Evol.*, **16**, 321–331.

Jobard, M., Rasconi, S. and Sime-Ngando, T. (2010) Diversity and functions of microscopic fungi: a missing component in pelagic food webs. *Aquat. Sci.*, **72**, 255–268.

King, N. (2005) Choanoflagellates. *Curr. Biol.*, **15**, 113–114.

Knefelkamp, B. (2009) An Investigation of the Eukaryotic Picoplankton Community in the German North Sea by Different Methods, University Bremen, Bremen.

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.

Kubiszyn, A.M., Piwosz, K., Wiktor, J.M. and Wiktor, J.M. (2014) The effect of inter-annual Atlantic water inflow variability on the planktonic protist community structure in the West Spitsbergen waters during the summer. *J. Plankton Res.*, **36**, 1190–1203.

LaJeunesse, T. C., Lambert, G., Andersen, R. A., Coffroth, M. A. and Galbraith, D. W. (2005) Symbiodinium (Pyrrhophyta) genome sizes (DNA content) are smallest among Dinoflagellates. *J. Phycol.*, **41**, 880–886.

Lewandowska, A. M., Striebel, M., Feudel, U., Hillebrand, H. And Sommer, U. (2015) The importance of phytoplankton trait variability in spring bloom formation. *ICES J. Mar. Sci.*, **72**, 1908–1915.

Löder, M. G. J., Meunier, C., Wiltshire, K.H., Boersma, M. and Aberle, N. (2011) The role of ciliates, heterotrophic dinoflagellates and copepods in structuring spring plankton communities at Helgoland roads, North Sea. *Mar. Biol.*, **158**, 1551–1580.

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.

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.*, **6**, 1823–1833.

Lucas, J., Koester, I., Wichels, A., Niggemann, J., Dittmar, T., Callies, U., Wiltshire, K. H. and Gerdts, G. (2016) Short-term dynamics of north sea bacterioplankton-dissolved organic matter coherence on molecular level. *Front. Microbiol.*, **7**, 1–14.

Lucas, J.,Wichels, A., Teeling, H., Chafee, M., Scharfe,M. and Gerdts, G. (2015) Annual dynamics of North Sea bacterioplankton: seasonal variability superimposes short-term variation. *FEMS Microbiol. Ecol.*,**91**.

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

Mahé, F., Rognes, T., Quince, C., DE Vargas, C. and Dunthorn, M. (2015) Swarm v2: highly-scalable and high-resolution amplicon clustering. *Peer J.*, **3**, e1420.

Marquardt, M., Vader, A., Stübner, E. I., Reigstad, M. and Gabrielsen, T. M. (2016) Strong seasonality of marine microbial eukaryotes in a high-Arctic fjord (Isfjorden, West Spitsbergen). *Appl. Environ. Microbiol.*, **82**, 1868–1880.

Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.*, **17**, 10–12.

Massana, R. (2011) Eukaryotic picoplankton in surface oceans. *Annu. Rev. Microbiol.*, **65**, 91–110.

Massana, R., Del Campo, J., Sieracki, M. E., Audic, S. and Logares, R. (2014) Exploring the uncultured microeukaryote majority in the oceans: re-evaluation of ribogroups within stramenopiles. *ISME J.*, **8**, 854–866.

Massana, R., Gobet, A., Audic, S., Bass, D., Bittner, L., Boutte, C., Chambouvet, A., Christen, R. *et al.* (2015) Marine protist diversity in European coastal waters and sediments as revealed by high-throughput sequencing. *Environ. Microbiol.*, **17**, 4035–4049.

McQuatters-Gollop, A., Edwards, M., Helaouët, P., Johns, D. G., Owens, N. J. P., Raitsos, D. E., Schroeder, D., Skinner, J. *et al.* (2015) The continuous plankton recorder survey: how can long-term phytoplankton datasets contribute to the assessment of good environmental status? *Estuar. Coast. Shelf Sci.*, **162**, 88–97.

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.

Medlin, L. K., Piwosz, K. and Metfies, K. (2017) Uncovering hidden biodiversity in the cryptophyta: clone library studies at the Helgoland time series site in the southern German bight identifies the cryptophycean clade potentially responsible for the majority of its genetic diversity during the spring. *Life Environ.*, **67**, 27–32.

Metfies, K., Gescher, C., Frickenhaus, S., Niestroy, R., Wichels, A., Gerdts, G., Knefelkamp, B., Wiltshire, K. *et al.* (2010) Contribution of the class cryptophyceae to phytoplankton structure in the German bight. *J. Phycol.*, **46**, 1152–1160.

Mieruch, S., Freund, J. A., Feudel, U., Boersma, M., Janisch, S. and Wiltshire, K. H. (2010) A new method of describing phytoplankton blooms: examples from Helgoland roads. *J. Mar. Syst.*, **79**, 36–43.

Neuwirth, E. (2014) Color brewer palettes.

Nixon, S.W., Fulweiler, R.W., Buckley, B.A., Granger, S. L., Nowicki, B. L. and Henry, K.M. (2009) The impact of changing climate on phenology, productivity, and benthic–pelagic coupling in Narragansett Bay. *Estuar. Coast. Shelf Sci.*, **82**, 1–18.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., Minchin, P. R., O'hara, R. B. *et al.* (2019) *Vegan: Community Ecology Package*, R package version 2.5-6, available at https://CRAN.R-project.org/package=vegan.

Pruesse, E., Quast, C., Knittel, K., Fuchs, B. M., Ludwig, W., Peplies, J. and Glöckner, F. O. (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res.*, **35**, 7188–7196.

R Core Team (2018) *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org/

Racault, M.-F., Le Quéré, C., Buitenhuis, E., Sathyendranath, S. and Platt, T. (2012) Phytoplankton phenology in the global ocean. *Ecol. Indic.*, **14**, 152–163.

Reid, P. C., Colebrook, J.M., Matthews, J. B. L. and Aiken, J. (2003) The continuous plankton recorder: concepts and history, from plankton indicator to undulating recorders. *Prog. Oceanogr.*, **58**, 117–173.

Reid, P. C., Lancelot, C., Gieskes, W.W. C., Hagmeier, E. and Weichart, G. (1990) Phytoplankton of the North Sea and its dynamics: a review. *Netherlands J. Sea Res.*, **26**, 295–331.

Riebesell, U. (1993) Aggregation of Phaeocystis during phytoplankton spring blooms in the southern North Sea. *Mar. Ecol. Prog. Ser.*, **96**, 281–289.

Rognes, T., Flouri, T., Nichols, B., Quince, C. and Mahé, F. (2016) VSEARCH: a versatile open source tool for metagenomics. *PeerJ*, **4**, e2584.

Ryota Suzuki, A. and Shimodaira, H. (2015) *pvclust: Hierarchical Clustering with P-Values via Multiscale Bootstrap Resampling, 2.0–0*, available at https://CRAN.R-project.org/package=pvclust.

Sapp, M., Wichels, A., Wiltshire, K. H. and Gerdts, G. (2007) Bacterial community dynamics during the winter-spring transition in the North Sea. *FEMS Microbiol. Ecol.*, **59**, 622–637.

Scharfe, M. and Wiltshire, K. H. (2019) Modeling of intra-annual abundance distributions: constancy and variation in the phenology of marine phytoplankton species over five decades at Helgoland roads (North Sea). *Ecol. Model.*, **404**, 46–60.

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

Schlüter, M. H., Kraberg, A. and Wiltshire, K. H. (2012) Long-term changes in the seasonality of selected diatoms related to grazers and environmental conditions. *J. Sea Res.*, **67**, 91–97.

Schoemann, V., Becquevort, S., Stefels, J., Rousseau, V. and Lancelot, C. (2005) Phaeocystis blooms in the global ocean and their controlling mechanisms: a review. *J. Sea Res.*, **53**, 43–66.

Sherr, E. B. and Sherr, B. F. (2002) Significance of predation by protists in aquatic microbial food webs. *Antonie Van Leeuwenhoek*, **81**, 293–308.

Simon, N., Cras, A.-L., Foulon, E. and Lemée, R. (2009) Diversity and evolution of marine phytoplankton. *C. R. Biol.*, **332**, 159–170.

Sournia, A., Chrétiennot-Dinet, M.-J. and Ricard, M. (1991) Marine phytoplankton: how many species in the world ocean? *J. Plankton Res.*, **13**, 1093–1099.

Steinberger, A. J. (2018) *asteinberger9/seq_scripts: Release v1 (Version v1)*, available at https://zenodo.org/record/1458243.

Stelfox-Widdicombe, C. E., Archer, S. D., Burkill, P. H. and Stefels, J. (2004) Microzooplankton grazing in Phaeocystis and diatom-dominated waters in the southern North Sea in spring. *J. Sea Res.*, **51**, 37–51.

Stern, R., Kraberg, A., Bresnan, E., Kooistra, W. H. C. F., Lovejoy, C., Montresor, M., Morán, X. A. G., Not, F. *et al.* (2018) Molecular analyses of protists in long-term observation programmes—current status and future perspectives. *J. Plankton Res.*, **40**, 519–536.

Suzuki, R. and Shimodaira, H. (2006) Pvclust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics*, **22**, 1540–1542.

Tange, O. (2011) GNU parallel: the command-line power tool login. *USENIX Mag.*, **36**, 42–47.

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

Teeling, H., Fuchs, B. M., Bennke, C. M., Krüger, K., Chafee, M., Kappelmann, L., Reintjes, G., Waldmann, J. *et al.* (2016) Recurring patterns in bacterioplankton dynamics during coastal spring algae blooms. *Elife*, **5**, 1–31.

Tragin, M., Zingone, A. and Vaulot, D. (2018) Comparison of coastal phytoplankton composition estimated from the V4 and V9 regions of the 18S rRNA gene with a focus on photosynthetic groups and especially Chlorophyta. *Environ. Microbiol.*, **20**, 506–520.

DE Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., Lara, E., Velayoudon, D. *et al.* (2015) Eukaryotic plankton diversity in the sunlit ocean. *Science (80-).*, **348**, 1261605–1261605.

Warnes, A. G. R., Bolker, B., Bonebakker, L., Huber, W., Liaw, A., Lumley, T., Magnusson, A., Moeller, S. *et al.* (2019) *gplots: Various R Programming Tools for Plotting Data*, version 3.0.1.1, available at https://CRAN.R-project.org/package=gplots.

Whittaker, R. H. (1960) Vegetation of the Siskiyou mountains, Oregon and California. *Ecol. Monogr.*, **30**, 279–338.

Wickham, H. (2016) *Ggplot2: Elegant Graphics for Data Analysis*, 2nd edn, Springer International Publishing, New York.

Widdicombe, C. E., Eloire, D., Harbour, D., Harris, R. P. and Somerfield, P. J. (2010) Long-term phytoplankton community dynamics in the western English Channel. *J. Plankton Res.*, **32**, 643–655.

Wiltshire, K. H., Boersma, M., Carstens, K., Kraberg, A. C., Peters, S. and Scharfe, M. (2015) Control of phytoplankton in a shelf sea: determination of the main drivers based on the Helgoland roads time series. *J. Sea Res.*, **105**, 42–52.

Wiltshire, K. H. and Dürselen, C. D. (2004) Revision and quality analyses of the Helgoland Reede long-term phytoplankton data archive. *Helgol. Mar. Res.*, **58**, 252–268.

Wiltshire, K. H., Kraberg, A., Bartsch, I., Boersma, M., Franke, H. D., Freund, J., Gebühr, C., Gerdts, G. *et al.* (2010) Helgoland roads, north sea: 45 years of change. *Estuaries Coast*, **33**, 295–310.

Wiltshire, K. H., Malzahn, A. M., Wirtz, K., Greve, W., Janisch, S., Mangelsdorf, P., Manly, B. F. J. and Boersma, M. (2008) Resilience of North Sea phytoplankton spring bloom

dynamics: an analysis of long-term data at Helgoland roads. *Limnol. Oceanogr.*, **53**, 1294–1302.

Wiltshire, K. H. and Manly, B. F. J. (2004) The warming trend at Helgoland roads, North Sea: phytoplankton response. *Helgol. Mar. Res.*, **58**, 269–273.

Wollschläger, J., Wiltshire, K. H., Petersen, W. and Metfies, K. (2015) Analysis of phytoplankton distribution and community structure in the German bight with respect to the different size classes. *J. Sea Res.*, **99**, 83–96.

Yilmaz, P., Kottmann, R., Field, D., Knight, R., Cole, J. R., Amaral-Zettler, L., Gilbert, J. A., Glöckner, F. O. *et al.* (2011) Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIxS) specifications. *Nat. Biotechnol.*, **29**, 415–420.

Zapata, M., Rodríguez, F. and Garrido, J. (2000) Separation of chlorophylls and carotenoids from marine phytoplankton: a new HPLC method using a reversed phase C8 column and pyridine-containing mobile phases. *Mar. Ecol. Prog. Ser.*, **195**, 29–45.

Zhang, J., Kobert, K., Flouri, T. and Stamatakis, A. (2014) PEAR: a fast and accurate Illumina paired-end reAd mergeR. *Bioinformatics*, **30**, 614–620.

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

5 Publication II

Publication II, namely, “Host–parasitoid associations in marine planktonic time series: Can metabarcoding help reveal them?”, focused on the identification of planktonic parasitoids at the Helgoland Roads LTER station. Marine planktonic parasitoids play an important role in plankton community dynamics, but they remain understudied, and their occurrence in the marine food web has mostly been ignored. By creating a three-year 18S metabarcoding dataset, **Publication II** aimed to determine the extent of the eukaryotic parasitoid presence in the plankton community (aims 1a and c of this thesis) and its potential links to environmental conditions (aim 1b of this thesis). Parasitoids were observed in high abundances and diversity, with a wide range of potential hosts. Some parasitoids were always present while others appeared only during certain seasons and in individual patterns. We also tried to detect host–parasitoid dynamics (aim 1c of this thesis); however, only known host–parasitoid associations could be identified. The high dynamics and variability of the parasitoids impeded the identification of potential hosts and new associations, so their impact on the different food web components is still unclear, which emphasizes the need for future investigations.

Host-parasitoid associations in marine planktonic time series: Can metabarcoding help reveal them?

Laura Käse[1*], Katja Metfies[2,3], Stefan Neuhaus[2], Maarten Boersma[1,4], Karen Helen Wiltshire[1,5], Alexandra Claudia Kraberg[2]

[1]Alfred-Wegener-Institut, Helmholtz-Zentrum für Polar- und Meeresforschung, Biologische Anstalt Helgoland, Helgoland, Schleswig-Holstein, Germany

[2]Alfred-Wegener-Institut, Helmholtz-Zentrum für Polar- und Meeresforschung, Bremerhaven, Bremen, Germany

[3]Helmholtz-Institut für Funktionelle Marine Biodiversität, Oldenburg, Germany

[4]University of Bremen, Bremen, Bremen, Germany

[5]Alfred-Wegener-Institut, Helmholtz-Zentrum für Polar- und Meeresforschung, Wadden Sea Station, List auf Sylt, Schleswig-Holstein, Germany

* laura.kaese@awi.de

Editor: Alberto Amato, IRIG-CEA Grenoble, FRANCE

Received: June 11, 2020, Accepted: December 16, 2020, Published: January 7, 2021

Published in PLoS ONE:

Citation: Käse L, Metfies K, Neuhaus S, Boersma M, Wiltshire KH, Kraberg AC (2021) Host-parasitoid associations in marine planktonic time series: Can metabarcoding help reveal them? PLoS ONE 16(1): e0244817. https://doi.org/10.1371/journal.pone.0244817

Data Availability Statement: Sequence data for this study have been deposited in the European Nucleotide Archive under accession number PRJEB37135 (https://www.ebi.ac.uk/ena/data/view/PRJEB37135). Additionally the full OTU table (280 samples with 59.284 OTUs) was archived in PANGAEA (DOI: 10.1594/PANGAEA.921026). All other relevant data are within the manuscript and its Supporting Information files.

Funding: This work was supported by institutional funds of the Alfred-Wegener-Institut Helmholtz Zentrum für Polar- und Meeresforschung, Germany. LK furthermore

acknowledges support by the Open Access Publication Funds of Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung.

Competing interests: The authors have declared that no competing interests exist.

All Supporting Information files (5 figures, 6 tables and 1 text file) can be found online (https://doi.org/10.1371/journal.pone.0244817), supporting figures and text can also be found in the appendix of this thesis.

Abstract

In this study, we created a dataset of a continuous three-year 18S metabarcoding survey to identify eukaryotic parasitoids, and potential connections to hosts at the Long-Term Ecological Research station Helgoland Roads. The importance of parasites and parasitoids for food web dynamics has previously been recognized mostly in terrestrial and freshwater systems, while marine planktonic parasitoids have been understudied in comparison to those. Therefore, the occurrence and role of parasites and parasitoids remains mostly unconsidered in the marine environment. We observed high abundances and diversity of parasitoid operational taxonomic units in our dataset all year round. While some parasitoid groups were present throughout the year and merely fluctuated in abundances, we also detected a succession of parasitoid groups with peaks of individual species only during certain seasons. Using co-occurrence and patterns of seasonal occurrence, we were able to identify known host-parasitoid dynamics, however identification of new potential host-parasitoid interactions was not possible due to their high dynamics and variability in the dataset.

Introduction

Parasitism is a common lifestyle for a wide variety of species, including planktonic ones. It is one of the multiple biotic factors that can influence food web structure. For example, there can be changes in food chain length, connectivity, and stability [1–3]. Such effects have previously been shown for planktonic freshwater systems [4, 5] but little information is available for the marine realm especially with regards to eukaryotic parasitoids [6, 7]. Parasitoids, those organisms that ultimately kill their hosts, in the marine environment range from viruses and bacteria to several protist taxa. Whereas some progress has been made in recent years on bacterial and viral infections [8–14], studies on eukaryotic parasites and parasitoids have focused mainly on single host-parasitoid/parasite systems (in the following only named as host-parasitoid systems) or species groups, in short-term microscopy-based

projects [15–19]. Currently, long-term (multi-year) investigations are largely missing. These kind of investigations could yield important information on the dynamics of the interactions.

While it is known that infection by a parasite affects the fitness of the host and most parasites are transferred through several different hosts, parasitoids often complete their life cycle in a single host and kill the host in the process [20, 21]. Since protist parasites are often classified as parasitoids [22–24] and a definite distinction between parasites and parasitoids is difficult for some planktonic taxa, we will only use the term parasitoid in the following manuscript to describe all taxa that have been found to be related to the parasitism strategy. Parasitoid microbes can be drivers of phytoplankton bloom dynamics, play important roles in host population regulation [20, 21] and can influence phytoplankton succession due to their selectiveness of host species [25]. The infection by parasitoids can even cause a phytoplankton bloom to collapse [26–28]. For example, Tillmann *et al*. [25] indicated that parasitic infections of phytoplankton compete with zooplankton in marine food webs, as algal cells are killed and consequently no longer available to higher trophic levels such as mesozooplankton. Indeed, even classic Lotka-Volterra dynamics, defined as periodic and alternating fluctuations of predator and prey, have been observed in host-parasitoid relationships [29], and peaks in abundance of a host are followed by peaks in abundance of a parasitoid [30].

Even though the examples cited above may suggest otherwise, our knowledge on the role of parasitoids in marine ecosystems is still incomplete [20]. This paucity of information is strongly related to insufficient monitoring capacity and methodological constraints [20], and even the identification of organisms as parasitoids and their subsequent taxonomic determination is difficult and needs improvement. Otherwise, it is not possible to make some inference about the impact of parasitoids on marine ecosystems.

Considerable diversity exists in marine parasitoid protists and an equally diverse range of known hosts, including marine algae, nematodes, crustaceans and fish has been described [20]. So far, several eukaryotic taxa are known to include parasitoid classes: Dinoflagellata, Stramenopiles, Cercozoa, Ciliophora, Apicomplexa, Mesomycetozoa, Metazoa, Lobosa, Perkinsida and true Fungi. The hosts of many of those parasitoid protists are protists themselves. Syndiniales, for example, a class of dinoflagellates, is composed exclusively of parasitoid species, and occur globally, including the Arctic and Antarctic [31] and may, as a result, be rather abundant in metabarcoding datasets [32–34]. They can infect several hosts, ranging from dinoflagellates and ciliates to copepods, crabs and fish. For example they have been found to be lethal to the eggs or newly hatched fish larvae [35]. Another example of a class of mostly protistan parasitoids are the heterokont oomycetes. These

belong to the kingdom of Stramenopiles [36, 37], and infect a wide range of hosts such as brown algae, diatoms, crustaceans and fish in marine environments [21]. While some parasitoids are host-specific, others can infect different (plurivorous) species, and in return, hosts can be infected by several parasitoids simultaneously [37].

As one of the longest running long-term observatories, Helgoland Roads Long-Term Ecological Research site (LTER) provides abiotic and biotic data at a very high temporal resolution, including phytoplankton, temperature, salinity and inorganic nutrients [38, 39]. During the course of this long-term observation programme at Helgoland, several diatom-infecting parasitoids were already detected, using light microscopic observation. These include Cryomonadida such as the nanoflagellate *Cryothecomonas aestivalis*, which is known to infect the diatom *Guinardia delicatula* [19, 27], the Oomycete *Lagenisma coscinodisci*, which is known to infect the diatom *Coscinodiscus* sp. [40, 41] and also two recently described oomycete parasitoid species: *Miracula helgolandica* in the host *Pseudo-nitzschia pungens* [24, 42] and *Olpidiopsis drebesii* in *Rhizosolenia imbricata* [42]. *Cryothecomonas longipes*, which can infect a broad spectrum of diatoms including *Thalassiosira rotula* [18], and several *Pirsonia sp*. with possible hosts like *Rhizosolenia sp*. [15] were detected in the North Sea but not yet at Helgoland.

As indicated above, most of the evidence on host-parasitoid interactions at Helgoland was derived from microscopic methods. However, many of the organisms involved are small and without conspicuous characteristics. They can -if at all- only be identified as flagellates in the pico- and nanoplankton fractions in their free living states or by spotting inside of infected host cells [37]. Therefore, there is great scope for improvement. Next generation sequencing (NGS) and other molecular methods have great potential to close this gap, but we do not know enough yet, to be able to implement these techniques in a long-terms series approach. Open questions are, for example, whether relevant temporal dynamics in a host-parasitoid system can be observed if the parasitoid changes from free living to parasitic stages. Furthermore, it also remains to be seen whether host-parasitoid dynamic behaviour follows the Lotka-Volterra type dynamics in a complex ecological context, with predators and competitors also present. The fact that several host-parasitoid systems have already been identified for Helgoland offers us the unique opportunity to test these open questions. It allows us to investigate the potential benefits and drawbacks of molecular methods in this context.

It was the aim of this study to create a high resolution and unique 18S metabarcoding dataset of continuous, high frequency sampling of three years duration (1) to identify the extent of planktonic eukaryotic parasitoid occurrence within the community at Helgoland

Roads throughout the year, and potential links to environmental conditions. Furthermore, we want (2) to assess if it is possible to detect known host-parasitoid systems, which have been described by conventional microscope analysis, and their dynamics using the sequencing dataset. By using the knowledge gleaned from the dynamics analysis of (2), we aim to (3) examine if potential host-parasitoid systems, that are not known at Helgoland but elsewhere, can be detected with these data based upon identification of alternating cyclical dynamics, plus if dynamical behaviour of host-parasitoid pairs allows for the identification of thus far unknown host-parasitoid associations.

Materials and methods

Study site and sampling

We took water surface samples from the Helgoland Roads LTER sampling site. The sampling site (54˚11.03' N, 7˚54.00'E) is situated between the main island of Helgoland and the dune island [38]. Secchi depth and temperature were measured directly. Other parameters include salinity, nutrients such as silicate, phosphate, inorganic nitrogen and chlorophyll, which were measured in the laboratory according to the LTER sampling protocol [38, 43, 44], for nutrients [45]. Daily observations of sunshine duration in hours were downloaded from the Deutscher Wetterdienst, Climate Data Centre [46]. Seasons were defined as follows: Spring = March to May, Summer = June to August, Autumn = September to November, Winter = December to February.

In total, three different sampling phases from the same station were combined to build a comprehensive dataset of over 3 years. In short, the first sampling phase was conducted from March 2016 to May 2016 (work-daily sampling) [47]. The second phase included samples from June to October 2016 (in total 6 samples, irregular sampling) [48]. The third phase was conducted from December 2016 until March 2019, where samples were taken twice a week. In the period between May to July 2018 we intensified sampling by increasing the frequency to three samples per week (see S1 Table for further information on the samples belonging to each sampling phase).

For sequencing, we filtered 1 L of the water sample. For the sampling phase 1, a sequential filtration was used as part of another sampling program for bacterial long-term monitoring [49, 50]. The sample was filtered through 10 µm polycarbonate filters, 3 µm PC filters and 0.2 µM polyvinylidene fluoride filters (Millipore, Schwalbach, Germany) according to the protocol by Teeling *et al.* [49]. Samples from sampling phase 2 and 3 were filtered with 0.45 µm nylon filters (Whatman, 47 mm). Following filtration, all filters were immediately frozen at -20˚C. It needs to be mentioned that the different pore sizes of the sampling

phases do not influence the detection of the eukaryotic picoplankton, due to their general size being bigger than 0.45 µm.

DNA-extraction

We used the Macherey-Nagel NucleoSpin® Plant II Kit for DNA extraction from the 10 µm, 3 µm of sampling phase 1 and all 0.45 µm filters from sampling phase 2 and 3, before the extracts were stored at -20˚C. DNA extraction from 0.2 µm filters from sampling phase 1 was conducted as described previously by Sapp *et al.* [51]. In short, cells were lysed with lysozyme and sodium dodecyl sulfate, a phenol/chloroform/isoamyl alcohol solution was used for DNA extraction with isopropanol used in the precipitation step. Here the DNA was eluted in sterile water. Then we pooled the separate DNA extracts from the sequentially filtered samples to obtain one sample per sampling date. The nucleic acid content of all samples was measured with a Quantus Fluorometer using the QuantiFluor® dsDNA System (Promega, USA).

MiSeq™ Illumina sequencing and data processing

We used the Nextera XT DNA Library Preparation protocol (Illumina, USA) to prepare the DNA isolates for the MiSeq™ Illumina sequencing. We identified a fragment of the V4 region of the 18S rDNA using the following primer set: 528iF (GCG GTA ATT CCA GCT CCA A) and 964iR (AC TTT CGT TCT TGA TYR R) [52]. For polymerase chain reactions (PCRs) KAPA HiFi HotStartReadyMix (Kapa Biosystems, Inc., USA) was used to avoid contamination. Afterwards, we confirmed the success of this amplicon PCR by using 2 µL of the PCR product for gel electrophoresis. 5 additional cycles were added to the original PCR program, if an increase of template (up to 5 µL) was not sufficient. About 43 million 2x300 bp paired-end sequences were produced using an Illumina MiSeq™ sequencer (Illumina, USA).

We then used our in-house developed pipeline for bioinformatic processing of the samples as described below (for more information see S1 File and https://github.com/PyoneerO/qzip).

The low-quality 3’-ends of the reads were trimmed by *Trimmomatic* (version 0.38) [53] and the paired-ends were merged by *VSEARCH* (version 2.3.0) [54]. *Cutadapt* (version 1.19) [55] was used to adjust the sequence orientation and to remove the forward and reverse primer matching sequence segments. Sequences were only kept if both primer matching segments could be detected. The remaining sequences were filtered by *VSEARCH* and sequences were discarded, i) if they were shorter than 300 bp or longer than 550 bp, ii) if

they carried any ambiguity or iii) if the expected base error (sum of all base error probabilities) of a sequence was above 0.25.

Chimeric sequences were sample-wise predicted by *VSEARCH* in *de novo* mode with default settings and removed from the sample files. Only samples with at least 10000 sequences after filtering were considered for further analyses.

The remaining 21 million sequences were clustered into operational taxonomic units (OTUs) by the tool s*warm* (version 2.2.2) [56, 57] with default settings. For each OTU the most abundant amplicon was selected as representative and taxonomically annotated with the default classifier implemented in *mothur* (version 1.38.1) [58]. As reference the *Protist Ribosomal Reference database* (PR2), version 4.11.1 [59], was chosen and the minimum confidence cut-off for annotation was set to a value of 80. The sequence data is available in the European Nucleotide Archive (ENA) at the European Bioinformatics Institute (EMBL-EBI) under accession number PRJEB37135 (https://www.ebi.ac.uk/ena/data/view/PRJEB37135), using the data brokerage service of the German Federation for Biological Data (GFBio) [60], in compliance with the Minimal Information about any (X) Sequence (MIxS) standard [61].

Data analysis and statistics

We reviewed the entire dataset of all 59,284 OTUs (in total 20,476,979 reads) for parasitoid taxa. For this we used information of literature focusing on known parasitoids in the North Sea and of the Tara Oceans Database W3 from the Companion Website of the article of de Vargas *et al.* [22]. Afterwards a threshold of 0.001% of total reads was applied to the full dataset. Hereby all OTUs remained, which had a total read count of 205 or higher, resulting in a limited dataset of 2790 OTUs. Out of this dataset, parasitoids that are known to be parasitizing plankton were extracted to get an overview of present parasitoids. Host-parasitoid relationships were identified by comparing occurrences of several parasitoids with potential hosts as described in the literature. Here, we defined peaks as local maxima during a certain period. The relative abundance needed to be at least 10% or more of the maximum relative abundance of the respective OTU or group. For diatom hosts, the word bloom was used, if various peaks could be identified in several consecutive samples or if high abundances above 10% were reached. Our goal was to find the relationships in the first place rather than describing the dynamics as a model. Also distinct time lags between host and parasitoid occurrence are either unknown for known relationships or can not be assumed to be correct for new potential relationships. Therefore, we focused on identifying two cases: 1. Alternating associations of potential hosts and parasitoid were considered to indicate typical Lotka-Volterra dynamics of the host-parasitoid system and time lags of up

to several days as they have been identified by microscopic analysis in the past. 2. Simultaneous appearance of potential host and parasitoid were expected to indicate a current infection.

For investigation of new host-parasitoid relationships two different approaches were tested. Parasitoid occurrences were compared with different hosts as they are known from the literature from other areas as well as closely related species. The limited dataset (2790 OTUs) was used to identify potential relationships that were found to be relevant based on the two cases of identification as described above. By using the known sequences, parasitoid OTUs and their possible hosts were verified with the Basic Local Alignment Search Tool (BLAST), when specific host-parasitoid systems were investigated.

A constrained ordination model based on the OTU table (based on relative abundances) and available environmental parameters was conducted in R, version 4.0.0 [62], using the vegan package [63]. Seasons and total parasitoid occurrence (as relative abundance) were included as additional parameters. Single parameters were combined with an analysis of variance- like permutation test for Canonical Correspondence Analysis (CCA) to assess the significance of the constraining factors [63]. The variables were chosen by their significance ($p < 0.05$). If several variables were given as significant in the same step, the variable with the lowest Akaike information criterion (AIC) value was chosen to minimize the information loss [64]. Environmental parameters that were included in the model development were temperature, salinity, Secchi depth, tide and sunshine duration as well as silicate, phosphate and nitrate concentrations. Due to missing parameters on seven different sampling dates (phosphate: 4 dates; silicate, nitrate temperature and salinity: 1 date each) the analysis was conducted with 273 samples.

Results

Baseline survey of parasitoid diversity

The 280 samples of the entire 18S metabarcoding dataset included 59,284 OTUs in total, of which 6056 OTUs (10.2%) were identified as potential parasitoids based on literature (see S2 Table for sequencing statistic). Over 55 percent of the dataset remained of unknown trophic mode due to insufficient taxonomic identification or missing reports on trophic modes. After setting a threshold of 0.001% of total reads, 2790 OTUs remained, of which 461 (16.5%) were identified as potential parasitoids based on their taxonomy and literature knowledge (S3 Table, see S1 File for comparison of results of different pipeline settings). For at least 124 parasitoid OTUs occurrence of taxa were known for Helgoland or nearby regions in the North Sea. Additionally, the assignment of parasitism or other trophic modes

was not possible for at least 50 percent of the remaining OTUs, which shows that there is still a great need for autecological studies on the plankton. Total reads of parasitoids were about ten times lower than non parasitoid reads and total relative abundances of parasitoids reached up to 45% (S1 Fig).

Parasitoid diversity, succession and influence of environmental conditions

The parasitoid OTUs belonged to ten different phyla (Table 1). These could be divided into 15 different classes, which are known to infect a wide range of hosts.

Table 1. Overview of parasitoid diversity on phylum and class level.

Phylum	OTU Count	Classes	Known hosts	References
Dinoflagellata	206	Syndiniales	Radiolaria, Dinoflagellata, Ciliates, Crustacea like Copepoda and Amphipoda, Cnidaria, Fish eggs, Chaetognatha	[20, 26, 31, 32, 65, 66]
Cercozoa	140	Endomyxa, Endomyxa-Phytomyxea, Filosa-Imbricatea, Filosa-Thecofilosea	Green plants, Brown algae, Diatoms and Stramenopiles	[18, 19, 27, 37, 67–72]
Stramenopiles_X	51	Oomycota, Pirsonia_Clade	Diatoms, Crustacea, Macro algae, Fish	[15, 17, 20, 23, 25, 36, 41, 73, 74]
Fungi	20	Ascomycota, Chytridiomycota	Cyanobacteria, Diatoms	[21, 75]
Apicomplexa	19	Apicomplexa_X	Arthropoda, Polychaeta, Chaetognatha, Copepoda, Euphausiacea, Dinoflagellata	[20, 31, 76]
Mesomycetozoa	14	Ichthyosporea	Diatoms, Fish, Mollusca Crustaceae	[77–79]
Ciliophora	5	Oligohymenophorea	Copepoda, Euphausiacea, Chaetognatha,	[80–83]
Metazoa	3	Nematoda	Hexapoda, Mollusca, Clitellata, Myriapoda, Crustacea, Annelida, Arthropoda	[84]
Lobosa	2	Tubulinea	Diatoms	[77]
Perkinsea	1	Perkinsida	Mollusca, Dinoflagellata	[85, 86]

https://doi.org/10.1371/journal.pone.0244817.t001

The dinoflagellate phylum contributed to this **amount of OTUs** with more than 44% of all parasitoid OTUs (Table 1). All of these belonged to the exclusively parasitoid Syndiniales. We identified Syndiniales from four out of the five different Dino-Groups as they are named by the PR2 database (also known as Syndiniales-Groups) (Fig 1). Dino-Group-II, also known as Syndiniales-Group II, contributed the most OTUs (76.7%), followed by Group I (17.5%). Most OTUs of the Syndiniales could not be assigned further than family level. In all Dino-Groups only three genera of Syndiniales could be identified by PR2: *Syndinium*, *Euduboscquella* and *Hematodinium*. BLAST alignment revealed that eight out of ten OTUs found in Group III most probably belonged to the genus *Amoebophyra*.

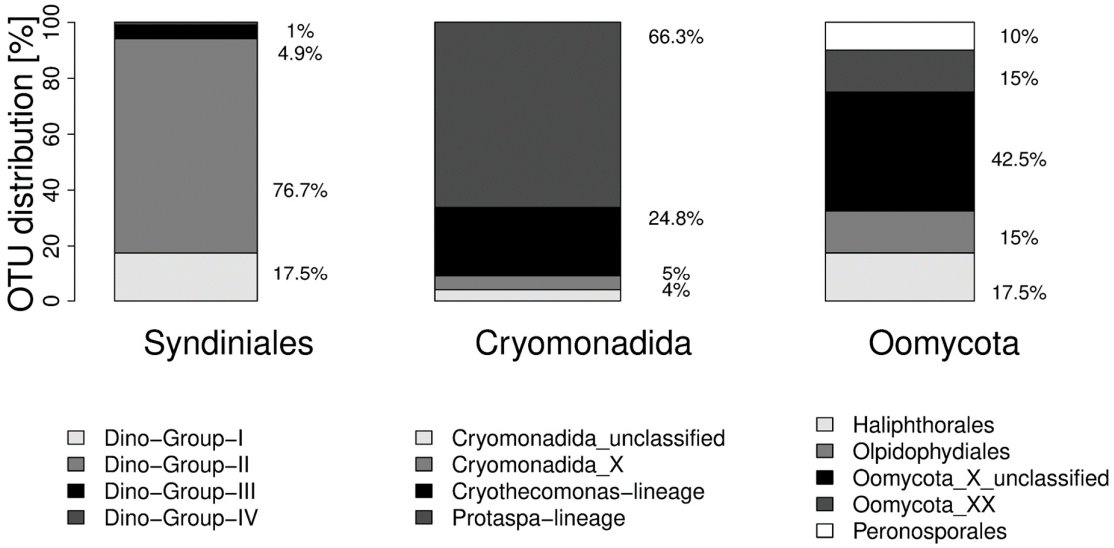


Fig 1. Distribution of OTUs in the different families of three different parasitoid taxa: Syndiniales, Cryomonadida and Oomycota.

Syndiniales were also the biggest contributor in **relative read abundance** of all parasitoids. 22.5% of all dinoflagellate reads (including non-parasitoids) belonged to Syndiniales. With regard to the distribution of Syndiniales reads, 73% belonged to Group II, followed by Group I with 22%. Group III (3.9%) and Group IV (0.3%) was detected in lower read abundances. (S4 Table). Syndiniales, as the only dinoflagellate parasitoids, could be found throughout all years and seasons with declines in relative parasitoid abundance during spring as well as during July (Fig 2A).

The next biggest contributor (30%) in terms of OTU numbers was the phylum Cercozoa (Table 1). The phylum had its highest relative abundances during March and April, especially in 2018 as well as during summer in 2017 (Fig 2B). It included four classes, namely Endomyxa, Endomyxa-Phytomyxea, Filosa-Imbricatea and Filosa-Thecofilosea (Table 1), with known hosts such as green plants, brown algae and Stramenopiles including diatoms. Of these classes, Filosa-Thecofilosea and Filosa-Imbricatea had the highest relative parasitoid abundances. The order Cryomonadida was the most abundant out of all parasitoid Cercozoa taxa. 9% of the Cryomonadida OTUs could not be identified further (Fig 1). The highest number of OTUs belonged to the Protaspa lineage.

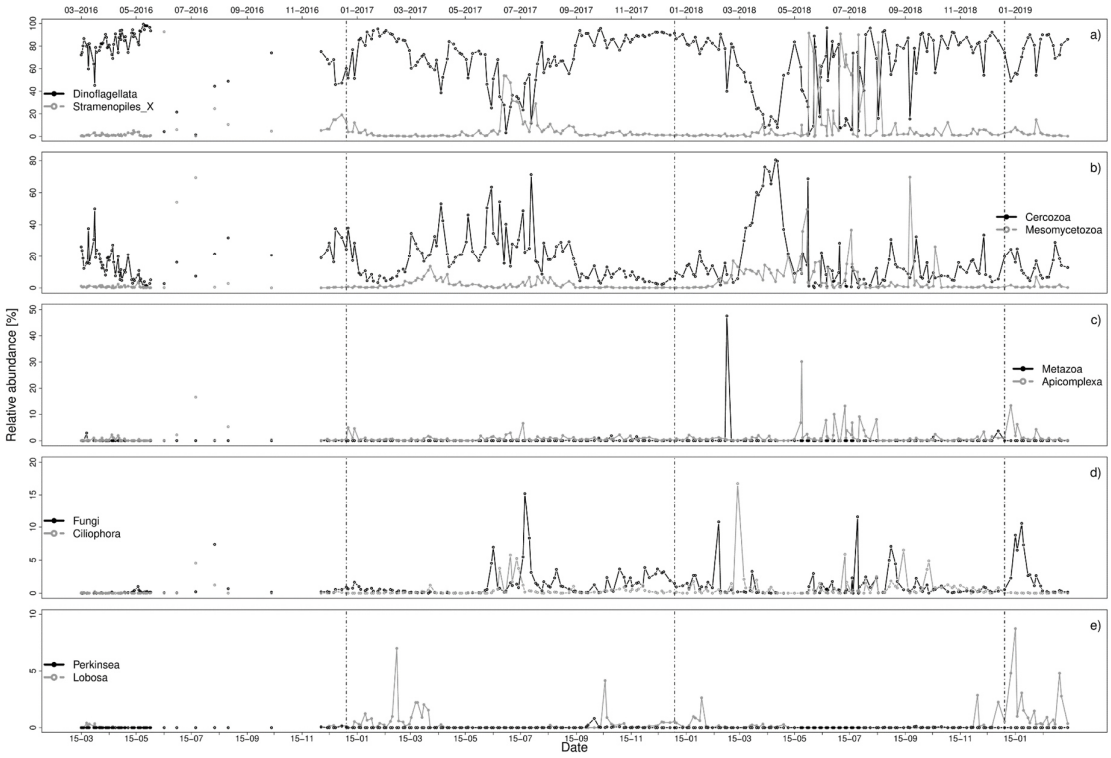


Fig 2. Relative parasitoid abundances [%] of parasitoid phyla, a) Dinoflagellata and Stramenopiles, b) Cercozoa and Mesomycetozoa, c) Metazoa and Apicomplexa, d) Fungi and Ciliophora, e) Lobosa and Perkinsea. Relative abundance is based on parasitoid taxa only. Note the different scaling of the axes. Vertical lines indicate turn of the years.

Parasitoid Stramenopiles made up over 10% of the parasitoid community (Table 1). While the phylum could be found in nearly all samples, the relative abundances of parasitoids were mostly low throughout the years, with peaks during summer months (Fig 2A). Highest relative parasitoid abundances were found in June 2016 (15-06-16), June/July 2017 and May to August 2018. We found two parasitoid Stramenopiles classes, namely *Pirsonia*-Clade (11 OTUs) and Oomycota (40 OTUs). Three families could be identified: Haliphthorales, Olpidophydiales and Peronosporales (Fig 1).

The phylum Mesomycetozoa included parasitoids of the class Ichthyosporea (Table 1), a group that can parasitize fish and crustaceans, which were mostly abundant during spring months (Fig 2B). In the phylum Fungi, parasitoid taxa in the classes Ascomycota and Chytridiomycota were found. Fungi were mainly present in June, July and August as well as during January 2019 (Fig 2D). Additional classes, some of which also included macro-parasite sequences in addition to parasitoids, were found mostly in low relative parasitoid abundances (Table 1 and Fig 2): Oligohymenophorea (Ciliophora), Apicomplexa_X (Apicomplexa), Nematoda (Metazoa), Tubulinea (Lobosa), Perkinsida (Perkinsea).

Each environmental parameter showed seasonal patterns as described below (S2 Fig, see also S1 Table), environmental conditions, therefore, were similar throughout all three years.

Water temperature ranged from 1.9˚C to 19.9˚C depending on the season, while salinity ranged from 29.0 to 34.2. Secchi depth varied between 0.3 and 8.7 meter with several fluctuations. Silicate and nitrate both ranged from 0 to over 29 µmol L^{-1}, highest concentrations were measured in winter and early spring months. Highest chlorophyll *a* concentrations were found in spring and summer with concentrations varying between 0.05 to 6.77 µg L^{-1}. Daily sunshine duration varied greatly from day to day and ranged from 0 hours of sunshine to 15.6 hours.

Based on the CCA model, which included all 2790 OTUs, all implemented parameters except for tide were found to be significantly associated to the community structure: season (AIC = 2020.9, p = 0.005), total parasitoid occurrence (AIC = 2019.2, p = 0.005), temperature (AIC = 2018.4, p = 0.005), salinity (AIC = 2018.1, p = 0.005), silicate (AIC = 2017.7, p = 0.005), sunshine duration (AIC = 2017.5, p = 0.005), phosphate (AIC = 2017.4, p = 0.04), nitrate (AIC = 2017.5, p = 0.005) and Secchi depth (AIC = 2017.7, p = 0.005). In total, only 12.1% of inertia could be explained by all variables in full space. In restricted space the first axis explained 21.9% of the variance (2.7% in full space) and the second axis explained 20.4% (2.5% in full space). The CCA plot (S3 Fig) indicated that high parasitoid occurrences were not clearly correlated with any environmental parameter nor any specific season.

Examples of known host-parasitoid systems at Helgoland

In the following, we display known host-parasitoid relationships, which were previously described in the literature and known to occur at Helgoland Roads, in order to check if the relationships can be found in the molecular dataset.

***Rhizosolenia imbricata–Olpidiopsis drebesii*.** OTU 39 was identified as *Rhizosolenia imbricata* by BLAST alignment with a Score of 701 (PR2: *Rhizosolenia* sp.) and compared to occurrences of OTUs that were identified as Oomycota by PR2. BLAST alignment revealed 18 OTUs as potential *Olpidiopsis* species. Inter alia, OTU 95 was assigned to *Olpidiopsis drebesii*. Host and parasitoid OTU occurred every year (Fig 3A). Blooms of the host (OTU39) occurred in June 2016 and 2017. In June 2016 and 2017 the parasitoid reached peaks as well. In 2017, *Rhizosolenia imbricata* reached its peak on June 20th, while a peak of *O. drebesii* followed 7 days later, resembling our assumed case 1.

Several *Olpidiopsis* and *Rhizosolenia* OTUs that were identified to genus level (Fig 3B) revealed additional peaks of parasitoids infections. In August 2016, peaks of the host (*Rhizosolenia* spp. and OTU 39) and *Olpidiopsis* spp. occurred on the same day, which represents our case 2 (Fig 3A and 3B). The five OTUs of *Rhizosolenia* spp. revealed

another bloom of the diatom in April and May 2018, however most peaks of that year were not closely linked to *Olpidiopsis* peaks.

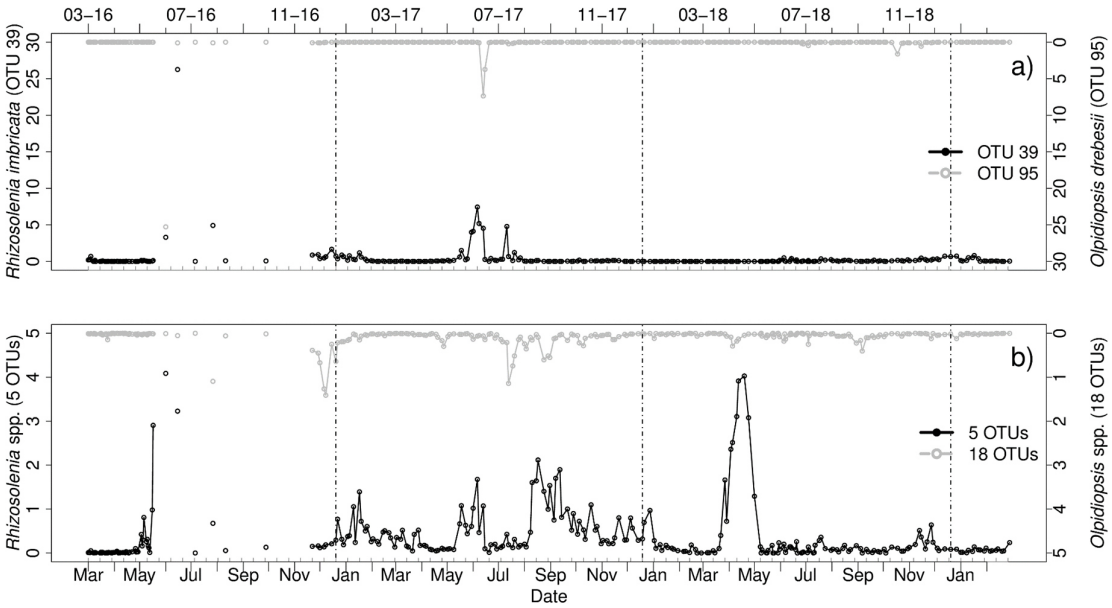


Fig 3. Relative abundances [%] of a) OTU 39 identified as *Rhizosolenia imbricata* (BLAST) and the parasitoid OTU 95 identified as *Olpidiopsis drebesii* (BLAST), and b) 5 OTUs identified as *Rhizosolenia* spp. (PR2) and 18 OTUs identified as *Olpidiopsis* spp. (BLAST) from March 2016 to March 2019. Vertical lines indicate turn of the years. Note the different scaling of the axes. Grey ticks on the x-axis indicate intervals of two weeks.

***Pseudo-nitzschia pungens–Miracula helgolandica*.** OTU 89 (Fig 4A), which was identified as *Pseudo-nitzschia pungens* (PR2), was found to be co-occurring with the parasitoid OTU 267 *Miracula helgolandica* (identification by BLAST, Score: 678). *Pseudo-nitzschia pungens* usually occurred in the spring and summer months. It was blooming during April 2016 (26– 04 to 29-04-16) and had further peaks in mid-May (06–05 to 12-05-16). In August, another peak was observed. In 2017, it was blooming in June and the highest peak was reached on June 06 (over 3%), followed by several smaller peaks in July (18–07 and 27-07-17) and August. The diatom was also blooming in summer 2018. It first peaked on June 13, followed by a second peak on June 19. The next big peak (over 4%) occurred in July (26-07-18). Afterwards a smaller peak followed on August 07.

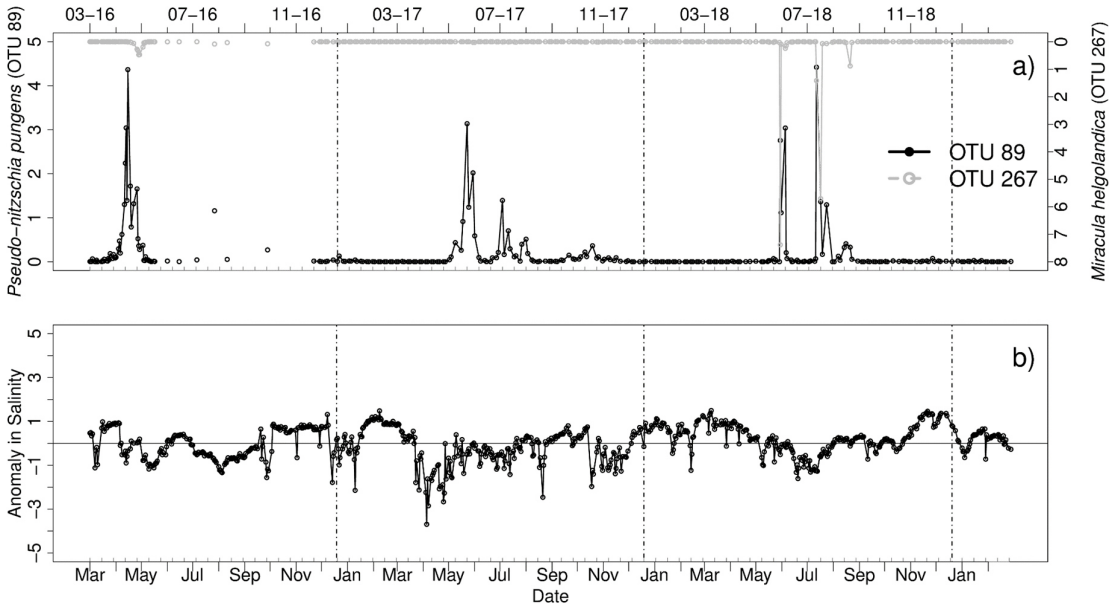


Fig 4. a) Relative abundances [%] of OTU 89 identified as *Pseudo-nitzschia pungens* (PR2) and the parasitoid OTU 267 identified as *Miracula helgolandica* (BLAST); b) anomaly in salinity from March 2016 to March 2019. Vertical lines indicate turn of the years. Note the different scaling of the axes. Grey ticks on the x-axis indicate intervals of two weeks.

The parasitoid OTU had its first occurrence during April and May 2016. For the rest of the year the parasitoid was either absent or occurrent without any distinct peak in abundance. In 2017, relative abundances were also low throughout the year and no distinct peak was detected. Several peaks, however, could be found in 2018, a first peak was reached in June (13-06-18) and a second peak appeared in July (31-07-18). The last smaller peak (below 1%) occurred in September (04-09-2018).

There were periods in 2016 and 2018, where host and parasitoid were closely aligned as defined for case 2. However, in 2017, large *P*. *pungens* blooms occurred without concurrent infection events. Comparison of host and parasitoid data with environmental conditions indicated that the absence of infections in 2017 coincided with a previous period of reduced salinity (Fig 4B).

***Coscinodiscus* sp.–*Lagenisma coscinodisci*.** Six OTUs were identified as *Coscinodiscus* sp., which included *Coscinodiscus wailesii* (OTU 113), two *C*. *radiatus* sp. (OTU 901 and 953) and three *Coscinodiscus* sp. which could not be further identified. OTU 2009 was identified as *Lagenisma coscinodisci* in BLAST (Score: 715). The parasitoid was found in 24 samples and in low relative abundances, as the maximum relative abundance was 0.25% on 31-07-18 (Fig 5A). Parasitoid read abundances peaked in August 2016 and 2017 (25-08-16, 08-08-17), and in June and July 2018 (13-06-18, 31-07-18). At these days no peaks of the host were found (Fig 5A, 5B and 5C).

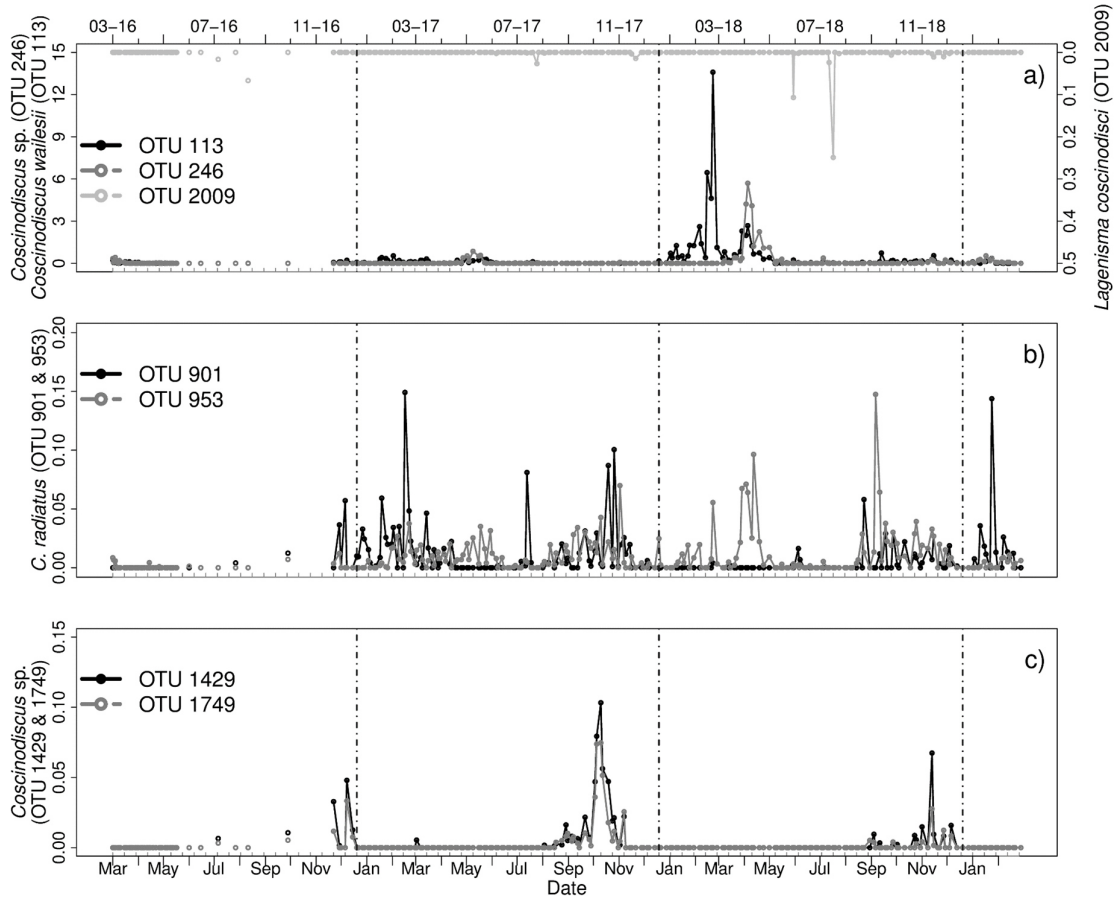


Fig 5. Relative abundances [%] of a) the parasitoid *Lagenisma coscinodisci* (BLAST, OTU 2009), the hosts *Coscinodiscus wailesii* (OTU 113), *Coscinodiscus* sp. (OTU 246); b) two potential *C. radiatus* sp. (OTU 901 and 953) and c) two *Coscinodiscus* sp. (OTU 1429 and 1749) from March 2016 to March 2019. Vertical lines indicate turn of the years. Note the different scaling of the axes. Grey ticks on the x-axis indicate intervals of two weeks.

All host OTUs occurred every year. In 2016 *Coscinodiscus wailesii* (OTU 113) was abundant in early spring and winter, in 2017 and 2018 in spring and summer and in winter 2018 until February 2019 (Fig 5A). It was blooming in February and March 2018 and had its biggest peaks during March 2018 (01-03-18: over 6%, 08-03-18: over 13%). A similar pattern was observed for OTU 246 (Fig 5A), which could only be identified up to genus level. Here the highest peak (over 5%) was found in April 2018. Two OTUs of *C. radiatus* (OTU 901 and 953) were only present in low relative abundances (below 0.02%) Both OTUs were continuously present during 2017 and 2018. OTU 901 had its biggest peaks in March 2017 and February 2019, OTU 953 in April and September 2018 (Fig 5B). The last two OTUs of *Coscinodiscus* sp. (1429, 1749) were also always below 0.2% in relative abundance and mostly present at the end of 2016, in autumn of 2017 and in winter 2018 (Fig 5C).

Co-occurrence as described by case 2 to the parasitoid was found for several of the host OTUs (OTU 113, 246, 953). However, no host peaks were aligned to peaks in the parasitoid.

Instead these hosts were always low in abundant. A peak of OTU 901 might be linked to a parasitoid peak in 2018, which would resemble our case 1 (12 days).

***Guinardia* sp.—Cryomonadida and *Pirsonia* clade.** Four OTUs of the diatom genus *Guinardia* (Figs 6 and 7) were found in the dataset: *Guinardia delicatula* (OTU 162, PR2), *Guinardia flaccida* (OTU 225, PR2). *Guinardia striata* (OTU 725, identified in BLAST, Score: 699) and OTU 1702 identified as *Guinardia striata* (BLAST, Score: 701). BLAST alignment of OTU 225 resulted in similar scores (701) for *G*. *flaccida* and *G*. *delicatula*, alignment of other OTUs confirmed the respective species as identified by PR2.

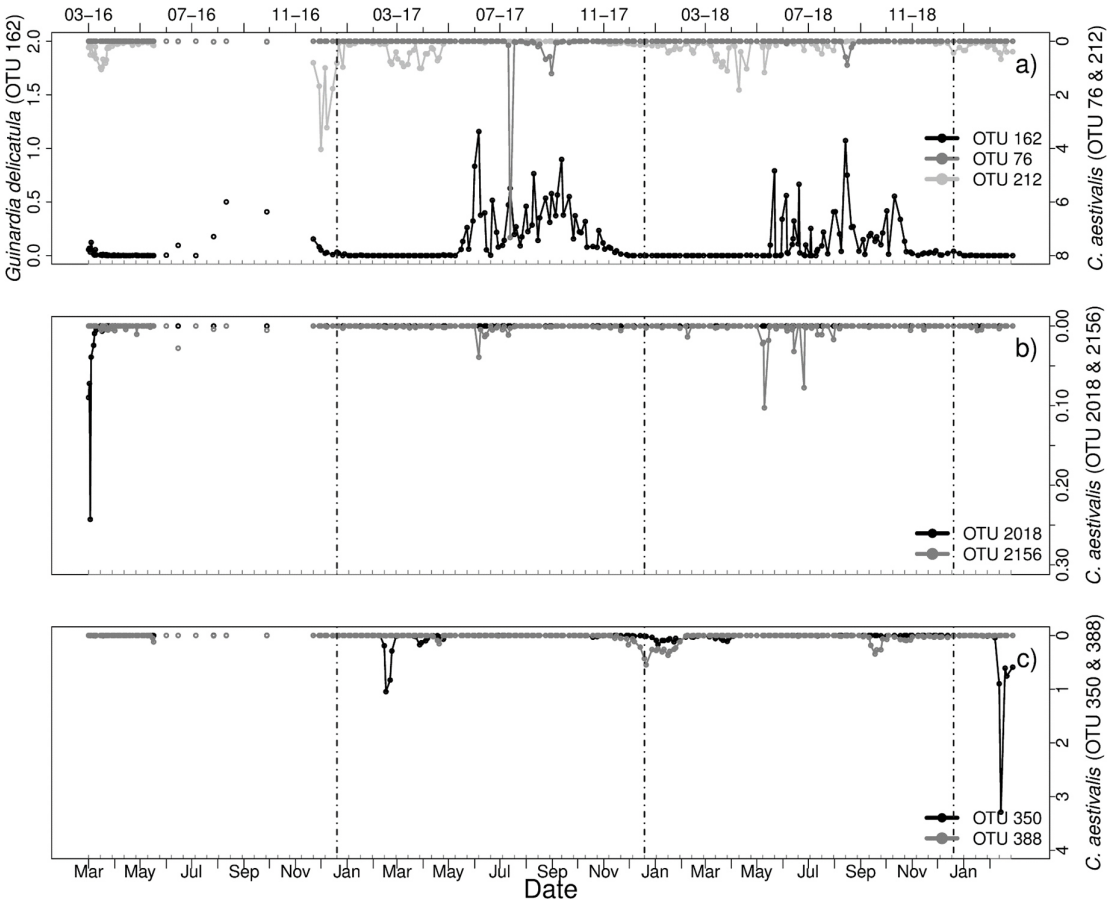


Fig 6. Relative abundances [%] of a) OTU 162 identified as *Guinardia delicatula*, and OTU 76 & 212 identified as *Cryothecomonas aestivalis* (BLAST), b) OTU 2018 & 2156 (*Cryothecomonas aestivalis*, BLAST) and c) OTU 350 & 388 (*Cryothecomonas aestivalis*, BLAST) from March 2016 to March 2019. Vertical lines indicate turn of the years. Note the different scaling of the axes. Grey ticks on the x-axis indicate intervals of two weeks.

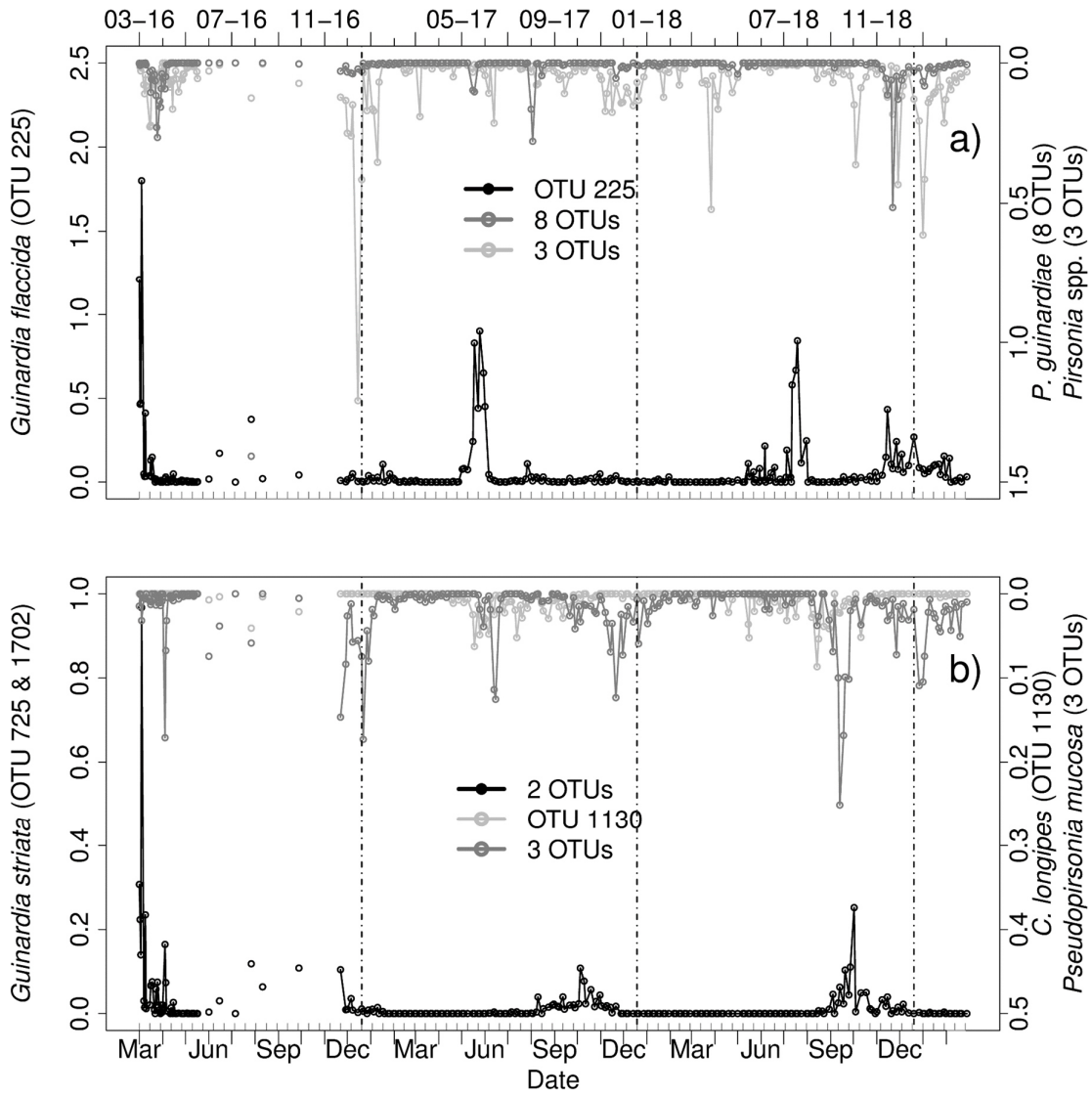


Fig 7. Relative abundances [%] of a) OTU 225 identified as *Guinardia flaccida* (PR2), *Pirsonia guinardiae* (8 OTUs) and *Pirsonia* spp. (3 OTUs), b) *Guinardia striata* (BLAST, OTU 725 and 1702), the parasitoid OTU 1130 identified as *Cryothecomonas longipes* (BLAST) and the parasitoid *Pseudopirsonia mucosa* (BLAST, 3 OTUs) from March 2016 to March 2019. Vertical lines indicate turn of the years. Note the different scaling of the axes. Grey ticks on the x-axis indicate intervals of two weeks.

First, the known host-parasitoid system of *G*. *delicatula* and *Cryothecomonas aestivalis* was investigated. Out of all Cryomonadida OTUs (in total 101 OTUs) 27 OTUs were found as potential *Cryothecomonas aestivalis* (see S5 Table for PR2 and BLAST results of potential parasitoids). These OTUs were checked for co-occurrences to the host *G*. *delicatula*. The parasitoid was found in all samples. Most parasitoid OTUs were also present while the host was not present in the dataset (Fig 6).

The host *G*. *delicatula* was present in every year (Fig 6A). In spring 2016 *G*. *delicatula* was mainly present in March with a peak on March 18. During summer 2016 two peaks were

detected in August (10–08 and 25–08). Furthermore, it was peaking on October 12 and in December 2016. In 2017 and 2018 *G*. *delicatula* was mostly occurring from May to December with several peaks and was blooming during June and July 2017 (e.g. between 15–06 to 20–06).

The association between *Guinardia delicatula* and *C*. *aestivalis* appeared to be complex and showed matches with different *C*. *aestivalis* OTUs throughout the sampling period as defined for case 2. For example for OTU 2018 in spring 2016, for OTU 2156 in June 2016, 2017 and 2018, and in July 2017, 2018 (Fig 6B), for OTU 76 in December 2016 and 2018 and for OTU 212 in summer 2017 (Fig 6A). For most OTUs the patterns hereby followed case 2, with simultaneous high abundances. Some parasitoid OTUs also showed high relative abundances after decline of the host OTU, such as OTU 76 in spring 2016, which indicates a relationship as described by case 1 in addition to co-occurrence as described by case 2. Additional peaks in parasitoid abundances did not match the occurrence of *G*. *delicatula*. These peaks, mainly occurring in late winter and early spring, included OTU 76 (January 2017, 2018 and February 2019), OTU 350 (February 2019) and OTU 388 in January 2018 (Fig 6A and 6C).

Cryothecomonas aestivalis is not the only parasitoid species known to infect *Guinardia* species. Additional *Cryothecomonas* species and *Pirsonia* clade were therefore also checked for cooccurrences with *G*. *delicatula* and other *Guinardia* OTUs (S5 Table). It is noteworthy that *G. flaccida* (OTU 225) had its highest relative abundances in March 2016 (Fig 7A) and occurred in low relative abundances without distinct peaks in February 2018, where other *Guinardia* OTUs were absent. BLAST alignment revealed eight out of eleven OTUs as potential *Pirsonia guinardiae*. Several co-occurrences (case 2) to their potential hosts were found throughout all years.

Furthermore, additional parasitoid OTUs were found to have similar occurrences compared to *Guinardia* OTUs (Fig 7A and 7B). These included for example OTU 1130, identified as *Cryothecomonas longipes* (BLAST, Score: 654) and three OTUs identified as *Pseudopirsonia sp*. and *P*. *muscosa*, respectively (PR2, verified in BLAST, S5 Table), indicative of possible additional infections as assumed by case 2 (Fig 7B).

Examples of known host-parasitoid systems recorded at Helgoland for the first time

In addition to known host-parasitoid relationships the data set revealed some potential host-parasitoid associations which had not been described before for the area of Helgoland but are known from other areas in the world.

Dinoflagellates–Perkinsida. We found one OTU belonging to the Perkinsida, which was identified as *Parvilucifera* sp. (PR2: *Parvilucifera prorocentri*). In BLAST it was identified as another Perkinsida species *Dinovorax pyriformis* (Score 516). As Perkinsida are known to infect dinoflagellates, the occurrence of this OTU (Fig 8A) was compared to the occurrence of known host species as well as additional dinoflagellates. *Parvilucifera prorocentri* peaked in September and October 2017, as well as in October 2018, with its highest peak occurring in 2017 on October 5. The two known host genera *Prorocentrum* sp. and *Dinophysis* sp. did not show a clear association with *P. prorocentri* as no peaks were detected in October 2017 (Fig 8B). However, corresponding to case 1, a time delay of seven days was observed between the maximum occurrence of *Akashiwo sp*., which was blooming in autumn 2017, and the parasitoid (Fig 8A).

***Eucampia zodiacus*–Cercozoa.** As the diatom *Eucampia zodiacus* is known to be infected by different species, the dataset was used to check for these potential parasitoids. Additionally, a parasitic infection was visible in several microscopic images (retrieved from planktonnet.awi.de, S4 Fig). The infections were visible in live cells from July as well as August 2017.

In our dataset *Eucampia zodiacus* was mostly present in summer 2017. The diatom host *Eucampia sp*. had a first peak (over 2%) on 25-07-17, a second bigger peak on 29-08-17 (over 2.8%) and a third smaller peak (over 0.5%) on 07-09-17 (Fig 9A). *Pirsonia*-Clade, which includes taxa that can infect *Eucampia zodiacus*, as well as Oomycota and Filosa-Thecofilosea abundances were compared to the occurrence of this host (Fig 9B). Several co-occurrences (case 2) and alternating associations (case 1) between the host and different parasitoids were found, including inter alia OTU 212 identified as Cryothecomonas aestivalis (BLAST Score: 673) and several OTUs belonging to *Pirsonia*-Clade (see S6 Table for PR2 and BLAST results of potential parasitoids).

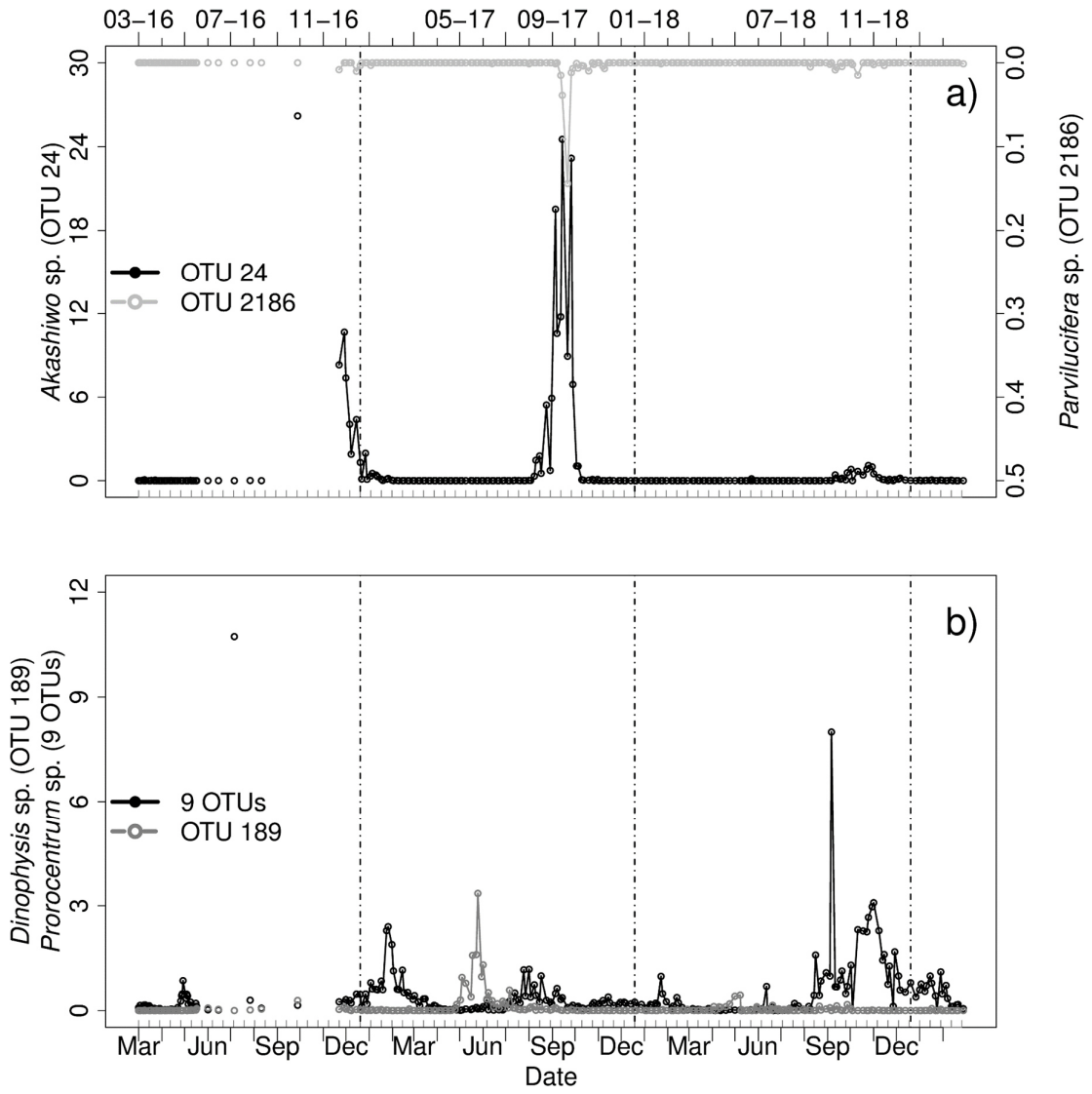


Fig 8. Relative abundances [%] of a) *Parvilucifera prorocentri* as identified by PR2 (OTU 2186) and *Akashiwo* sp. (OTU 24), b) *Prorocentrum* sp. (9 OTUs combined) and *Dinophysis* sp. (OTU 189) from March 2016 to March 2019. Vertical lines indicate turn of the years. Note the different scaling of the axes. Grey ticks on the x-axis indicate intervals of two weeks.

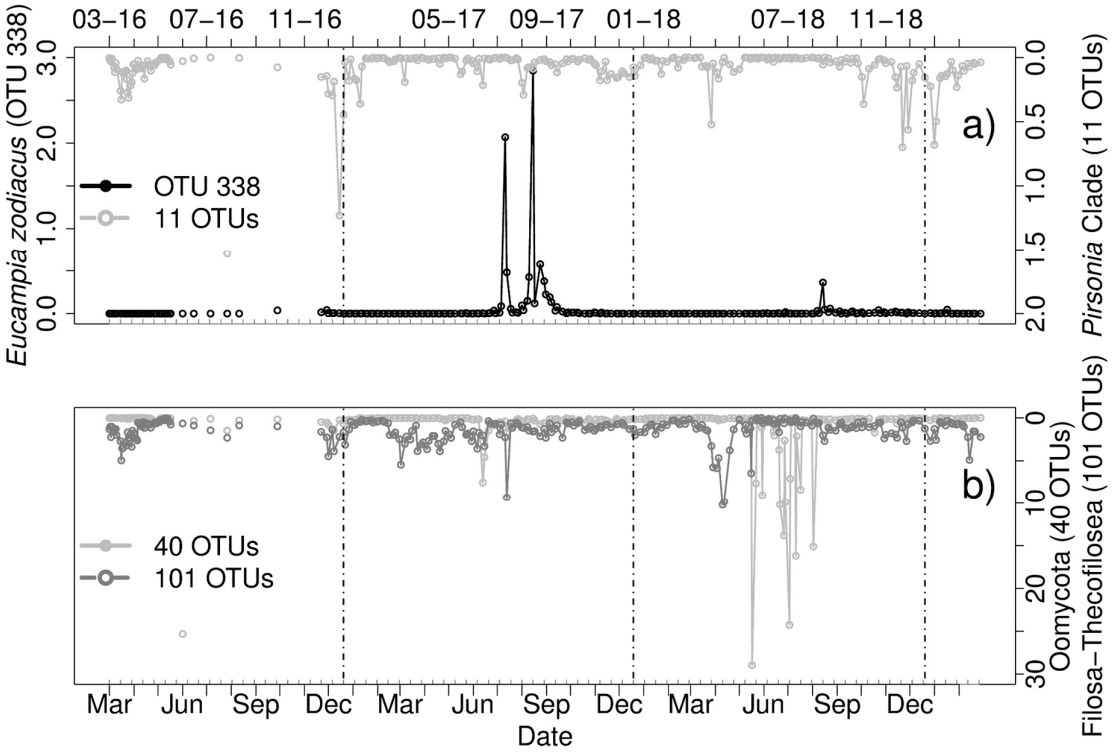


Fig 9. Relative abundances [%] of a) OTU 338 identified as *Eucampia* sp. (PR2) and the parasitoid taxa *Pirsonia* Clade (11 OTUs), b) Oomycota (40 OTUs) and Filosa-Thecofilosea (101 OTUs) from March 2016 to March 2019. Vertical lines indicate turn of the years. Note the different scaling of the axes. Grey ticks on the x-axis indicate intervals of two weeks.

Syndiniales genera–Crustacea & Tintinnida. Three different genera of Syndiniales (*Hematodinium* sp., *Eudubosquella* sp. and *Syndinium* sp.) could be identified and were compared to potential host OTUs. For *Hematodinium* sp. two peaks in relative abundance were found (02-01-18 and 27-12-18). The peak at the end of 2018 was co-occurring with high relative abundances of Crustacea (Fig 10A). This high abundance was mainly caused by 4 OTUs (identification by PR2): *Paracalanus* sp. (OTU 1), *Temora* sp. (OTU 2), unclassified Maxillopoda (OTU 27) and *Tachidius* sp. (OTU 38).

Favella sp. a known host of *Eudubosquella* sp. had its biggest peaks in occurrence from 27- 07-2017 to 03-08-2017 and in September 2018.The parasitoid occurred during all years with several peaks in abundance (Fig 10B). On August 24 2017, *Eudubosquella* sp. reached a peak in relative abundance of over 0.4%, where the host was also present. In 2018, the peak of the parasitoid occurred in absence of the host OTU. Some of the parasitoid peaks were also cooccurring with other Tintinnida.

Syndinium sp. also had several peaks in abundance, for example in December 2016, in August 2017 and from August to December 2018 (Fig 10C). Other peaks of *Syndinium* sp. were also co-occurring with *Paracalanus* sp. during all years.

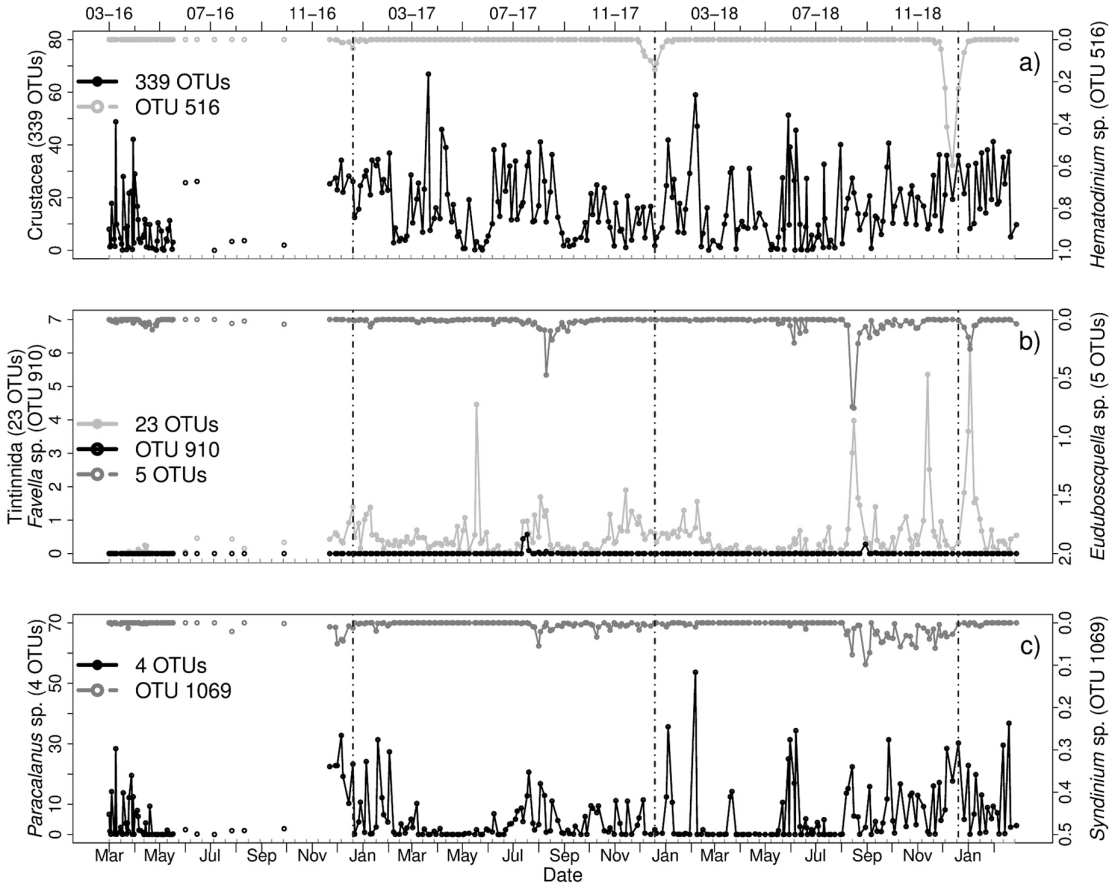


Fig 10. Relative abundances [%] of a) 339 OTUs identified as Crustacea (PR2) and the parasitoid *Hematodinium* sp. (OTU 516, Syndiniales, PR2), b) Tintinnida (23 OTUs), *Favella* sp. (OTU 910, PR2) and the parasitoid *Euduboscquella* sp. (5 OTUs, Syndiniales, PR2) and c) *Paracalanus* sp. (4 OTUs, PR2) and the parasitoid *Syndinium* sp. (OTU 1069, Syndiniales, PR2) from March 2016 to March 2019. Vertical lines indicate turn of the years. Note the different scaling of the axes. Grey ticks on the x-axis indicate intervals of two weeks.

Identification of potentially new host-parasitoid systems

Identification of new potential systems proved to be very difficult, since known systems as described in the previous paragraphs did not show consistent dynamics (see also Table 2). Thus using population dynamical information to identify other pairs based just on temporal dynamics of known interactions is not a promising venue. Particularly, the high diversity of potential parasitoid and hosts leaves a high level of speculation even on co-occurring OTUs.

Table 2. Overview of parasitoid dynamics.

System	Observed dynamics	Observed Time delay
Rhizosolenia imbricata–Olpidiopsis drebesii	Case 1 and 2	7 days
Pseudo-nitzschia pungens–Miracula helgolandica	Case 2	
Coscinodiscus sp.–*Lagenisma coscinodisci*	Case 1 and 2	12 days
Guinardia sp.—Cryomonadida and *Pirsonia* clade	Case 1 and 2	up to several days
Akashiwo sp.—*Parvilucifera prorocentri*	Case 1	7 days
Eucampia zodiacus—Cercozoa	Case 1 and 2	2 to 7 days
Syndiniales genera–Crustacea & Tintinnida	Case 2	

https://doi.org/10.1371/journal.pone.0244817.t002

Discussion

Identifying parasitoids

A wide diversity of parasitoids, which are known to be associated with a suite of different hosts, could be identified at Helgoland Roads. At the same time, the variability in the dynamics of known host-parasitoid pairs was considerable with many instances. For example, either hosts or parasitoids occurred separately, they showed some sort of Lotka-Volterra type alternating cycles or they co-occurred. Hence, our goal to use the dynamics of known pairs to identify potential thus far unknown host-parasitoid sets was essentially doomed from the start.

Due to the high abundances in parasitoids and the number of species present at different times of the year, infections can essentially occur throughout the year. For example, some parasitoid phyla were found as isolated events in a specific year such as Fungi, Apicomplexa, Metazoa and Perkinsea. Other taxa were present nearly throughout the whole sampling periods (e.g. Syndiniales and Cercozoa). Importantly, many trophic levels from primary producers to secondary consumers can potentially be affected. The potential hosts range from diatoms (e.g. Oomycota) to fish (e.g. Ichthyosporea) depending on the parasitoid species or group.

Highest abundances were found for the parasitoid dinoflagellates from the Syndiniales class. However, it was impossible to find clear correlations to potential hosts. The high read abundances are in accordance with generally high read abundances of dinoflagellates at Helgoland. Moreover, since it has been known that Syndiniales have low chromosome numbers compared to Dinophyceae [87], we can conclude that the high abundances are not caused by potential sequencing biases. Besides different Dino-Groups that cannot be further identified, we found known genera such as *Eudubosquella*, *Syndinium*, and *Hematodinium* present in our dataset. Among others, the three genera are known to infect tintinnid ciliates [88, 89], and crustaceans such as calanoid copepods, crabs and lobsters [20, 65, 90], respectively.

There have been suggestions about Syndiniales not always having a clear host-specificity [33]. For known genera, such as the parasitoid *Amoebophyra*, it has been shown that even though hosts were killed, other potential hosts in the same water mass were not declining even though a large number of dinospores were released [26]. The dinospores, that are released in large numbers, are short-lived and so far, they are known to complete their life cycle in a few days [33]. The high abundances are in accordance with other environmental studies, where Syndiniales showed high abundances especially in pico- and nanoplankton size fractions [91, 92], also in Antarctic winter [31]. It has been suggested that the free-living dinospores are mostly picoplanktonic, while an increase of abundances in bigger size fractions represent the parasitoids in their infectious stage in their host cell [22]. The fact that Syndiniales sequences can be found in high diversity throughout the year, could be explained in a number of scenarios. For instance it might be that they are only facultatively parasitoid, that production of new spores is either constant or that additional, so far unknown, life cycle stages exist [33], but this will require further investigation.

It needs to be noted that a majority of parasitoids is still poorly investigated on the molecular level as well. DNA sequences on species level are scarce for some groups including host taxa, which implies that protistan parasitoids can be even more diverse than known today [20]. As discussed before [47], there are several methodological issues such as choice of target region and database that influence identification. For example, comparison of V4 and V9 sequencing revealed differences in community diversity and weaknesses regarding identification of specific taxonomic groups like Chlorophyta, Ciliates or full eukaryotic communities [93–95]. The combination of different primer pairs and addition of mock communities to the analysis to decrease these weaknesses were suggested so far [95, 96]. Additionally, the V4 region has been found to have a bigger taxonomic resolution compared to the V9 region [97, 98]. The use of different pipelines results in not-reproducible outputs and differences in assigned taxa as it has been shown for diatoms [99], which makes it important to include all potential parameters in the methodology. While tuning on parameters might increase coverage of community composition, we focused on using a strict parameter set and a high confidence cut-off of annotation aiming for a high reliability. Furthermore, comparison of different parameter sets revealed that our main findings are pretty robust against changes in the parameter values. The drawback in molecular identification is also noticeable for the whole plankton community as identification not only on species level is scarce and assignment of trophic modes therefore is not possible for big parts of the community. It is also evident when comparing identification results from the PR2 database and BLAST alignment, where contradictory results occurred even for potential hosts, not only on species level (e.g. OTU 225: *Guinardia flaccida* or *G. delicatula*), but also

when comparing higher taxonomic levels. For example, while PR2 could identify OTU 725 only up to family level (Radial-centric-basal-Coscinodiscophyceae), through BLAST alignment it could be identified as *Guinardia striata* (Score: 699). Furthermore, PR2 identified several OTUs as belonging to the parasitoid *Protaspa*-lineage, whereas BLAST results indicate that the OTUs belong to *Cryothecomonas longipes*. Hereby, the BLAST results could be supported by construction of a maximum likelihood tree of the Cryomonadida OTUs (S5 Fig) in MEGA X [100] by use of Tamura-Nei model [101].

With respect to the influence of environmental conditions on parasitoids occurrences and infections, correlations with temperature are known. For example, for *Cryothecomonas aestivalis* infecting *Guinardia delicatula* on the New England Shelf, the highest infection rates only occurred at water temperatures of above 4˚C. The host on the other hand was blooming at a greater range of temperature below and above 4˚C [27]. This indicates that environmental conditions influence the presence of parasitoid and the opportunity for infections and that host and parasitoid are not necessarily perfectly synchronized in terms of their environmental tolerances. In our study, we cannot confirm this phenomenon. The host *G. delicatula* (OTU 162) was only found to be abundant, when the water temperature was above 5˚C, while *C. aestivalis* was present at all temperatures, that ranged from 2.7˚C to 19.7˚C. However, another host-parasitoid system indicates influence of the environmental conditions to development of the parasitoid. *Miracula helgolandica* was described and isolated from *P. pungens* at Helgoland [42]. While the host was present in high abundances during 2017, the parasitoid did not notably peak in abundance. Highest peak abundances in the host were found for temperatures above 10˚C (up to 19.7˚C) and the parasitoid occurred at similar temperature ranges except 2017. Anomalies in salinity might have influenced the availability of *Miracula* instead. Additionally, differences in timing and life cycle developments can be influential, especially since *P. pungens* occurred in short time corridors throughout the sampling period.

Recognizing known host-parasitoid systems using NGS

It was possible to find co-occurrences of known host-parasitoid systems at Helgoland such as host *Rhizosolenia imbricata* which was infected by *Olpidiopsis drebesii* [42] and *Pseudonitzschia pungens*, which is known to be infected by *Miracula helgolandica* [24, 42]. These parasitoids have been described as new species at Helgoland and since then could be observed frequently. The parasitoid *Lagenisma coscinodisci* has been observed in detail in the past [40, 41, 73, 74] and was found in our dataset, however, *Lagenisma coscinodisci* relative abundances were generally low throughout the sampling period.

Identification of other known systems turned out to be more complex with respect to host specificity and therefore their potential contribution to the seasonal dynamics within the plankton at Helgoland Roads. An example is presented by the genus *Guinardia*. The three species known to be present at Helgoland Roads, are all known to be parasitized by the parasitoids *Cryothecomonas*, *Pirsonia* and *Pseudopirsonia* [15, 18, 19, 23–25, 27]. In our study, parasitoid occurrences were overlapping with different species. For example, the peak in abundance of *Guinardia flaccida* during February 2017 and December 2018 was matching with several different parasitoid taxa such as *C. aestivalis*, *C. longipes*, *Pseudopirsonia muscosa* and *Pirsonia guinardiae*. This suggests that coincident infections of the identified *Cryothecomonas* OTUs and *Pirsonia* took place in this taxon. While this indicates that simultaneous infections by different parasitoids are likely, the loss or lack of host specificity of certain parasitoids also increase the complexity of the system.

A new potential host-parasitoid system for Helgoland was found for *Parvilucifera prorocentri* and an OTU of the genus *Akashiwo*. The parasitoid is known to have dinoflagellate hosts such as *Dinophysis* sp. and *Prorocentrum* sp. [86]. However, comparison of the occurrences showed, that these known host species were not associated with the parasitoid in our study. Our first assumption was loosely based on the Lotka-Volterra model, defined as periodic fluctuations with a certain time lag [30]. For *Akashiwo sp*. this assumption in predator-prey dynamics was observed. The example hints at the potential of parasitoids for controlling plankton blooms and their consequences for the food web. However, linking these rapid changes in host abundance to further potential host-parasitoid associations is not easy. So far, *Parvilucifera* infections of the dinoflagellate *Akashiwo sanguinea* were only observed in Masan Bay, Korea in April 2015 [102].

After comparison with other host-parasitoid systems, it was hard to detect alternating associations with time lags between host and parasitoid in addition to the *Akashiwo–P. prorocentri* system. For June 2017 we could find a delay of several days between the peaks of host *Rhizosolenia imbricata* and parasitoid *Olpidiopsis drebesii*, while this delay was not visible during other co-occurrences. For *Guinardia delicatula* and OTU 76 also both cases could be suggested, however simultaneous appearances, and therefore current infections (case 2), were mostly observed for all other *C. aestivalis* OTUs.

In addition to inspection of sequencing data, we could find microscopic evidence for a parasitic infection of the diatom *Eucampia zodiacus*. Thus far, known parasitoids for *Eucampia* are *Pirsonia* sp. like *Pirsonia eucampiae* and *Pirsonia formosa* [15] or *Paulsenella kornmannii* [103]. While *P. eucampiae* and *P. kornmannii* were not found in the sequencing dataset, *P. formosa* was identified as a potential parasitoid species. However,

this OTU was present during times, where *Eucampia zodiacus* was not detected and BLAST identification was inconclusive. Therefore, this is an indication that additional parasitoids are infecting *Eucampia*, which are still unknown. One potential parasitoid might be OTU 212, which peaked in abundance shortly after *Eucampia*. If looking at the dataset, several additional potential parasitoids were occurring simultaneously. However, some of these potential parasitoids are not likely infecting *Eucampia*. Some co-occurrences might happen by chance, since other potential hosts could be present at the same time. For example, several OTUs were identified as *Protaspa grandis*, which is bigger in size than the parasitoid which was found by microscopy. This species is known to reach sizes from 32.5–55.0 mm in length and 20.0–35.0 mm width [104]. In addition, visual comparison of known parasitoids indicates that some OTUs are unlikely to be a potential parasitoid of *Eucampia*. One example is *Olpidiopsis drebesii*, which can be excluded, if we inspect and compare the morphology as described for infections in *Rhizosolenia imbricata* [42].

Is identification of unknown host-parasitoid systems possible using NGS data?

In regard to high temporal resolution sequencing studies, previously observed host-parasitoid systems might not follow the expected dynamics. Since other co-occurrences were mostly found to be happening simultaneously and without delay between host and parasitoid and since DNA of the parasitoid should be able to be detected from its host, a match in peak abundance between host and parasitoid hints towards a current infection. In addition, for both – host and parasitoid – the environmental conditions need to be favourable for an infection to occur [105, 106]. Additional shifts in the physico-chemical environment, pertinently, in temperature and differences in thermal tolerances, in addition to changes in timing of occurrence, might cause the decoupling of existing host-parasitoid systems and the development of new relationships, increasing of infection rates and shifts in local food webs [107, 108]. In case of short-lived infections, long gaps in time between sampling might reduce recognition of this phenomenon. However, this is unlikely here due to our high sampling frequency in sampling phase 1 and 3, even though an even higher sampling frequency might cover short-lived infections that might occur within one day. Furthermore, knowledge of survival of parasitoids without their host and the life cycle of free-living states is scarce for most new described parasitoids since they are hard to detect with microscopy and mostly based on culturing experiments. While it is not possible to distinguish different stages in sequencing, the presence of the parasitoid can still be detected with this method. Another issue is the potential mismatch in timing of host and parasitoid occurrences and the influence of environmental conditions on the life cycles. Given the complexity of the life cycles, the diversity of parasitoid-host relationships within the system as well as their interaction with environmental conditions, it might be too simple to expect a typical Lotka-

Volterra type dynamic for identifying host-parasitoid systems, since typical and clear parasitoid-host phenomenon as described by Alves-de-Souza *et al.* [29] might be the exception rather than the rule.

The high dynamics of parasitoid occurrence and the variability in infection dynamics made it hard to detect host-parasitoid relationships using our sequencing dataset. Reasons for this might be the possibility of infections by several parasitoids either simultaneously or at different times, the fact that parasitoids could be plurivorous and that free-living stages cannot be distinguished by sequencing.

Conclusions

Our study is, to our knowledge the first, investigating multiple host-parasitoid systems and dynamics of parasitoids over a number of years. We have shown the high prevalence of parasitoids at Helgoland in high temporal resolution. The flexibility in parasitoid infections might have a big impact to the seasonal dynamics of the plankton community at Helgoland Roads. This highly detailed study also revealed several host-parasitoid systems with different temporal patterns such as simultaneous appearances, alternating cycles (with or without regular lags) and persistent parasitoid occurrence (Syndiniales). Potential systems that have been mentioned here, might be verified by microscopic and further molecular analysis such as newly developed fluorescence in situ hybridization probes. To adequately capture the complexity and high variability of host-parasitoid interactions and dynamics, further research on the dataset are necessary, especially since it was impossible to identify new systems with NGS alone.

Due to the high abundances, broad temporal occurrence patterns and their considerable diversity, we consider there to be a high likelihood of parasitoid infections on different components of the food web. The high diversity also shows that effects on the whole food web are likely, since parasitoids found are known to infect hosts of all trophic levels. While a high chance of parasitic infections adversely affects single hosts throughout the food web, this phenomenon might in contrast positively affect the whole community and the resilience of the system. The infection of one component of the food web can help the growth of other populations, which would not have evolved with the other population present. This in turn makes this topic even more relevant for future investigations on food web dynamics.

Supporting information

S1 Fig. Relative abundances [%] of parasitoids and non-parasitoid OTUs. Non-parasitoid OTUs include all remaining OTUs, that were not identified as Parasitoids; Vertical lines indicate turn of the years.

S2 Fig. Overview of environmental conditions, a) water temperature, Secchi depth, b) Salinity, Tide, c) Silicate, Nitrate, d) Chlorophyll a, Sunshine duration from March 2016 to March 2019. Vertical lines indicate turn of the years. Note the different scaling of the axes.

S3 Fig. Canonical Correspondence Analysis (CCA) of the samples (grey asterisks with sampling date) including significant parameters in black: Temperature (temp), salinity (sal), silicate (SiO4), nitrate (NO3), sunshine duration (sun), total parasitoid occurrence (parasitoids), seasons (spring, summer, autumn, winter) and tide (low tide, high tide). 12.2% of total inertia could be explained by all variables in full space, in restricted space CCA1 explained 23.8% of the variance and CCA2 explained 20.9%.

S4 Fig. Live cells of the centric diatom Eucampia zodiacus collected at Helgoland Roads, a) without parasitic infection (3rd August 2017), b)-d) with parasitic infection (b) 27th July 2017, c-d) 29th August 2017). Figures retrieved from planktonnet.awi.de.

S5 Fig. Maximum likelihood tree of Cryomonadida OTUs. The evolutionary history was inferred by using the Maximum Likelihood method and Tamura-Nei model [101]. The tree with the highest log likelihood (-3807.40) is shown. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Tamura-Nei model, and then selecting the topology with superior log likelihood value. This analysis involved 101 nucleotide sequences. There were a total of 397 positions in the final dataset. Evolutionary analyses were conducted in MEGA X [100].

S1 Table. Sampling information and environmental parameters.

S2 Table. Sequencing statistics. Raw: raw sequences after demultiplexing; Trimmed: remaining sequences after 3'-quality trimming; Assembled: remaining sequences after paired-end merging; Primer filtered: remaining sequences after removing primers; Feature filtered: remaining sequences after filtering for length; Sample derep: amount of unique sequences; Chimera filtered: remaining unique sequences after chimera removal; Final rerep: remaining sequences if we would rereplicate the sequences; Avg length: average length of each sequence in the sample.

S3 Table. Relative parasitoid abundances of parasitoid OTUs. Relative abundance is based on parasitoid taxa only.

S4 Table. Proportional distribution of Syndiniales clades detected over the whole timeframe.

S5 Table. Relative abundances [%] of potential parasitoids of *Guinardia* sp. after manual identification.

S6 Table. Relative abundances [%] of potential parasitoids of *Eucampia* sp. after manual identification.

S1 File. Additional information on bioinformatic pipeline and analysis.

Acknowledgments

We wish to thank the crew of the research vessel “Aade”, our technicians Kristine Carstens and Silvia Peters for collecting all LTER data and the MPI sampling team, especially Bernhard Fuchs, for sharing their samples. Additionally, we want to thank Swantje Rogge, Kerstin Oetjen and Pim Sprong for their support on the sequencing procedure.

Author Contributions

Conceptualization: Laura Käse, Katja Metfies, Maarten Boersma, Karen Helen Wiltshire, Alexandra Claudia Kraberg.
Data curation: Laura Käse, Stefan Neuhaus.
Formal analysis: Laura Käse, Stefan Neuhaus.
Funding acquisition: Maarten Boersma, Karen Helen Wiltshire.
Investigation: Laura Käse.
Methodology: Laura Käse, Katja Metfies, Stefan Neuhaus, Alexandra Claudia Kraberg.
Project administration: Katja Metfies, Maarten Boersma, Karen Helen Wiltshire, Alexandra Claudia Kraberg.
Resources: Katja Metfies, Maarten Boersma, Karen Helen Wiltshire, Alexandra Claudia Kraberg.
Software: Laura Käse, Stefan Neuhaus.
Supervision: Katja Metfies, Maarten Boersma, Karen Helen Wiltshire, Alexandra Claudia Kraberg.
Validation: Katja Metfies, Alexandra Claudia Kraberg.
Visualization: Laura Käse.
Writing – original draft: Laura Käse.
Writing – review & editing: Laura Käse, Katja Metfies, Stefan Neuhaus, Maarten Boersma, Karen Helen Wiltshire, Alexandra Claudia Kraberg.

References

1. Lafferty KD, Allesina S, Arim M, Briggs CJ, De Leo G, Dobson AP, *et al.* Parasites in food webs: The ultimate missing links. Ecol Lett. 2008; 11: 533–546. https://doi.org/10.1111/j.1461-0248.2008.01174.x PMID: 18462196

2. Lafferty KD, Dobson AP, Kuris AM. Parasites dominate food web links. Proc Natl Acad Sci. 2006; 103: 11211–11216. https://doi.org/10.1073/pnas.0604755103 PMID: 16844774

3. Dunne JA, Williams RJ, Martinez ND. Network structure and biodiversity loss in food webs: robustness increases with connectance. Ecol Lett. 2002; 5: 558–567. https://doi.org/10.1046/j.1461-0248.2002.00354.x

4. Lefèvre E, Roussel B, Amblard C, Sime-Ngando T. The Molecular Diversity of Freshwater Picoeukaryotes Reveals High Occurrence of Putative Parasitoids in the Plankton. Ibelings B, editor. PLoS One. 2008; 3: e2324. https://doi.org/10.1371/journal.pone.0002324 PMID: 18545660

5. Valois AE, Poulin R. Global drivers of parasitism in freshwater plankton communities. Limnol Oceanogr. 2015; 60: 1707–1718. https://doi.org/10.1002/lno.10127

6. Huxham M, Raffaelli D, Pike A. Parasites and Food Web Patterns. J Anim Ecol. 1995; 64: 168. https://doi.org/10.2307/5752

7. Marcogliese DJ. Pursuing parasites up the food chain: Implications of food web structure and function on parasite communities in aquatic systems. Acta Parasitol. 2001; 46: 82–93.

8. Arsenieff L, Simon N, Rigaut-Jalabert F, Le Gall F, Chaffron S, Corre E, *et al.* First Viruses Infecting the Marine Diatom *Guinardia delicatula*. Front Microbiol. 2019; 9. https://doi.org/10.3389/fmicb.2018.03235 PMID: 30687251

9. Kazamia E, Helliwell KE, Purton S, Smith AG. How mutualisms arise in phytoplankton communities: building eco-evolutionary principles for aquatic microbes. Fussmann G, editor. Ecol Lett. 2016; 19: 810–822. https://doi.org/10.1111/ele.12615 PMID: 27282316

10. Rohwer F, Thurber RV. Viruses manipulate the marine environment. Nature. 2009; 459: 207–212. https://doi.org/10.1038/nature08060 PMID: 19444207

11. Grossart H, Simon M. Interactions of planktonic algae and bacteria: effects on algal growth and organic matter dynamics. Aquat Microb Ecol. 2007; 47: 163–176. https://doi.org/10.3354/ame047163

12. Miki T, Jacquet S. Complex interactions in the microbial world: underexplored key links between viruses, bacteria and protozoan grazers in aquatic environments. Aquat Microb Ecol. 2008; 51: 195–208. https://doi.org/10.3354/ame01190

13. Munn CB. Viruses as pathogens of marine organisms—from bacteria to whales. J Mar Biol Assoc United Kingdom. 2006; 86: 453–467. https://doi.org/10.1017/S002531540601335X

14. Middelboe M, Brussaard CPD. Marine viruses: Key players in marine ecosystems. Viruses. 2017; 9: 1–6. https://doi.org/10.3390/v9100302 PMID: 29057790

15. Kühn SF, Drebes G, Schnepf E. Five new species of the nanoflagellate *Pirsonia* in the German Bight, North Sea, feeding on planktic diatoms. Helgoländer Meeresuntersuchungen. 1996; 50: 205–222. https://doi.org/10.1007/BF02367152

16. Kühn SF. Victoriniella multiformis, gen. et spec. nov. (incerta sedis), a polymorphic parasitoid protist infecting the marine diatom Coscinodiscus wailesiiGran & Angst (North Sea, German Bight). Arch für Protistenkd. 1997; 148: 115–123. https://doi.org/10.1016/S0003-9365(97)80039-5

17. Schweikert M, Schnepf E. Light and electron microscopical observations on *Pirsonia punctigerae* spec, nov., a nanoflagellate feeding on the marine centric diatom *Thalassiosira punctigera*. Eur J Protistol. 1997; 33: 168–177. https://doi.org/10.1016/S0932-4739(97)80033-8

18. Schnepf E, Kühn SF. Food uptake and fine structure of *Cryothecomonas longipes* sp. nov., a marine nanoflagellate incertae sedis feeding phagotrophically on large diatoms. Helgol Mar Res. 2000; 54: 18–32. https://doi.org/10.1007/s101520050032

19. Drebes G, Kühn SF, Gmelch A, Schnepf E. *Cryothecomonas aestivalis* sp. nov., a colourless nanoflagellate feeding on the marine centric diatom *Guinardia delicatula* (Cleve) Hasle. Helgoländer Meeresuntersuchungen. 1996; 50: 497–515. https://doi.org/10.1007/BF02367163

20. Skovgaard A. Dirty tricks in the plankton: Diversity and role of marine parasitic protists. Acta Protozool. 2014; 53: 51–62. https://doi.org/10.4467/16890027AP.14.006.1443

21. Rasconi S, Jobard M, Sime-Ngando T. Parasitic fungi of phytoplankton: Ecological roles and implications for microbial food webs. Aquat Microb Ecol. 2011; 62: 123–137. https://doi.org/10.3354/ame01448

22. de Vargas C, Audic S, Henry N, Decelle J, Mahé F, Logares R, *et al.* Eukaryotic plankton diversity in the sunlit ocean. Science. 2015; 348: 1261605–1261605. https://doi.org/10.1126/science.1261605 PMID: 25999516

23. Kühn S, Medlin L, Eller G. Phylogenetic Position of the Parasitoid Nanoflagellate Pirsonia inferred from Nuclear-Encoded Small Subunit Ribosomal DNA and a Description of *Pseudopirsonia* n. gen. and *Pseudopirsonia mucosa* (Drebes) comb. nov. Protist. 2004; 155: 143–156. https://doi.org/10.1078/143446104774199556 PMID: 15305792

24. Garvetto A, Nézan E, Badis Y, Bilien G, Arce P, Bresnan E, *et al.* Novel Widespread Marine Oomycetes Parasitising Diatoms, Including the Toxic Genus *Pseudo-nitzschia*:

Genetic, Morphological, and Ecological Characterisation. Front Microbiol. 2018; 9: 1–19. https://doi.org/10.3389/fmicb.2018.00001 PMID: 29403456

25. Tillmann U, Hesse K-J, Tillmann A. Large-scale parasitic infection of diatoms in the Northfrisian Wadden Sea. J Sea Res. 1999; 42: 255–261. https://doi.org/10.1016/S1385-1101(99)00029-5

26. Chambouvet A, Morin P, Marie D, Guillou L. Control of Toxic Marine Dinoflagellate Blooms by Serial Parasitic Killers. Science. 2008; 322: 1254–1257. https://doi.org/10.1126/science.1164387 PMID: 19023082

27. Peacock E, Olson R, Sosik H. Parasitic infection of the diatom *Guinardia delicatula*, a recurrent and ecologically important phenomenon on the New England Shelf. Mar Ecol Prog Ser. 2014; 503: 1–10. https://doi.org/10.3354/meps10784

28. Park MG, Yih W, Coats DW. Parasites and Phytoplankton, with Special Emphasis on Dinoflagellate Infections. J Eukaryot Microbiol. 2004; 51: 145–155. https://doi.org/10.1111/j.1550-7408.2004.tb00539.x PMID: 15134249

29. Alves-de-Souza C, Benevides TS, Menezes M, Jeanthon C, Guillou L. First report of vampyrellid predator–prey dynamics in a marine system. ISME J. 2019; 13: 1110–1113. https://doi.org/10.1038/s41396-018-0329-0 PMID: 30523275

30. Papkou A, Gokhale CS, Traulsen A, Schulenburg H. Host–parasite coevolution: why changing population size matters. Zoology. 2016; 119: 330–338. https://doi.org/10.1016/j.zool.2016.02.001 PMID: 27161157

31. Cleary AC, Durbin EG. Unexpected prevalence of parasite 18S rDNA sequences in winter among Antarctic marine protists. J Plankton Res. 2016; 38: 401–417. https://doi.org/10.1093/plankt/fbw005

32. Guillou L, Alves-de-Souza C, Siano R Dr, González H. The ecological significance of small, eukaryotic parasites in marine ecosystems. Microbiol Today. 2010; 37: 92–95.

33. Guillou L, Viprey M, Chambouvet A, Welsh RM, Kirkham AR, Massana R, *et al.* Widespread occurrence and genetic diversity of marine parasitoids belonging to Syndiniales (Alveolata). Environ Microbiol. 2008; 10: 3349–3365. https://doi.org/10.1111/j.1462-2920.2008.01731.x PMID: 18771501

34. Chambouvet A, Alves-de-Souza C, Cueff V, Marie D, Karpov S, Guillou L. Interplay Between the Parasite *Amoebophrya* sp. (Alveolata) and the Cyst Formation of the Red Tide Dinoflagellate *Scrippsiella trochoidea*. Protist. 2011; 162: 637–649. https://doi.org/10.1016/j.protis.2010.12.001 PMID: 21349764

35. Skovgaard A, Meneses I, Angélico MM. Identifying the lethal fish egg parasite *Ichthyodinium chabelardi* as a member of Marine Alveolate Group I. Environ Microbiol. 2009; 11: 2030–2041. https://doi.org/10.1111/j.1462-2920.2009.01924.x PMID: 19453613

36. Beakes GW, Glockling SL, Sekimoto S. The evolutionary phylogeny of the oomycete "fungi." Protoplasma. 2012; 249: 3–19. https://doi.org/10.1007/s00709-011-0269-2 PMID: 21424613

37. Scholz B, Guillou L, Marano A V., Neuhauser S, Sullivan BK, Karsten U, *et al.* Zoosporic parasites infecting marine diatoms–A black box that needs to be opened. Fungal Ecol. 2016; 19: 59–76. https://doi.org/10.1016/j.funeco.2015.09.002 PMID: 28083074

38. Wiltshire KH, Kraberg A, Bartsch I, Boersma M, Franke HD, Freund J, *et al.* Helgoland roads, north sea: 45 years of change. Estuaries and Coasts. 2010; 33: 295–310. https://doi.org/10.1007/s12237-009-9228-y

39. Wiltshire KH, Dürselen CD. Revision and quality analyses of the Helgoland Reede long-term phytoplankton data archive. Helgol Mar Res. 2004; 58: 252–268. https://doi.org/10.1007/s10152-004-0192-4

40. Thines M, Nam B, Nigrelli L, Beakes G, Kraberg A. The diatom parasite *Lagenisma coscinodisci (*Lagenismatales, Oomycota) is an early diverging lineage of the Saprolegniomycetes. Mycol Prog. 2015; 14: 75. https://doi.org/10.1007/s11557-015-1099-y

41. Schnepf E, Deichgräber G, Drebes G. Development and ultrastructure of the marine, parasitic oomycete, *Lagenisma coscinodisci* Drebes (Lagenidiales). Arch Microbiol. 1978; 116: 141–150. https://doi.org/10.1007/BF00406029

42. Buaya AT, Ploch S, Hanic L, Nam B, Nigrelli L, Kraberg A, *et al.* Phylogeny of *Miracula helgolandica* gen. et sp. nov. and *Olpidiopsis drebesii* sp. nov., two basal oomycete parasitoids of marine diatoms, with notes on the taxonomy of *Ectrogella*-like species. Mycol Prog. 2017; 16: 1041–1050. https://doi.org/10.1007/s11557-017-1345-6

43. Hickel W, Mangelsdorf P, Berg J. The human impact in the German Bight: Eutrophication during three decades (1962–1991). Helgoländer Meeresuntersuchungen. 1993; 47: 243–263. https://doi.org/10.1007/BF02367167

44. Wiltshire KH, Malzahn AM, Wirtz K, Greve W, Janisch S, Mangelsdorf P, *et al.* Resilience of North Sea phytoplankton spring bloom dynamics: An analysis of long-term data at Helgoland Roads. Limnol Oceanogr. 2008; 53: 1294–1302. https://doi.org/10.4319/lo.2008.53.4.1294

45. Grasshoff K. Methods of Seawater Analysis. 1st ed. Weinheim and New York: Verlag Chemie; 1976.

46. DWD Climate Data Center (CDC). Daily station observations of sunshine duration in hours for Germany. 2019. p. version v19.3, last accessed: 24.08.2019.

47. Käse L, Kraberg AC, Metfies K, Neuhaus S, Sprong PAA, Fuchs BM, *et al.* Rapid succession drives spring community dynamics of small protists at Helgoland Roads, North Sea. J Plankton Res. 2020; 42: 305–319. https://doi.org/10.1093/plankt/fbaa017 PMID: 32494090

48. Sprong PAA, Fofonova V, Wiltshire KH, Neuhaus S, Ludwichowski KU, Käse L, *et al.* Spatial dynamics of eukaryotic microbial communities in the German Bight. J Sea Res. 2020; 163: 101914. https://doi.org/10.1016/j.seares.2020.101914

49. Teeling H, Fuchs BM, Bennke CM, Krüger K, Chafee M, Kappelmann L, *et al.* Recurring patterns in bacterioplankton dynamics during coastal spring algae blooms. Elife. 2016; 5: 1–31. https://doi.org/10.7554/eLife.11888 PMID: 27054497

50. Chafee M, Fernàndez-Guerra A, Buttigieg PL, Gerdts G, Eren AM, Teeling H, *et al.* Recurrent patterns of microdiversity in a temperate coastal marine environment. ISME J. 2018; 12: 237–252. https://doi.org/10.1038/ismej.2017.165 PMID: 29064479

51. Sapp M, Wichels A, Wiltshire KH, Gerdts G. Bacterial community dynamics during the winter-spring transition in the North Sea. FEMS Microbiol Ecol. 2007; 59: 622–637. https://doi.org/10.1111/j.1574-6941.2006.00238.x PMID: 17381518

52. Fadeev E, Salter I, Schourup-Kristensen V, Nöthig E-M, Metfies K, Engel A, *et al.* Microbial Communities in the East and West Fram Strait During Sea Ice Melting Season. Front Mar Sci. 2018; 5: 1–21. https://doi.org/10.3389/fmars.2018.00043 PMID: 29552559

53. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics. 2014; 30: 2114–2120. https://doi.org/10.1093/bioinformatics/btu170 PMID: 24695404

54. Rognes T, Flouri T, Nichols B, Quince C, Mahé F. VSEARCH: a versatile open source tool for metagenomics. PeerJ. 2016; 4: e2584. https://doi.org/10.7717/peerj.2584 PMID: 27781170

55. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal. 2011; 17: 10–12. doi:https://doi.org/10.14806/ej.17.1.200

56. Mahé F, Rognes T, Quince C, de Vargas C, Dunthorn M. Swarm v2: highly-scalable and high-resolution amplicon clustering. PeerJ. 2015; 3: e1420. https://doi.org/10.7717/peerj.1420 PMID: 26713226

57. Mahé F, Rognes T, Quince C, de Vargas C, Dunthorn M. Swarm: robust and fast clustering method for amplicon-based studies. PeerJ. 2014; 2: e593. https://doi.org/10.7717/peerj.593 PMID: 25276506

58. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, *et al.* Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. Appl Environ Microbiol. 2009; 75: 7537–7541. https://doi.org/10.1128/AEM.01541-09 PMID: 19801464

59. Guillou L, Bachar D, Audic S, Bass D, Berney C, Bittner L, *et al.* The Protist Ribosomal Reference database (PR2): A catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. Nucleic Acids Res. 2013; 41: 597–604. https://doi.org/10.1093/nar/gks1160 PMID: 23193267

60. Diepenbroek M, Glöckner F, Grobe P, Güntsch A, Huber R, König-Ries B, *et al.* Towards an Integrated Biodiversity and Ecological Research Data Management and Archiving Platform: The German Federation for the Curation of Biological Data (GFBio). In: Plödereder E, Grunske L, Schneider E, Ull D, editors Informatik 2014 –Big Data Komplexität meistern GI-Edition: Lecture Notes in Informatics (LNI)–Proceedings, Vol 232, Bonn: Köllen Verlag; 2014, pp 1711–1724. 2014. Available: https://dl.gi.de/handle/20.500.12116/2782

61. Yilmaz P, Kottmann R, Field D, Knight R, Cole JR, Amaral-Zettler L, *et al.* Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIxS) specifications. Nat Biotechnol. 2011; 29: 415–420. https://doi.org/10.1038/nbt.1823 PMID: 21552244

62. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2020. Available: https://www.r-project.org/

63. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, Mcglinn D, *et al.* vegan: Community Ecology Package. 2019. Available: https://cran.r-project.org/web/packages/vegan/vegan.pdf

64. Akaike H. A new look at the statistical model identification. IEEE Trans Automat Contr. 1974; 19: 716–723. https://doi.org/10.1109/TAC.1974.1100705

65. Coats DW. Parasitic Life Styles of Marine Dinoflagellates. J Eukaryot Microbiol. 1999; 46: 402–409. https://doi.org/10.1111/j.1550-7408.1999.tb04620.x

66. Lima-Mendez G, Faust K, Henry N, Decelle J, Colin S, Carcillo F, *et al.* Determinants of community structure in the global plankton interactome. Science. 2015; 348: 1262073–1262073. https://doi.org/10.1126/science.1262073 PMID: 25999517

67. Hess S, Sausen N, Melkonian M. Shedding Light on Vampires: The Phylogeny of Vampyrellid Amoebae Revisited. Woo PCY, editor. PLoS One. 2012; 7: e31165. https://doi.org/10.1371/journal.pone.0031165 PMID: 22355342

68. Neuhauser S, Kirchmair M, Gleason FH. The ecological potentials of Phytomyxea ("plasmodiophorids") in aquatic food webs. Hydrobiologia. 2011; 659: 23–35. https://doi.org/10.1007/s10750-010-0508-0 PMID: 21339888

69. Neuhauser S, Kirchmair M, Gleason FH. Ecological roles of the parasitic phytomyxids (plasmodiophorids) in marine ecosystems—a review. Mar Freshw Res. 2011; 62: 365. https://doi.org/10.1071/ MF10282 PMID: 22319023

70. Neuhauser S, Kirchmair M, Bulman S, Bass D. Cross-kingdom host shifts of phytomyxid parasites. BMC Evol Biol. 2014; 14: 33. https://doi.org/10.1186/1471-2148-14-33 PMID: 24559266

71. Schnepf E, Kühn SF, Bulman S. *Phagomyxa bellerocheae* sp. nov. and Phagomyxa odontellae sp. nov., Plasmodiophoromycetes feeding on marine diatoms. Helgol Mar Res. 2000; 54: 237–241. https://doi.org/10.1007/s101520000056

72. Ward GM, Neuhauser S, Groben R, Ciaghi S, Berney C, Romac S, *et al.* Environmental Sequencing Fills the Gap Between Parasitic Haplosporidians and Free-living Giant Amoebae. J Eukaryot Microbiol. 2018; 65: 574–586. https://doi.org/10.1111/jeu.12501 PMID: 29336517

73. Drebes G. Ein parasitischer Phycomycet (Lagenidiales) in *Coscinodiscus*. Helgoländer Wissenschaftliche Meeresuntersuchungen. 1966; 13: 426–435. https://doi.org/ 10.1007/BF01611959

74. Schnepf E, Drebes G. Über die Entwicklung des marinen parasitischen Phycomyceten *Lagenisma coscinodisci* (Lagenidiales). Helgoländer Wissenschaftliche Meeresuntersuchungen. 1977; 29: 291–301. https://doi.org/10.1007/BF01614265

75. Elbrächter M, Schnepf E. Parasites of harmful algae. In: Anderson MD, Cembella AD, Hallegraeff GM, editors. Physiological Ecology of Harmful Algal Blooms. Berlin: Springer-Verlag; 1998. pp. 351–369. https://doi.org/10.1016/S1434-4610(98)70041-0 PMID: 23194717

76. Takahashi KT, Kawaguchi S, Kobayashi M, Toda T, Tanimura A, Fukuchi M, *et al.* Eugregarine infection within the digestive tract of larval Antarctic krill, *Euphausia superba*. Polar Biol. 2011; 34: 1167–1174. https://doi.org/10.1007/s00300-011-0979-0

77. Adl SM, Simpson AGB, Farmer MA, Andersen RA, Anderson OR, Barta JR, *et al.* The New Higher Level Classification of Eukaryotes with Emphasis on the Taxonomy of Protists. J Eukaryot Microbiol. 2005; 52: 399–451. https://doi.org/10.1111/j.1550-7408.2005.00053.x PMID: 16248873

78. Herr RA, Ajello L, Taylor JW, Arseculeratne SN, Mendoza L. Phylogenetic analysis of *Rhinosporidium seeberi*'s 18S small-subunit ribosomal DNA groups this pathogen among members of the protoctistan Mesomycetozoa clade. J Clin Microbiol. 1999; 37: 2750–4. https://doi.org/10.1128/JCM.37.9.2750-2754.1999 PMID: 10449446

79. Mendoza L, Taylor JW, Ajello L. The Class Mesomycetozoea: A Heterogeneous Group of Microorganisms at the Animal-Fungal Boundary. Annu Rev Microbiol. 2002; 56: 315–344. https://doi.org/10.1146/annurev.micro.56.012302.160950 PMID: 12142489

80. Estes AM, Reynolds BS, Moss AG. *Trichodina ctenophorii* N. Sp., a Novel Symbiont of Ctenophores of the Northern Coast of the Gulf of Mexico. J Eukaryot Microbiol. 1997; 44: 420–426. https://doi.org/10.1111/j.1550-7408.1997.tb05718.x PMID: 9304811

81. Gómez-Gutiérrez, Peterson WT, De Robertis A, Brodeur RD. Mass Mortality of Krill Caused by Parasitoid Ciliates. Science. 2003; 301: 339–339. https://doi.org/10.1126/science.1085164 PMID: 12869754

82. Gómez-Gutiérrez J, Peterson W, Morado J. Discovery of a ciliate parasitoid of euphausiids off Oregon, USA: *Collinia oregonensis* n. sp. (Apostomatida: Colliniidae). Dis Aquat Organ. 2006; 71: 33–49. https://doi.org/10.3354/dao071033 PMID: 16921999

83. Gómez-Gutiérrez J, Strüder-Kypke M, Lynn D, Shaw T, Aguilar-Méndez M, López-Cortés A, *et al. Pseudocollinia brintoni* gen. nov., sp. nov. (Apostomatida: Colliniidae), a parasitoid ciliate infecting the euphausiid Nyctiphanes simplex. Dis Aquat Organ. 2012; 99: 57–78. https://doi.org/10.3354/dao02450 PMID: 22585303

84. Blaxter M, Koutsovoulos G. The evolution of parasitism in Nematoda. Parasitology. 2015; 142: S26–S39. https://doi.org/10.1017/S0031182014000791 PMID: 24963797

85. Chambouvet A, Berney C, Romac S, Audic S, Maguire F, De Vargas C, *et al.* Diverse molecular signatures for ribosomally 'active' Perkinsea in marine sediments. BMC Microbiol. 2014; 14: 110. https://doi.org/10.1186/1471-2180-14-110 PMID: 24779375

86. Reñé A, Alacid E, Ferrera I, Garcés E. Evolutionary Trends of Perkinsozoa (Alveolata) Characters Based on Observations of Two New Genera of Parasitoids of dinoflagellates, *Dinovorax* gen. nov. and *Snorkelia* gen. nov. Front Microbiol. 2017; 8. https://doi.org/10.3389/fmicb.2017.01594 PMID: 28970818

87. Loeblich AR. Dinoflagellate Evolution: Speculation and Evidence*†. J Protozool. 1976; 23: 13–28. https://doi.org/10.1111/j.1550-7408.1976.tb05241.x PMID: 944775

88. Coats DW, Bachvaroff TR, Delwiche CF. Revision of the Family Duboscquellidae with Description of *Eduduboscquella crenulata* n. gen., n. sp. (Dinoflagellata, Syndinea), an Intracellular Parasite of the Ciliate *Favella panamensis* Kofoid & Campbell. J Eukaryot Microbiol. 2012; 59: 1–11. https://doi.org/10.1111/j.1550-7408.2011.00588.x PMID: 22221918

89. Jung J-H, Choi JM, Coats DW, Kim Y-O. *Eduduboscquella costata* n. sp. (Dinoflagellata, Syndinea), an Intracellular Parasite of the Ciliate *Schmidingerella arcuata*: Morphology, Molecular Phylogeny, Life Cycle, Prevalence, and Infection Intensity. J Eukaryot Microbiol. 2016; 63: 3–15. https://doi.org/10.1111/jeu.12231 PMID: 25963420

90. Stentiford G, Shields J. A review of the parasitic dinoflagellates *Hematodinium* species and *Hematodinium*-like infections in marine crustaceans. Dis Aquat Organ. 2005; 66: 47–70. https://doi.org/10.3354/dao066047 PMID: 16175968

91. Hernández-Ruiz M, Barber-Lluch E, Prieto A, Álvarez-Salgado XA, Logares R, Teira E. Seasonal succession of small planktonic eukaryotes inhabiting surface waters of a coastal upwelling system. Environ Microbiol. 2018; 20: 2955–2973. https://doi.org/10.1111/1462-2920.14313 PMID: 30187625

92. Thiele S, Wolf C, Schulz IK, Assmy P, Metfies K, Fuchs BM. Stable Composition of the Nano- and Picoplankton Community during the Ocean Iron Fertilization Experiment LOHAFEX. Paranhos R, editor. PLoS One. 2014; 9: e113244. https://doi.org/10.1371/journal.pone.0113244 PMID: 25401706

93. Dunthorn M, Klier J, Bunge J, Stoeck T. Comparing the Hyper-Variable V4 and V9 Regions of the Small Subunit rDNA for Assessment of Ciliate Environmental Diversity. J Eukaryot Microbiol. 2012; 59: 185–187. https://doi.org/10.1111/j.1550-7408.2011.00602.x PMID: 22236102

94. Tragin M, Zingone A, Vaulot D. Comparison of coastal phytoplankton composition estimated from the V4 and V9 regions of the 18S rRNA gene with a focus on photosynthetic

groups and especially Chlorophyta. Environ Microbiol. 2018; 20: 506–520. https://doi.org/10.1111/1462-2920.13952 PMID: 28984410

95. Choi J, Park JS. Comparative analyses of the V4 and V9 regions of 18S rDNA for the extant eukaryotic community using the Illumina platform. Sci Rep. 2020; 10: 6519. https://doi.org/10.1038/s41598-020-63561-z PMID: 32300168

96. Bradley IM, Pinto AJ, Guest JS. Gene-Specific Primers for Improved Characterization of Mixed Phototrophic Communities. Appl Environmental Microbiol. 2016; 82: 5878–5891. https://doi.org/10.1128/AEM.01630-16 PMID: 27451454

97. Stoeck T, Bass D, Nebel M, Christen R, Jones MDM, Breiner H-W, *et al.* Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. Mol Ecol. 2010; 19: 21–31. https://doi.org/10.1111/j.1365-294X.2009.04480.x PMID: 20331767

98. Pawlowski J, Christen R, Lecroq B, Bachar D, Shahbazkia HR, Amaral-Zettler L, *et al.* Eukaryotic richness in the abyss: Insights from pyrotag sequencing. PLoS One. 2011; 6. https://doi.org/10.1371/journal.pone.0018169 PMID: 21483744

99. Bailet B, Apothéloz-Perret-Gentil L, Baričević A, Chonova T, Franc A, Frigerio J-M, *et al.* Diatom DNA metabarcoding for ecological assessment: Comparison among bioinformatics pipelines used in six European countries reveals the need for standardization. Sci Total Environ. 2020; 745: 140948. https://doi.org/10.1016/j.scitotenv.2020.140948 PMID: 32736102

100. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 2018; 35: 1547–1549. https://doi.org/10.1093/molbev/msy096 PMID: 29722887

101. Tamura K, Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol. 1993; 10: 512–526. https://doi.org/10.1093/oxfordjournals.molbev.a040023 PMID: 8336541

102. Jeon BS, Nam SW, Kim S, Park MG. Revisiting the *Parvilucifera infectans*/*P. sinerae* (Alveolata, Perkinsozoa) species complex, two parasitoids of dinoflagellates. ALGAE. 2018; 33: 1–19. https://doi.org/10.4490/algae.2018.33.3.6

103. Drebes G, Schnepf E. *Paulsenella* Chatton (Dinophyta), ectoparasites of marine diatoms: development and taxonomy. Helgoländer Meeresuntersuchungen. 1988; 42: 563–581. https://doi.org/10.1007/BF02365627

104. Hoppenrath M, Leander BS. Dinoflagellate, Euglenid, or Cercomonad? The Ultrastructure and Molecular Phylogenetic Position of Protaspis grandis n. sp. J Eukaryot Microbiol. 2006; 53: 327–342. https://doi.org/10.1111/j.1550-7408.2006.00110.x PMID: 16968450

105. Mackenzie K. Parasites as Pollution Indicators in Marine Ecosystems: a Proposed Early Warning System. Mar Pollut Bull. 1999; 38: 955–959. https://doi.org/10.1016/S0025-326X(99)00100-9

106. Marcogliese DJ. Parasites: Small Players with Crucial Roles in the Ecological Theater. Ecohealth. 2004; 1: 151–164. https://doi.org/10.1007/s10393-004-0028-3

107. Hance T, van Baaren J, Vernon P, Boivin G. Impact of Extreme Temperatures on Parasitoids in a Climate Change Perspective. Annu Rev Entomol. 2007; 52: 107–126. https://doi.org/10.1146/annurev.ento.52.110405.091333 PMID: 16846383

108. Jeffs CT, Lewis OT. Effects of climate warming on host-parasitoid interactions. Ecol Entomol. 2013; 38: 209–218. https://doi.org/10.1111/een.12026

6 Publication III

Publication III, namely, “Metabarcoding analysis suggests that flexible food web interactions in the eukaryotic plankton community are more common than specific predator-prey relationships at Helgoland Roads, North Sea”, dealt the whole plankton community as characterized through metabarcoding analysis. Instead of looking at certain parts of the plankton community separately, we used the DNA present in one litre of water to capture all potential food web compartments ranging from picoplanktonic all the way up to meso- and macrozooplanktonic taxa in our analysis. The aim was to use metabarcoding to identify plankton communities and their succession over three years (aims 1a and b of this thesis). The plankton dataset included several taxa that were present at all times. Seasonal changes in the communities were caused among other things by small, abrupt appearances by autotrophs (blooms). Furthermore, we sought to identify predator-prey relationships in the food web and investigated associations among all food web components found during each season (aim 1d of this thesis). While it was not possible to detect clear connections between known predator-prey pairings, our network analysis revealed various associations of all different types of taxa in the food web. These diverse associations may hint at a large number of unidentified predator-prey relationships in the plankton community that would have to be investigated in further research.

Metabarcoding analysis suggests that flexible food web interactions in the eukaryotic plankton community are more common than specific predator–prey relationships at Helgoland Roads, North Sea

Laura Käse [1*], Katja Metfies[2,3], Alexandra C. Kraberg[3], Stefan Neuhaus[3], Cédric L. Meunier[1], Karen H. Wiltshire[1,4], Maarten Boersma[1,5]

[1]Alfred-Wegener-Institut, Helmholtz-Zentrum für Polar- und Meeresforschung, Biologische Anstalt Helgoland, Helgoland, Schleswig-Holstein, 27498, Germany

[2]Helmholtz-Institut für Funktionelle Marine Biodiversität, Eukaryotic Microbial Biodiversity and Ecology, Lower Saxony, Oldenburg, 26129, Germany

[3]Alfred-Wegener-Institut, Helmholtz-Zentrum für Polar- und Meeresforschung, Bremerhaven, Bremen, 27570, Germany

[4]Alfred-Wegener-Institut, Helmholtz-Zentrum für Polar- und Meeresforschung, Wadden Sea Station, List auf Sylt, Schleswig-Holstein, 25992, Germany

[5]Biology Department, University of Bremen, Bremen, 28359, Germany

*Corresponding author: laura.kaese@awi.de.

Received 7 January 2021; revised 26 February 2021; accepted 13 March 2021.

Published in ICES Journal of Marine Science (2021), 78(9), 3372 3386.
doi:10.1093/icesjms/fsab058

Käse, L., Metfies, K., Kraberg, A. C., Neuhaus, S., Meunier, C. L., Wiltshire, K. H., Boersma, M. Metabarcoding analysis suggests that flexible food web interactions in the eukaryotic plankton community are more common than specific predator–prey relationships at Helgoland Roads, North Sea. – ICES Journal of Marine Science, doi:10.1093/icesjms/fsab058.

Sequence data for this study have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession number PRJEB37135. Details of our pipeline are available on GitHub (https://github.com/PyoneerO/qzip). Additionally, the full OTU table (280 samples with 59.284 OTUs) was archived in PANGAEA (DOI: 10.1594/PANGAEA.921026).
All supplementary data can be found at ICES Journal of Marine Science online, supplementary figures and text can also be found in the appendix of this thesis.

Abstract

Various field studies on plankton dynamics have broadened our understanding of seasonal succession patterns. Additionally, laboratory experiments have described consumers ranging from generalists to selective grazers. While both approaches can give us a good understanding of the ecosystem and its dynamics, drawbacks in identification and a limited coverage of the ecosystem have left open questions on the generality of previous results. Using an integrative approach, we investigated water samples taken at Helgoland Roads by metabarcoding to describe seasonal succession patterns of the whole plankton community. By use of network analysis, we also tried to identify predator–prey dynamics. Our data set depicted the strong seasonality typically found for temperate waters. Despite a stable background community surviving strong fluctuations, small and abrupt changes, such as pronounced blooms and random appearance of autotrophs, cause seasons to be quite different in an inter-year comparison. Main consumers were copepods, ciliates, and dinoflagellates, of which the latter were most abundant. Furthermore, our results suggest that zooplankton predators might favour specific prey during certain time periods but seem to be quite opportunistic otherwise throughout the year. The variability and potential for many different relationships in the plankton community might be an indicator of resilience in the system.

Keywords: diversity, German Bight, Illumina MiSeq sequencing, Long-Term Ecological Research (LTER), network analysis.

Introduction

The immense diversity and size spectrum of marine eukaryotic plankton makes it difficult to capture and study the whole community at once. As a result, our understanding of marine ecosystems is somewhat fragmentary, as field studies out of methodological necessity focus typically on individual community compartments, either in temporally explicit one-point

time-repeated measurements or spatially explicit (one-time many points) settings. The findings from these field studies, or the questions arising from them, are typically then addressed using laboratory, mesocosm, or whole-field experiments. The use of frequent monitoring at Long-Term Ecological Research (LTER) sites is another approach to broaden the view on seasonal and longer-term scales.

The unique observation programme Helgoland Roads studies the long-term development of abiotic factors such as temperature, and the resulting dynamics of the planktonic community, at the Helgoland Roads station in the North Sea (Wiltshire *et al*., 2010). This programme has been boosted by additional studies focusing on different, specific, plankton groups to provide a more detailed view on their dynamics (e.g. Wiltshire and Dürselen, 2004; Medlin *et al*., 2006; Sapp *et al*., 2007; Knefelkamp, 2009; Metfies *et al*., 2010; Löder *et al*., 2011, 2012; Schlüter *et al*., 2012; Yang *et al*., 2014; Boersma *et al*., 2015; Lucas *et al*., 2015; Wiltshire *et al*., 2015; Scharfe and Wiltshire, 2019). These studies helped to improve the understanding of the seasonal dynamics of the different food web compartments at Helgoland. The abundance of picoeukaryotes, for instance, was found to be the highest in spring and summer and rapid changes could be observed throughout the year at Helgoland (Knefelkamp, 2009). It was also shown that the contribution of cryptophytes to the picoplankton was the highest in winter and spring (Metfies *et al*., 2010). Autotrophic plankton such as diatoms is known to be highly abundant during spring, when environmental conditions such as temperature and light availability increase (Mieruch *et al*., 2010). While heterotrophic dinoflagellates were found to be the most important contributors to biomass in general, mixotrophic dinoflagellates and ciliates can also significantly contribute to the total planktonic biomass at certain times of the year (Löder *et al*., 2012). Generally, microzooplankton might exert a stronger top-down control on phytoplankton at Helgoland than mesozooplankton (Löder *et al*., 2011).

Whereas the above-described specific studies on plankton trophodynamics have certainly increased our knowledge of relatively short-term patterns, or of single components of the food web, an overall assessment of the complete planktonic component is still lacking. Furthermore, by focussing on conspicuous or short-term food web interactions these might be over-interpreted, if they are assumed to be a regular phenomenon. New technology and analytical approaches might fill this knowledge gap by providing information on the complete plankton community as well as on interactions between different food web components, such as between nano- and picoplankton, and micro- and macroplankton. For instance, metabarcoding approaches have been used to study planktonic microbial communities around Helgoland, allowing the identification of new succession patterns in bacterioplankton

throughout the year (Chafee *et al.*, 2018), and in small eukaryotic plankton during spring (Käse *et al.*, 2020).

Ideally, metabarcoding studies should be used to study the whole ecosystem at once, integrating all components. These integrated approaches are, however, rare (but see Abad *et al.*, 2016, 2017). Instead, most metabarcoding studies either focus on smaller components of the community, by studying water samples, after removing larger mesozooplankton (Hernández-Ruiz *et al.*, 2018; Rachik *et al.*, 2018; Giner *et al.*, 2019; Sprong *et al.*, 2020) or they study the larger zooplankton, typically using net samples (Lindeque *et al.*, 2013; Hirai and Tsuda, 2015; Hirai *et al.*, 2015; Sommer *et al.*, 2017; Bucklin *et al.*, 2019; Blanco-Bercial, 2020). Here, we aimed to integrate these approaches by investigating the whole planktonic community using metabarcoding at once.

With this method, knowledge of zooplankton biodiversity, which is probably much higher than known today, and the functional role of different zooplankton species in the planktonic food web can be further enhanced. The links in the planktonic food web, identified by metabarcoding, can be visualized by conducting network analysis (Kurtz *et al.*, 2015). Hereby, interactions or associations (for example, in terms of co-occurrence) are shown by edges (also called links) that connect different nodes (also called vertices). By using these kinds of techniques, we can break down complex community structures into distinct clusters at different times. These clusters can then be used to obtain new insights into the relationships within the planktonic community throughout the food web, and to discover potentially new or to verify previously observed predator–prey relationships.

Hence, for the first time, we conducted a metabarcoding study over a 3-year period to provide a comprehensive assessment of the annual succession of species constituting the planktonic food web presented in 1 L of water. In addition to identifying plankton community diversity as a whole, we suggest that metabarcoding analysis of natural seawater could provide unique insights into potential predator–prey relationships within the planktonic food web. Our aims were (i) to identify plankton communities that occurred from 2016 to 2019 and their succession using metabarcoding analysis and (ii) to identify predator–prey dynamics with regard to zoo- and phytoplankton. We used information on previously observed predator–prey pairings to check for co-occurrences of consumers/predators (in the following only referred to as predators) and prey in the metabarcoding data set. Furthermore, conducting of network analyses gave us a unique possibility to investigate associations and to corroborate potential relationships in the plankton community.

Materials and methods

Study site and sampling

Work daily water surface samples (depth: 1 m) were taken with a bucket at the Helgoland Roads LTER sampling site (54°11′N, 7°54′E) from mid-March 2016 to mid-March 2019. The LTER site is situated between the main island and the dune island of Helgoland (Figure 1). Due to the strong tidal currents, the surface samples are representative of the complete water column at the station (Hickel, 1998; Wiltshire *et al*., 2015). Depending on the tides, the well-mixed water column can fluctuate between 10 m depth (Callies and Scharfe, 2015). Measuring of Secchi depth and temperature was conducted directly at the sampling site. Other parameters were measured in the laboratory according to the LTER protocols (Hickel *et al*., 1993; Wiltshire *et al*., 2008, 2010). These include salinity and nutrients such as silicate, phosphate, and inorganic nitrogen (nitrate, nitrite, and ammonium; Grasshoff, 1976) and chlorophyll *a*. Daily observations of sunshine duration in hours were downloaded from the Deutscher Wetterdienst, Climate Data Centre (DWD Climate Data Center (CDC), 2019); and the seasons were defined using the meteorological calendar: Spring = March to May, Summer = June to August, Autumn = September to November, Winter = December to February.

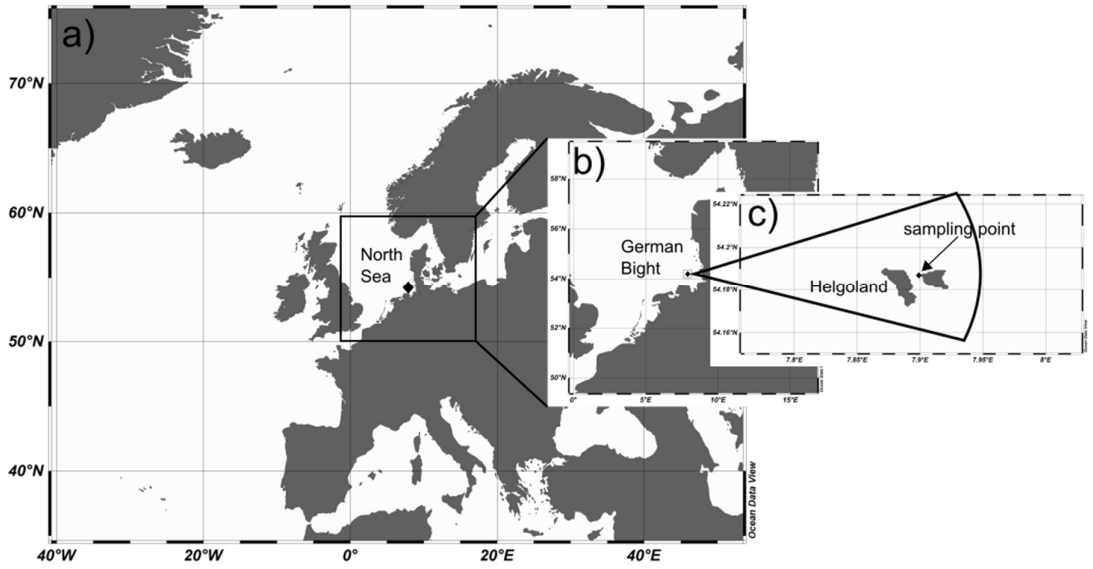


Figure 1. Map of the (a) North Sea, Europe, (b) German Bight in the North Sea, (c) Helgoland Roads sampling point in between the main island and dune island of Helgoland.

Molecular analysis

We combined samples from three different sampling periods for metabarcoding analysis. The sampling protocols on sampling frequency, filtration, and DNA extraction steps of the different sampling periods were not identical. However, each sample was taken with a

bucket and 1 L of seawater was filtered. The first sampling period from 15 March 2016 to 31 May 2016 included work daily sampling and a sequentially filtration using 10 µm polycarbonate filters, 3 µm PC filters, and 0.2 µm polyvinylidene fluoride filters (Millipore, Schwalbach, Germany; Teeling *et al*., 2016; Käse *et al*., 2020). The samples from the other two sampling periods (from June 2016 to March 2019) were filtered onto 0.45 µm nylon filters (Whatman, 47 mm). Samples from June to October 2016 were taken infrequently (in total six samples; Sprong *et al*., 2020). From December 2016 until 14 March 2019, the samples were analysed twice a week, with additional samples from mid-May to the end of July 2018 (three samples per week). All samples were stored at −20°C until DNA isolation.

General methods on DNA extraction, MiSeq™ Illumina sequencing, and data processing have been described elsewhere (Käse *et al*., 2021).

In short, two different protocols were used for DNA extraction. DNA extraction for the 10 µm, 3 µm filters of the sampling period in spring 2016 and all 0.45 µm filters from June 2016 to March 2019 was carried out with the Macherey-Nagel NucleoSpin® Plant II Kit. DNA extraction from 0.2 µm filters from spring 2016 was conducted as described previously by Sapp *et al*. (2007). It should be mentioned that DNA of multi-celled zooplankton or other big organisms is o based on DNA that occurred in the 1 l water sample. This can include DNA sticking to particle surfaces or faecal pellets as well as free DNA. No additional samples of these big organisms were included and macroalgae, copepods, or other mesozooplankton, that were visible to the naked eye, were not present during the DNA extraction process.

Sequentially filtered samples were pooled accordingly to achieve one sample per sampling date. A fragment (V4 region) of the 18S ribosomal (r) DNA was amplified using KAPA HiFi HotStartReadyMix (Kapa Biosystems, Inc., Roche, Germany) and the following primer set developed by Fadeev *et al*. (2018): 528iF (GCG GTA ATT CCA GCT CCA A) and 964iR (ACTTT CGT TCT TGA TYR R). About 43 million 2 × 300 bp paired-end sequences were produced using an Illumina MiSeq™ sequencer (Illumina, San Diego, CA, USA). After data processing, 21 million sequences remained and were clustered using s*warm* (version 2.2.2, Mahé *et al*., 2014, 2015). The Protist Ribosomal Reference database (PR2), version 4.11.1 (Guillou *et al*., 2013) was used as reference and all names are based on the taxonomy as it is given by the database. Sequence data for this study have been deposited in the European Nucleotide Archive (ENA) at European Bioinformatics Institute (EMBL-EBI) under accession number PRJEB37135, using the data brokerage service of the German Federation for Biological Data (GFBio, Diepenbroek *et al*., 2014), in compliance with the Minimal Information about any (X) Sequence (MIxS) standard (Yilmaz *et al*., 2011). Details of our pipeline are available on GitHub (https://github.com/PyoneerO/qzip) and the full table

of operational taxonomic units (OTU280 samples with 59.284 OTUs) was archived in PANGAEA (https://doi.pangaea.de/10.1594/PANGAEA.921026).

A threshold of 0.001% (of total reads) was applied to the full data set, resulting in a data set of 2790 OTUs, which was used for all further analyses. Identification up to genus level was accepted. Where necessary, identification on species level was verified with BLAST. For taxa that could not be further identified, the previous taxonomic level was assigned; these objects were indicated with an additional term (e.g. unclassified) and they were either included as a different taxon on the respective taxonomic level or not used for further analysis.

Statistical analysis

All follow-up analyses, as described below, were conducted in R (version 4.0.2, R Core Team, 2020). For significance tests, the significance level was set at $p < 0.05$.

Community diversity

For alpha diversity, the number of OTUs per season was counted. For principal component analysis (PCA) and beta diversity calculation the OTU table (280 samples with 2790 OTUs) was centred log-ratio (clr) transformed. The PCA was conducted with the “pca” function of the mixOmics package (Rohart *et al.*, 2017). Beta diversity was calculated with the Whittaker index (Whittaker, 1960) using the “betadiver” function of the vegan package (Oksanen *et al.*, 2019). The “hclust” function was used to convert the matrix into a cluster, which was then cut into five branches (at $h = 0.8$) and visualized. Additionally, clusters were defined at $h = 0.5$. The clusters at $h = 0.5$ were tested for significance with a permutational multivariate analysis of variance (PERMANOVA) using the “adonis2” function in the vegan package with 10 000 permutations. This analysis was also used to check for significances of the beta diversity matrix to the different seasons. In order to compare environmental parameters with the defined clusters, we applied the “mantel” function from the ade4 package (Dray and Dufour, 2007; Bougeard and Dray, 2018) with 10 000 permutations. Euclidean distance was used for the dissimilarity matrices of the environmental data and the phases. Missing values in environmental data resulted in deletion of samples before creation of the respective dissimilarity matrices for each analysis (see Supplementary Table S1). We calculated the Shannon and Simpson diversity indices using the “diversitycomp” command of the BiodiversityR package (Kindt and Coe, 2005). Hereby, we used the relative abundances and the seasons of each year and the four seasons combined as grouping variables. Additionally, we visualized the OTU intersection via the “upset” function of the UpSetR package (Conway *et al.*, 2017) for all OTUs per

season and for the 200 most abundant OTUs per season. Display of intersections was limited to at least 12 intersections for all OTUs and at least 5 intersections for 200 most abundant OTUs.

Predator–prey interactions and network analysis

As a starting point, we used previously observed predator–prey pairings (Table 1) to check for potential additional information on predator–prey dynamics and grazing impacts. Exemplary successions of main predators and prey were compared, and additional examples of predator–prey pairings were displayed.

Additionally, we inferred planktonic interactions by developing networks for each season that included at least 20 samples. Therefore, summer and autumn 2016 (five and one samples, respectively) and spring 2019 (three samples) were excluded from the analysis. Additionally, we created a network out of the 200 most abundant OTUs of all samples throughout all seasons. The OTU table was converted into relative abundances, prepared as a phyloseq object (McMurdie and Holmes, 2013), and the data were divided into subsets for each season. Due to the uneven number of samples and the high number of OTUs per season, which increased the computational complexity and duration of analysis, we tried to improve the comparison of the single networks by including the same amount of OTUs (200 most abundant each). We then used the method developed by Kurtz *et al*. (2015) called SParse InversE Covariance estimation for Ecological Association and Statistical Inference (SPIEC-EASI, version 1.1.0) for network construction. In contrast to other methods such as SparCC or Pearson/Spearman, which are based on empirical correlation or co-variance estimations, the SPIEC-EASI method uses the concept of conditional independence. Edges between any two OTUs (network nodes) therefore imply that there exists a relationship between the OTU abundances (association or interaction), which cannot be better explained by creating any other nodes in the network (Kurtz *et al*., 2015). The pipeline includes a data pre-processing and transformation step (clr transformation). The chosen graphical model was the Meinshausen–Buhlmann's neighbourhood selection, with lambda min ratio set to 1e-2, nlambda set to 20, and 999 repetitions. For visualization of the networks, the results were used to create igraph objects (Csardi and Nepusz, 2006) and plotted with the "plot_network" function of the phyloseq package. The edge width displays edge weights (strength of association). Therefore, positive and negative networks were plotted separately. Node sizes were set proportional to the abundances of the respective OTUs. Additionally, we created a network out of the 200 most abundant OTUs of all samples throughout all seasons using the same parameters as described for the seasonal networks.

Table 1. Potential predator–prey relationships as found by feeding experiments of exemplary taxa.

Predator	Prey or prey preferences	References
Calanus helgolandicus	*Chaetoceros pseudocurvisetus, Lauderia borealis, Gymnodinium splendens, Prorocentrum minimum, Skeletonema marinoi, Rhodomonas baltica*	(Paffenhöfer, 1970, 1971; Schnack, 1979; Lauritano *et al.*, 2011)
Centropages hamatus	Prefers ciliates over diatoms	(Saage *et al.*, 2009)
Gyrodinium dominans, G. spirale	*Prorocentrum minimum*	(Kim and Jeong, 2004)
Gyrodinium instriatum	Ciliates (*Favella* spp., *Eutintinnus tubulosus*)	(Uchida *et al.*, 1997)
Paracalanus sp.	Dinoflagellates (*Scrippsiella* sp., *Ceratium fusus*, *Gymnodinium* spp.), diatoms (*Skeletonema costatum*, *Chaetoceros lorenzianum*), ciliates	(Suzuki *et al.*, 1999)
Protoperidinium bipes	*Skeletonema costatum*	(Jeong *et al.*, 2004)
Protoperidinium pellucidum	Prefers diatoms over dinoflagellates (*Thalassiosira* sp., *Ditylum brightwellii*)	(Buskey, 1997)
Protoperidinium conicum, P. depressum, P. excentricum	*Ditylum brightwellii*	(Menden-Deuer *et al.*, 2005)
Protoperidinium cf. *divergens*	Copepod nauplii and eggs (*Acartia tonsa*)	(Jeong, 1994)

These networks may provide further insights into previously observed predator–prey interactions. We tested if new potential pairings can be described as they might be visible in the network, for example, due to the formation of subnetworks (small networks consisting of several OTUs, that are not connected to the main network) or the formation of clusters (OTUs of the same taxon that are connected in a main network). Hereby, we assume a positive association between prey and predators is caused by a bottom-up effect, since more prey might lead to more predators. Negative associations are assumed to portray a top-down effect as the higher predator occurrences would cause lower prey abundances.

Results

Successional patterns of different food web components

Especially during spring, we observed a shift from autotrophic Bacillariophyta to Prymnesiophyceae (haptophytes), Trebouxiophyceae, and Ulvophyceae (green algae). Maximum relative sequence abundances of several Bacillariophyta taxa were found in early spring and summer 2016, spring of 2017 and 2018. Highly abundant Bacillariophyta

included *Rhizosolenia*, *Thalassiosira*, *Coscinodiscus*, and *Ditylum*. While the two green macroalgal Ulvophyceae (*Ulva* and *Dilabifilum*) were most abundant in spring 2016, spring 2018, and summer 2018, the Prymnesiophyceae (mainly *Emiliania* and *Phaeocystis*) constituted more than 10% in spring of 2016 and 2018. The Trebouxiophyceae (*Picochlorum*) were found both in spring and summer 2018 (for more detailed information on succession of different taxa, see Supplementary Text).

Mixotrophic Dinophyceae shifted from genera such as *Gymnodinium* and *Heterocapsa* also occurring during spring to taxa such as *Alexandrium* in summer to *Akashiwo*, which mostly occurred in autumn. Other highly abundant genera included *Lepidodinium*, *Tripos*, and *Prorocentrum*, and *Chrysochromulina* sp. (Prymnesiophyceae).

Heterotroph microzooplankton was mostly represented by Dinophyceae (*Gyrodinium* sp.), which were most abundant in spring and summer of 2016, 2017, and 2018 and additionally in autumn 2018. *Gyrodinium* was one of the few genera, which occurred in every sample (Supplementary Table S2). Other heterotrophic microzooplanktons were Chrysophyceae, Spirotrichea, and Noctilucophyceae. Heterotrophic Chrysophyceae were mostly abundant in winters 2016/2017, spring 2018 and autumn 2018. Noctilucophyceae (*Noctiluca* sp.) was found in summers 2017 and 2018 and Spirotrichea (*Leegardiella* sp.) in spring 2018 (for more detailed information on succession of different taxa, see Supplementary Text).

The phylum Metazoa included 20 classes of potential meso- and macrozooplankton taxa, of which 130 genera were identified. The highest relative sequence abundances (i.e. above 10%) were found for different classes of worms (Annelida, Platyhelminthes, Nemertea), for fish (Craniata) as well as for Mollusca, Cnidaria, Ctenophora, and Brachiopoda. Arthropoda was found to be the most abundant class. Arthropoda were present in high relative sequence abundances during all seasons, except for autumn 2016. Brachiopoda (*Phoronis*) relative sequence abundances were especially high from autumn 2016 onwards until winter 2017/2018 and in summer 2018 (Supplementary Table S3, for more detailed information on succession of different taxa see Supplementary Text).

Comparison of planktonic predator occurrence

Most OTUs (out of 2790) represented Opisthokonta and Alveolata (Apicomplexa, Ciliophora, Dinoflagellata, and Perkinsea), 722 and 671 OTUs, respectively (Supplementary Figure 1a). These kingdoms had the highest relative sequence abundances, up to 89.5% and 87.0%, respectively (for details on the succession of other kingdoms see Supplementary Text).

Opisthokonta representatives included consumers such as Metazoa and Choanoflagellida. Decomposers such as Fungi, and Mesomycetozoa were also found (Table 2). Fungi were

highly abundant in summer 2018 with relative sequence abundances of up to 75% (mainly *Aspergillus* sp. and *Cryptococcus* sp., see Supplementary Tables S3 and S4), which were co-occurrent with several community shifts. Mesomycetozoa (including, e.g. the parasite *Pseudoperkinsus*) were most abundant during summers 2016 and 2018 at up to 18% relative abundance.

Table 2. Overview of presence and abundance of each phyla per kingdom, unclassified taxa on kingdom or phylum level are not included.

Kingdom	Phylum	Presence (max. 280 samples)	Total number of OTUs	Max. relative abundance in at least one sample (%)
Opisthokonta	Choanoflagellida	276	35	6.72
	Fungi	278	63	74.38
	Mesomycetozoa	279	16	17.99
	Metazoa	All	585	86.55
Alveolata	Apicomplexa	263	19	4.22
	Ciliophora	All	201	20.88
	Dinoflagellata	All	442	89.20
	Perkinsea	31	1	0.14
Stramenopiles	Ochrophyta	All	305	54.77
	Stramenopiles_X	All	228	40.49
Archaeplastida	Chlorophyta	All	79	58.91
	Streptophyta	160	14	3.00
Rhizaria	Cercozoa	All	420	15.85
	Radiolaria	211	7	3.54
Hacrobia	Centroheliozoa	257	11	3.49
	Cryptophyta	271	19	7.87
	Haptophyta	All	66	37.08
	Katablepharidophyta	268	6	2.40
	Picozoa	273	15	5.09
	Telonemia	262	22	6.01
Amoebozoa	Lobosa	247	27	3.25
Apusozoa	Apusomonadidae	211	11	0.43
	Hilomonadea	187	4	0.27
	Mantamonadidea	40	1	0.10

Out of the most abundant phyla, three groups of planktonic predators were identified: Dinoflagellata, Ciliophora, and Copepoda (Figure 2a and b). Planktonic prey included the autotrophic Bacillariophyta (Figure 2b) and additional highly abundant autotrophic and mixotrophic taxa (for further information see Supplementary Tables S3 and Supplementary Text). In most samples (198 out of 280), relative sequence abundances of Dinoflagellata (excluding Syndiniales, Supplementary Figure S2) were higher than those of Copepoda and Ciliophora. Copepoda and Dinoflagellata contrasted in relative sequence abundances. If Dinoflagellata relative sequence abundances were high, Copepoda relative sequence abundances were low and vice versa. In two samples (22 February 2018 and 08 March

2018), Ciliophora were more abundant than Dinoflagellata (excluding Syndiniales). In one sample (08 March 2018), Ciliophora had the highest abundance out of the three predator groups and relative sequence abundances of all three predator groups were below 5%. Higher relative sequence abundances of Ciliophora were not only found in occasional samples but throughout several samples of consecutive sampling days, in which Ciliophora had higher relative sequence abundances than Copepoda. This happened for example in May 2016, in May/June 2017, and in March 2018.

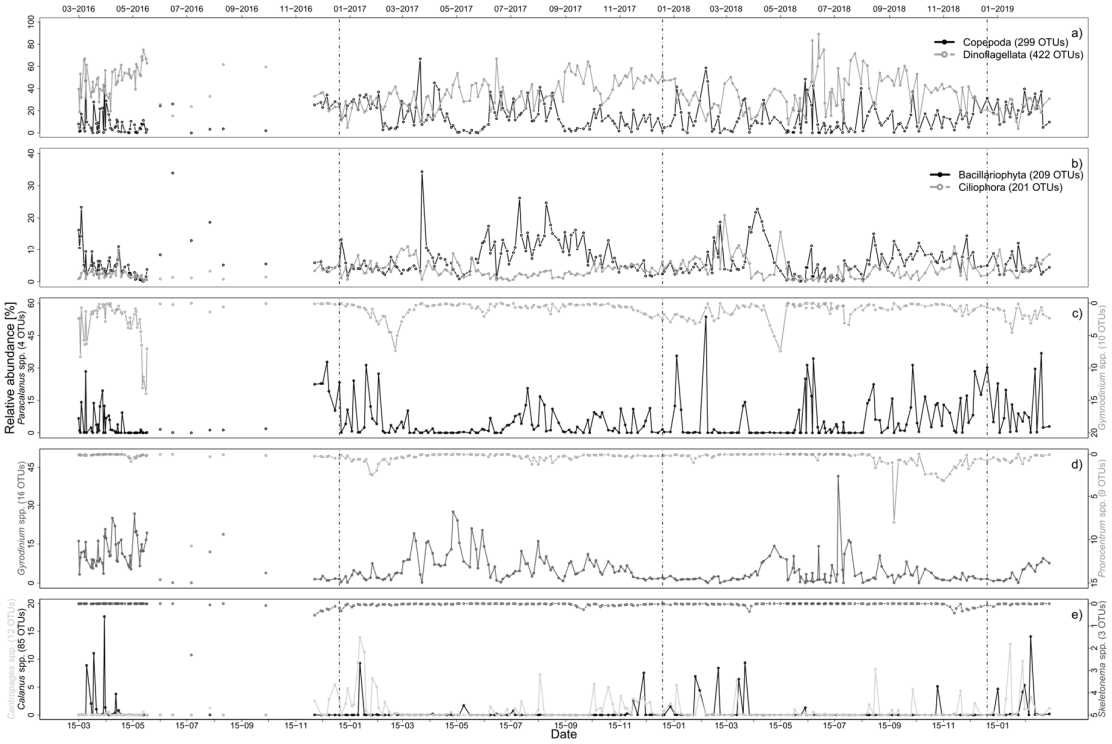


Figure 2. Relative abundance of (a) Copepoda and Dinoflagellata, (b) Bacillariophyta and Ciliophora, and of potential predators and prey combinations (c) *Paracalanus* spp. and *Gymnodinium* spp., (d) *Gyrodinium* spp. and *Prorocentrum* spp., and (e) *Calanus* spp., *Centropages* spp., and *Skeletonema* spp. from March 2016 to March 2019; perpendicular dotted lines mark the transition into a new year.

In general, Ciliophora were less abundant in relation to other predators, although they presented with a high diversity in OTUs (201). In comparison, over 74% of the crustacean OTUs (in total 339) belonged to only two genera (*Pseudocalanus* and *Calanus*) and most crustacean OTUs belonged to Copepoda (Supplementary Table S5). Over 45% of the Dinoflagellata OTUs (in total 442) were parasitic Syndiniales and over 26% remained unidentified and thus of unknown genus. Out of the 200 most abundant OTUs, except for spring 2018, Dinoflagellata always contributed the most OTUs. Comparison of the different spring communities revealed a much higher number of Dinoflagellata OTUs in 2016 than in 2017, 2018, and 2019. In autumn 2016 and autumn 2017, the second biggest contributor to the most abundant OTUs belonged to Ochrophyta, in all other seasons this was the case

for Metazoa or Ochrophyta, and Metazoa contributed the same number of OTUs. Additionally, during spring more OTUs of the class Cercozoa belonged to the 200 most abundant taxa than during other seasons.

Predator–prey interactions and network analysis

Potential links between predators and prey were investigated for the known predator–prey pairings (Table 1). In short, no clear connections between the known pairings were observed: several prey preferences are known for the genera *Protoperidinium* and some distinct species. In our data set three *Protoperidinium* OTUs were found in low relative sequence abundances (under 1%). Potential prey, such as *Skeletonema* (Figure 2), had much higher relative sequence abundances and therefore no clear relationships between predator and prey could be assumed. Additionally, even though *Thalassiosira* and *Ditylum* (both not displayed) were occasionally co-occurring with *Protoperidinium*, most peaks were not correlated with the predator. Due to the low relative sequence abundances, no *Protoperidinium* OTU was part of the network analysis.

Similarly, for copepod predators, such as *Paracalanus*, *Calanus*, or *Centropages*, no clear connection to potential prey could be found (Figure 2). For example, *Chaetoceros* OTUs were peaking either when *Calanus* sp. was absent or peaked after the occurrence of predators. However, BLAST alignment did not reveal any *Chaetoceros pseudocurvisetus* OTU, which was known as prey for *Calanus* (Table 1). Other potential prey of *Calanus* and *Paracalanus* was too low in abundance (*Lauderia* sp.), only present when the predator was absent (*Skeletonema*, *Rhodomonas*) or co-occurring and alternating in peaks (*Gymnodinium* sp., *Prorocentrum* sp.) without any distinct patterns. Instead, all investigated relationships revealed high variabilities in occurrences and relative sequence abundances.

We then performed a network analysis to inspect the interconnectivity of the food web during the different seasons and for identification of potential separate networks in the food web. Networks were thus constructed with the 200 most abundant OTUs (based on relative abundance) of each season (see the section Community diversity on comparison of the 200 most abundant OTUs).

We found that the different taxa were highly interconnected with each other in each season (Figures 3 and 4), as OTUs were positively associated with other OTUs of all kinds of taxa at high frequency in one network. A similar structure was also found for the positive association network, which included the 200 most abundant OTUs of all samples (Figure 5). However, interconnections were varying greatly in strength (thickness of the edges) (Figures 3–5, a list of all associations can be found in Supplementary Table S6). Strength of the association was not associated with abundance of the OTUs, as strong associations

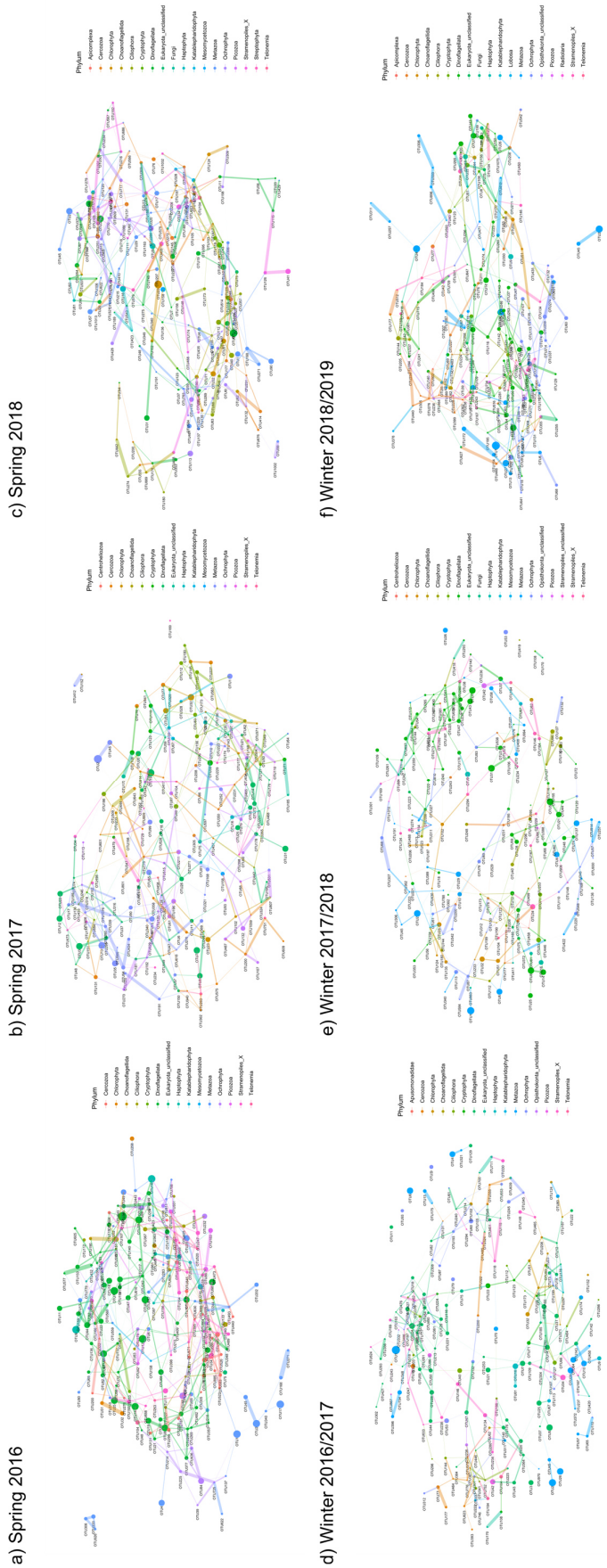


Figure 3. Co-occurrence networks per season (a) spring 2016, (b) spring 2017, (c) spring 2018, (d) winter 2016/2017, (e) winter 2017/2018 and (f) winter 2018/2018; displayed are the 200 most abundant OTUs per season, each season included at least 20 samples, colour code shows the associated phylum for each OTU. Only positive edges were plotted, and the width of the edges displays edge weights. Node sizes were set proportional to the relative sequence abundances of the respective OTUs.

were also found between OTUs of high and low relative sequence abundances. Some networks revealed small subnetworks of up to 3 OTUs, which were only positively associated with each other but not to the rest of the main network. These subnetworks did not reveal any separate food web connections; instead subnetworks rather consisted of OTUs of the same taxa, which hinted at these taxa sporadically occurring in high relative sequence abundances. Some OTUs were not found to be positively associated with any other OTU. These included especially OTUs of Opisthokonta, namely of the phyla Metazoa and Fungi, and some Dinoflagellata.

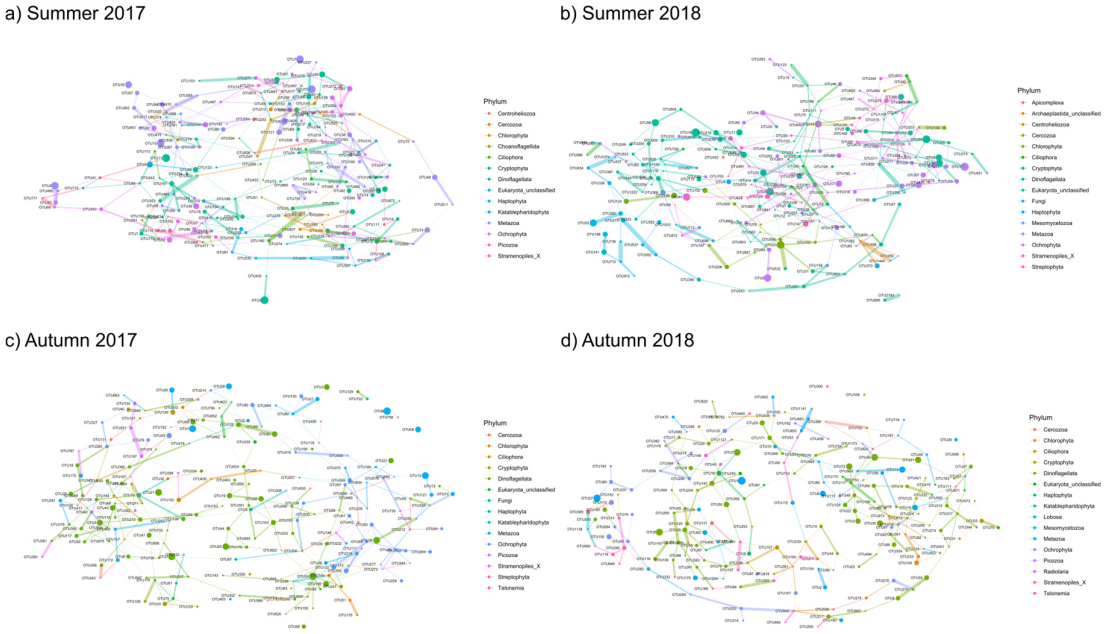


Figure 4. Co-occurrence networks per season (a) summer 2017 (b) summer 2018, (c) autumn 2017 and (d) autumn 2018; displayed are the 200 most abundant OTUs per season, each season included at least 20 samples, colour code shows the associated phylum for each OTU. Only positive edges were plotted, and the width of the edges displays edge weights. Node sizes were set proportional to the relative sequence abundances of the respective OTUs.

For negative association networks (Supplementary Figure S3) fewer links between different OTUs were found. In spring 2016, the most complex network of negative associations was detected. Overall, most negative associations were found between two single OTUs. Associations occurred between different phyla, but also within single phyla. For example, in winter 2017/2018 six Dinoflagellata OTUs formed a subnetwork, including two parasitic Syndiniales OTUs and four Dinophyceae OTUs.

The network including positive associations throughout all seasons of all samples revealed two subnetworks, which consisted of two OTUs each, as well as three OTUs that were not connected to any other OTU (Figure 5). One subnetwork consisted of *Aspergillus* (OTU 52 and 141), the other subnetwork consisted of *Temora* sp. (OTU 2 and 45). The three

separate OTUs all belonged to different Metazoa: OTU 13 (Ctenophora), OTU 29 (*Hiatella* sp.), and OTU 191 (*Metridium* sp.). A cluster of eight interconnected Ciliophora OTUs (Choreotrichida and Strombidiida) was found in the main network. This cluster was connected to several other taxa, most connections belonged to OTUs of Chlorophyta, Dinoflagellata, Cryptophyta, and Cercozoa. Additionally, Dinoflagellata OTUs were highly interconnected to each other in the main network and several clusters were formed. For example, several OTUs of Dinophyceae (OTU 21, 24, 30, 36, 69, 72, 85, 117, 129) and Syndiniales (OTU 54, 100, 163) were highly interconnected, but also connected to further Dinoflagellata OTUs, as well as to other taxa such as Metazoa, Ochrophyta, Chlorophyta, and Ciliophora.

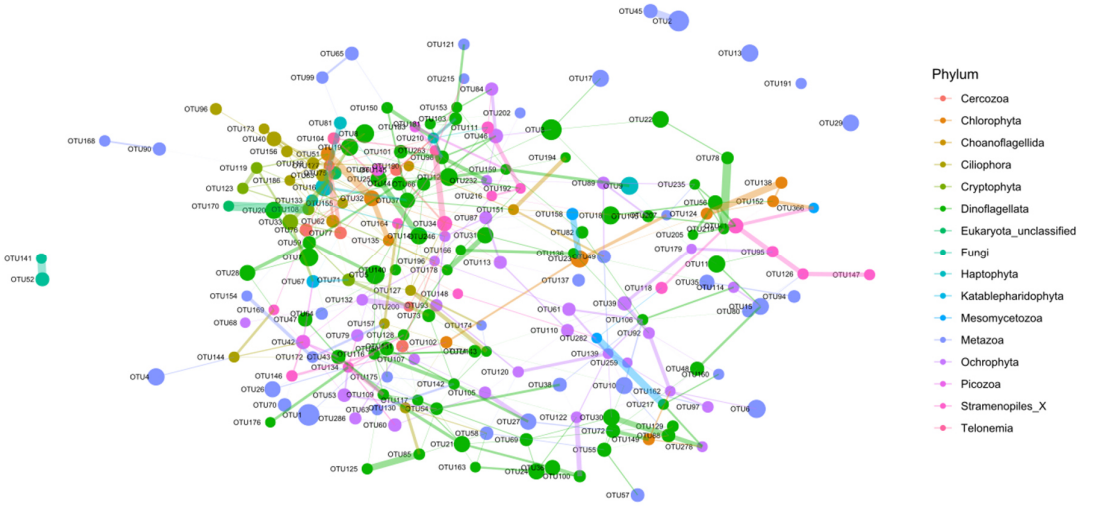


Figure 5. Co-occurrence network of all samples; displayed are the 200 most abundant OTUs, colour code shows the associated phylum for each OTU. Only positive edges were plotted, and the width of the edges displays edge weights. Node sizes were set proportional to the relative sequence abundances of the respective OTUs.

Analysis of network connections of OTUs belonging to the observed predator–prey interactions as listed in Table 1 revealed complex structures and a variety of potential predator–prey constellations.

For *Paracalanus* network analysis revealed positive and negative associations with 10 different phyla: Cercozoa, Chlorophyta, Ciliophora, Dinoflagellata, Fungi, Haptophyta, Metazoa, Ochrophyta, Picozoa, and Stramenopiles_X. Most associations (positive and negative) were found for different Dinoflagellates, including a weak negative association to *Gymnodinium* (OTU14) in spring 2016 and the strongest positive association to *Heterocapsa* sp. (OTU 8) in summer 2018. A positive association to diatom OTUs was found during spring 2016 (OTU 725) and 2018 (OTU68), a weak negative association in autumn

2018 (OTU 107 and 314). Three ciliate OTUs in three different seasons each showed positive (OTU 563, OTU 203) or negative (OTU40) associations.

Calanus OTUs were only present in networks from winter and spring. No negative associations were found. For positive associations, *Calanus* OTUs were mostly interconnected with other *Calanus* OTUs. In total, positive associations were found for six phyla: Ciliophora, Cryptophyta, Dinoflagellata, unclassified Eukaryota, Metazoa, and Ochrophyta. Potential prey as observed before (Table 1) did not show any connections. The only association to a diatom OTU (OTU 307) was found in the winter 2017/2018 network, displaying the strongest connection besides the interconnections of different *Calanus* OTUs. In winter 2018/2019, OTU 123 that was identified as Cryptomonadales was connected to two *Calanus* OTUs (OTU 333 and OTU 593). In spring 2018, a connection to OTU 306 (unclassified Gymnodiniales) and an even stronger connection to a Chrysophyceae OTU (OTU 844) were found.

Centropages was associated with 7 different phyla: Cercozoa, Choanoflagellida, Ciliophora, Dinoflagellata, Metazoa, Ochrophyta, and Stramenopiles_X. Most associations were positive, only one negative association to a ciliate (OTU 248) was found in winter 2017/2018. As indicated by Table 1, Centropages seems to prefer ciliates over diatoms, as more connections to different ciliate OTUs were found compared to connections to diatoms. In terms of strength of association, connections to ciliates were stronger in 2017 compared to 2018. Positive associations of ciliates were found in summer 2017, autumn 2017, and autumn 2018, for OTU 497, OTU 130, and OTU 590, respectively. A positive association to diatom OTU 487 was found in spring 2018, a weaker association to diatom OTU 84 in summer 2018.

Community diversity

Community composition significantly changed from one season to the next during all three seasonal cycles, while the communities of the individual seasons observed in the different years were highly similar. This is reflected in the PCA plot (Figure 6) showing that samples from the same season rather than samples from the same year cluster together. Samples collected in spring 2017 showed the highest inter-sample variability compared to other years, whereas spring samples were, in general, more similar to each other than autumn samples. The two outliers of the PCA plot from the 01 February 2018 (winter 2017/2018) and 18 December 2018 (winter 2018/2019) can be explained by the heterogeneity of library sizes, since these two samples consisted of a bigger library than all other samples.

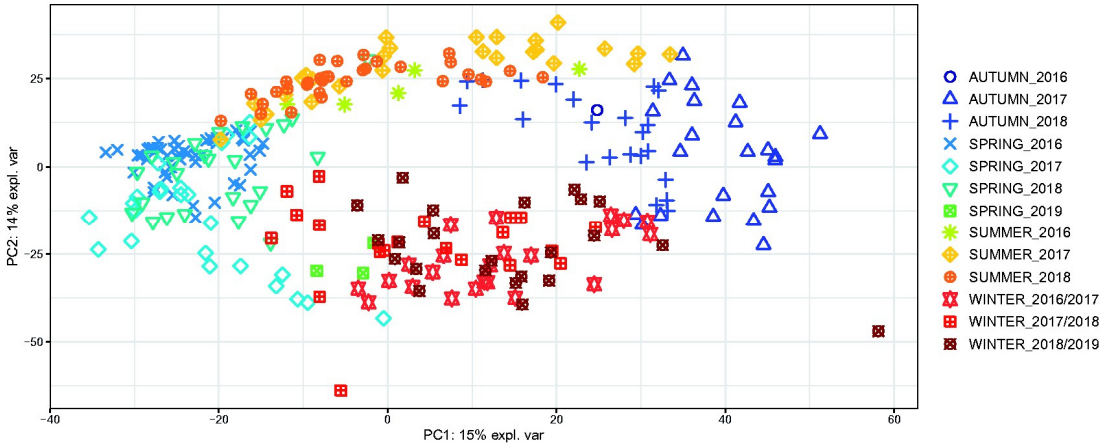


Figure 6. PCA plot with colour coded samples according to the respective season/year.

In most years, the diversity of the plankton communities was higher in autumn and winter than in spring and summer, while summer displayed the lowest diversity. This is reflected by both, OTU numbers (Figure 7, Supplementary Figure S4) and diversity indices (Tables 3 and 4), whereas Shannon and Simpson diversity indices were not differing much in general. Sampling intensity for the different seasons of the different years was variable due to logistic constraints. For most of the seasons, more samples collected for a respective season did not result in a higher diversity (sample size above 20). For example, even though the number of samples taken in spring 2016 was nearly twice as much as in spring 2017 and spring 2018, 49 samples compared to 25 samples each, the number of OTUs was similar, 2098, 2092, and 2016 OTUs, respectively. Considering Shannon and Simpson indices, species composition of autumn and winter samples was most diverse, while spring diversity was lower and summer communities had the lowest diversity (Table 3). However, the subset of three seasons with the fewest number of samples, also had the least number of OTUs as well as the lowest Shannon and Simpson diversity, whereas both indices resulted in high values in general (Table 4). Only one sample with 834 OTUs in total was taken in autumn 2016, the five samples of summer 2016 consisted of 1638 OTUs, and spring 2019 (three samples) consisted of 1276 OTUs (Figure 7).

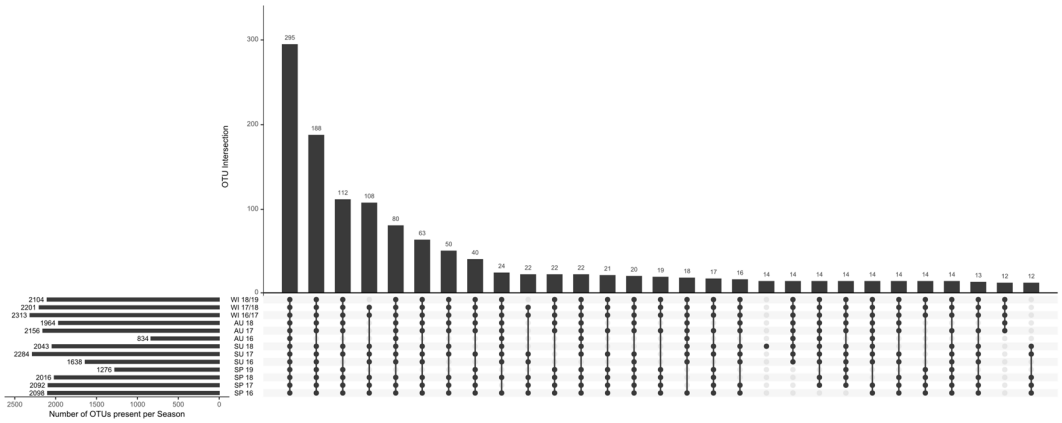


Figure 7. Total number of OTUs present during each season and comparison of shared OTUs, OTU intersection displays the number of shared OTUs; SP, Spring; SU, Summer; AU, Autumn; WI, Winter, display of intersections was limited to at least 12 intersections.

Table 3. Shannon and Simpson diversity indices for four seasons combined.

Season	Number of samples	Shannon	Simpson
Spring	102	5.338	0.9855
Summer	65	5.167	0.9831
Autumn	46	5.443	0.9869
Winter	67	5.375	0.9810

Table 4. Shannon and Simpson diversity indices for each season per year.

Season_Year	Number of samples	Shannon	Simpson
Spring_2016	49	4.822	0.9769
Spring_2017	25	5.084	0.9791
Spring_2018	25	5.115	0.9803
Spring_2019	3	4.243	0.9546
Summer_2016	5	4.637	0.9722
Summer_2017	27	4.997	0.9783
Summer_2018	33	4.761	0.9776
Autumn_2016	1	3.661	0.9005
Autumn_2017	23	5.223	0.9844
Autumn_2018	22	5.309	0.9842
Winter_2016/2017	23	5.114	0.9754
Winter_2017/2018	21	5.235	0.9838
Winter_2018/2019	23	5.102	0.9759

Out of the 200 most abundant OTUs (Supplementary Figure S4), except for spring 2018, Dinoflagellata always contributed the most OTUs. Comparison of the different spring communities revealed a much higher number of Dinoflagellata OTUs in 2016 than in 2017, 2018, and 2019. Ochrophyta contributed the second most OTUs in autumn 2016 and autumn 2017. In all other seasons, this was the case for Metazoa or both, Ochrophyta and

Metazoa, contributed the same number of OTUs. Additionally, during spring more OTUs of the class Cercozoa belonged to the 200 most abundant taxa than during other seasons.

In total, our data set for analysis consisted of 2790 OTUs. Between 2104 and 2313 OTUs were found during winter. Alpha diversity in autumn was between 834 and 2156 OTUs. Out of all OTUs, 295 OTUs (10.6%) were present throughout all seasons (Figure 7). Furthermore, 168 OTUs were shared by all seasons, except for autumn 2016, which were sampled just once. Only 13 OTUs that belonged to the most abundant OTUs were present in all seasons (Supplementary Figure S4). A higher number of unique OTUs (Figure 7) and the highest proportion of unique OTUs of the 200 most abundant OTUs (Supplementary Figure S4) indicated that the community of summer 2018 was different from other years and seasons. The beta diversity analysis further indicated that spring and summer 2018 were different from other years, as the samples from these seasons resulted in several significant small clusters (Supplementary Figures S5 and S6, more details in Supplementary Text, $R^2 = 0.69727$, $F = 16.597$, $p < 0.0001$). PERMANOVA also confirmed that the communities occurring during each season (13 seasons, $R^2 = 0.47509$, $F = 20.138$ $p < 0.0001$) and the OTUs of the four seasons across all years ($R^2 = 0.31119$, $F = 41.564$, $p < 0.0001$) were significantly different. While this difference was also caused by the different environmental conditions, the Mantel test revealed a significant but mostly weak correlation of the metabarcoding data set with several environmental factors: temperature, nitrate, sunshine duration, salinity, and Secchi depth (see Supplementary Table S1).

Discussion

Our 3-year metabarcoding study revealed a highly variable system, in which blooms of single prey taxa are only occurring occasionally and without any distinct pattern, whereas potential predators are found in high relative sequence abundances throughout. While the overall abundance of Bacillariophyta was highest in spring followed by summer and autumn, certain genera did not bloom during every year. Throughout the years, similar findings on changes in abundances of diatoms, but also of shifts in occurrence were recorded (Scharfe and Wiltshire, 2019).

In the following paragraphs, we, therefore, want to discuss the following important results: We found a very strong inter-annual variation in the algae, but not in the grazers and existing predator–prey relationships could not be found in the metabarcoding data. Instead, our networks show very strong connections with many nodes, indicating that the webs are probably very robust, and the predators seem to be able to shift without any problems from one prey item to another.

Community diversity

Using water samples, we detected a high diversity of taxa, ranging from potential meroplankton, such as fish and other Metazoa, to holoplanktonic mesoplankton down to the smallest picoplankton and parasites. Parasitic taxa, which mostly are represented by marine parasitoids, hereby can include free-living stages, but also parasitoids currently infecting other plankton (Käse *et al*., 2021).

Intensive sampling revealed a high diversity and likely increased the probability of catching rarer taxa, as confirmed by both, Shannon and Simpson, diversity indices. However, a maximum diversity was reached as more sampling did not result in a higher diversity, e.g. when comparing sampling efforts during different spring seasons.

Furthermore, despite spring blooms being frequently monitored, spring revealed less OTUs than other seasons, especially compared to winter and resulted in lower diversity indices. Instead, autumn was the most diverse season. Comparison of OTUs per season revealed a steady background community consisting of various taxa (295 OTUs), which were sampled during every season, despite the strong seasonality at the sampling site. The main taxa on phylum level were also, with few exceptions, present nearly all the time. A diverse background community of protists was reported in spring 2016 (Käse *et al*., 2020), which therefore can now be extrapolated to other seasons and includes also bigger sized zooplankton.

However, the presence of a seemingly stable background community, which has now been shown in two studies, does not mean considerable fluctuations such as blooms of unusual species are not possible. This is exemplified by the year 2018. This year was unique in terms of community composition, with summer samples, in particular, differing markedly as seen in the amount of unique OTUs. The occurrence of fungi, *Picochlorum* and *Dilabifilum* hints at a community of the algae and lichen-forming fungi which was previously observed for different Ulvophyceae and lichen-forming fungi (Nelsen *et al*., 2011; Thüs *et al*., 2011). In addition, benthic taxa, such as worms, were occasionally found in high relative sequence abundances. There are several potential explanations for this. The most parsimonious is that due to the relatively shallow sampling site, combined with the strong tidal influence, and the occasional storm, material, and organisms were suspended into the water column and sampled in our water samples. However, it is also possible that these species were sampled in their (mero-)planktonic state instead of the adult state, which is indistinguishable by metabarcoding (Bucklin *et al*., 2016).

Predator–prey interactions and network analysis

We did not observe any strong support of the predator–prey pairs that we know, nor did we find any close connections to others. One reason for this might be our observation that even though the prey communities rapidly change and do not re-occur with the same species from 1 year to the next, this is completely different for predators. Hence, a strong link between individual taxa cannot be expected.

Dinoflagellates made the highest contribution to the community, most likely, playing a key role not only as a predator but also as potential prey for bigger sized taxa. A general bias of our primers in favour of dinoflagellates is unlikely, even though dinoflagellates have a higher copy number, as Sprong *et al*. (2020) showed that more coastal stations were not dinoflagellate dominated using the same primers compared to our sampling area. Similar results were found in the Estuary of Bilbao, where high relative sequence abundances of copepods were found in larger size fractions, and no dominance of dinoflagellates even in small size fractions was seen (Abad *et al*., 2017). Therefore, our results were most likely a reflection of the unique ecology at our sampling point and not caused by a bias in the molecular method.

The high relative sequence abundances of heterotrophic dinoflagellates are supported by a previous study, using traditional microscopy, which also detected high contribution of heterotrophic dinoflagellates in biomass (Löder *et al*., 2012). Our observation of ciliates peaking during spring is also supported by previous studies (Löder *et al*., 2011, 2012), which found ciliates to be an important but highly selective grazer in spring. Besides the possibility of methodological constraints regarding ciliate detection, this specificity could explain the low relative sequence abundances in our assemblage during certain years. The variability of diatom occurrences as prey might be reflected by the grazer relative sequence abundances as well. However, a potential cluster of ciliates was found by network analysis, and a variety of taxa were associated with this cluster, indicating several potential predator–prey relationships. Furthermore, it needs to be noted that bacteria, which were not part of this study, are known as an additional important prey option for ciliates (Albright *et al*., 1987; Sherr and Sherr, 1987).

The copepods *Paracalanus* and *Centropages*, which are able to feed on ciliates (Suzuki *et al*., 1999; Saage *et al*., 2009), were associated with ciliates in the network analysis. However, in general, connections were weak, and stronger connections to other taxa were found. In contrast to ciliates, crustacean OTUs peaked every year no matter which potential food was present. While copepods can feed selectively on microzooplankton (Nejstgaard *et al*., 2001; Löder *et al*., 2011), they are often omnivorous, feeding on a diverse range of organisms

such as diatoms, dinoflagellates, and other zooplankton including their own eggs and larvae (e.g. Calbet and Saiz, 2005; Boersma *et al*., 2014; Yeh *et al*., 2020).

The difficulty of identifying predator–prey dynamics could be explained by potential coexistence within different plankton communities due to different feeding and motility traits as it has been shown, e.g. for the western English Channel, (Kenitz *et al*., 2017). The authors also linked seasonal succession in the community, besides the interannual variation in dominant species, to a succession of activity traits. Furthermore, Kenitz *et al*. (2017) suggested that strong turbulence benefit passive feeding zooplankton, and leads to an enhanced grazing pressure on motile prey, which would benefit the growth of non-motile prey. In our data set, *Oithona*, as a passive feeder, occurred in high relative sequence abundances in summer but was also abundant during other seasons and we observed additional blooms of non-motile prey during summer months, which might have benefitted from decreased grazing pressure.

Similar to the differences in copy numbers for dinoflagellates, a large proportion of DNA of multi-celled zooplankton such as copepods or fish might indicate a bias of these taxa in the metabarcoding assemblage. However, in our approach a dominance of these taxa was not evident, as generally, dinoflagellates were most abundant, nor did high relative sequence abundances prevent the formation of seasonal patterns of small-sized plankton. This further indicates that the data on these taxa was mostly based on DNA, which was not retrieved through extraction of whole organisms or their extremities, but rather from environmental DNA. Due to continuously high relative sequence abundances and diversity of the predator and prey as well as the complexity of the food web, a distinct grazing impact on single taxa or distinct links between potential predator–prey pairings could not be extracted from our data set. While it is known, that especially zooplankton biomass responds in much longer time scales (Wiltshire *et al*., 2015), e.g. due to the complexity of the metazoan life stages, we pose that co-occurrence might already be hinting at a potential relationship. For example, *Calanus* might prey on *Chaetoceros* or peaks of predators might indicate that prey is eaten up and therefore no longer detected. It could also be the case that peaks of predators are caused by feeding on other prey instead, which makes it difficult to define distinct predator–prey relationships. However, other investigated potential distinct predator–prey relationships could also not be observed.

In general, network analysis of the 200 most abundant OTUs revealed Metazoa OTUs, whose respective organisms would be bigger in size than the rest of the sampled plankton community, were not as tightly connected to the rest of the network. This might be the case, especially if potential consumers are highly abundant on rare occasions. In general, these

networks could be a potential tool for detection of specific relationships. However, the networks were tightly interconnected and only few OTUs did not connect to the main seasonal networks. Links between edges of the same taxonomic group have been observed before (Faust *et al*., 2012; Kurtz *et al*., 2015) and are commonly described as “assortativity.” These associations could also be detected because occasionally two OTUs represent the same species but could not be merged to a single OTU because of potential errors in the sequence or other biases. Additionally, the risk of spurious coincidences might be increased in the network analysis and associations might be depicted by chance. This is why interpretation needs to be careful and additional analyses are necessary to verify potential associations. As each season also comprises several communities as depicted by the clusters in the beta diversity, the networks might also include associations in between these different communities. Especially since the PCA analysis also indicated that few samples might rather belong to a previous or follow-up season, more or other associations might have been found without the focus on the seasonality. Based on these complex networks, clear dynamics cannot be identified easily and interactions between food web components seem to occur on much bigger scales. The tight links in between various components of the food web emphasize that they are co-occurring throughout the different seasons and indicate a high variability in food options for potential consumers.

Previous investigations of copepod faecal pellets (Turner, 1984) and metabarcoding of the gut content (Yeh *et al*., 2020) revealed that copepods ingest a wide variety of food items. Besides this high variability in ingested food, combining the known predator–prey pairings of previous grazing experiments also demonstrates this high variability. The selective feeding on certain taxa has mostly been observed in grazing experiments, where potential prey is limited to certain taxa and provided constantly, whereas the complex hydrography at Helgoland might disturb the actual formation of a system sufficiently stable for allowing specific predator–prey relationships. Therefore, it is possible that predators are more flexible and less selective in natural environments than in experiments, but they are also provided with a higher choice and variability of prey options. Alternatively, we are not able to distinguish existing predator–prey relationships as the high variability in the system conceals explicit dynamics.

The fact that we found a background community in addition to the very variable occurrence of taxa, which might include mostly opportunistic species, hints at a rather flexible but nevertheless stable and healthy ecosystem. While shifts of species occurrence due to environmental change were observed at Helgoland (Scharfe and Wiltshire, 2019), stability of the ecosystem might be high enough due to the natural fluctuation and high adaptability of the system. The tight links in the seasonal networks indicate furthermore, that the

robustness of the food web to species loss is potentially quite high (Dunne *et al*., 2002; Estrada, 2007), which is also supported by the random occurrence of taxa throughout the years.

Conclusion

Metabarcoding of water samples is suitable for capturing taxa of the whole community and for obtaining information on plankton succession in relation to time and environmental conditions, without exclusion of large-sized taxa from the analysis. Therefore, new technologies, like next-generation sequencing, may be used in addition to traditional methodologies on a long-term basis. Comparability and practicability of combining these different methods still need to be tested in future studies. The system is characterized by a high variability of potential prey or predators, which are not necessarily co-occurring or displaying typical patterns. Predator–prey relationships in the planktonic community are diverse and plentiful and specialist relationships are rather uncommon. Instead, generalists seem to be the norm, which makes it difficult to extract distinct predator–prey dynamics in the field. This offers an enormous potential of relationships in the plankton community that might be verified by traditional laboratory experiments. Furthermore, it remains under question how the resilience of the North Sea is influenced by the high variability.

Supplementary data

Supplementary material is available at the *ICESJMS* online version of the manuscript.

Data availability

Sequence data for this study have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession number PRJEB37135, using the data brokerage service of the German Federation for Biological Data (GFBio, Diepenbroek *et al*., 2014), in compliance with the Minimal Information about any (X) Sequence (MIxS) standard (Yilmaz *et al*., 2011). Details of our pipeline are available on GitHub (https://github.com/PyoneerO/qzip). Additionally, the full OTU table (280 samples with 59.284 OTUs) was archived in PANGAEA (DOI: 10.1594/PANGAEA.921026). All other relevant data are within the manuscript and its Supplementary material files.

Käse, L., Metfies, K., Kraberg, A. C., Neuhaus, S., Meunier, C. L., Wiltshire, K. H., Boersma, M. Metabarcoding analysis suggests that flexible food web interactions in the eukaryotic plankton community are more common than specific predator–prey relationships at Helgoland Roads, North Sea. – ICES Journal of Marine Science,00: 000–000.

Acknowledgements

We gratefully acknowledge the crew of the research vessel “Aade,” our technicians Kristine Carstens and Silvia Peters for collecting all LTER data and the MPI sampling team, especially Bernhard Fuchs, for sharing their samples. Additionally, we want to thank Swantje Rogge, Kerstin Oetjen, and Pim Sprong for their support of the sequencing procedure.

Funding

This work was supported by the Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung. We further acknowledge support by the Open Access Publication Funds of Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung.

References

Abad D., Albaina A., Aguirre M., Laza-Martínez A., Uriarte I., Iriarte A., Villate F. *et al.* 2016. Is metabarcoding suitable for estuarine plankton monitoring? A comparative study with microscopy. *Marine Biology*, 163: 1 – 13.

Abad D., Albaina A., Aguirre M., Estonba A. 2017. 18S V9 metabarcoding correctly depicts plankton estuarine community drivers. *Marine Ecology Progress Series*, 584: 31 – 43.

Albright L., Sherr E., Sherr B., Fallon R. 1987. Grazing of ciliated protozoa on free and particle-attached bacteria. *Marine Ecology Progress Series*, 38: 125 – 129.

Blanco-Bercial L. 2020. Metabarcoding analyses and seasonality of the zooplankton community at BATS. *Frontiers in Marine Science*, 7: 1 – 16.

Boersma M., Wesche A., Hirche H. J. 2014. Predation of calanoid copepods on their own and other copepods’ offspring. *Marine Biology*, 161: 733 – 743.

Boersma M., Wiltshire K. H., Kong S. M., Greve W., Renz J. 2015. Long-term change in the copepod community in the southern German Bight. *Journal of Sea Research*, 101: 41 – 50.

Bougeard S., Dray S. 2018. Supervised multiblock analysis in R with the ade4 package. *Journal of Statistical Software*, 86: 1 – 17.

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

Bucklin A., Yeh H. D., Questel J. M., Richardson D. E., Reese B., Copley N. J., Wiebe P. H. 2019. Time-series metabarcoding analysis of zooplankton diversity of the NW Atlantic continental shelf. *ICES Journal of Marine Science*, 76: 1162 – 1176.

Buskey E. 1997. Behavioral components of feeding selectivity of the heterotrophic dinoflagellate *Protoperidinium pellucidum*. *Marine Ecology Progress Series*, 153: 77 – 89.

Calbet A., Saiz E. 2005. The ciliate-copepod link in marine ecosystems. *Aquatic Microbial Ecology*, 38: 157 – 167.

Callies U., Scharfe M. 2015. Mean spring conditions at Helgoland Roads, North Sea: graphical modeling of the influence of hydro-climatic forcing and Elbe River discharge. *Journal of Sea Research*, 101: 1 – 11.

Chafee M., Fernàndez-Guerra A., Buttigieg P. L., Gerdts G., Eren A. M., Teeling H., Amann R. I. 2018. Recurrent patterns of microdiversity in a temperate coastal marine environment. *ISME Journal*, 12: 237 – 252.

Conway J. R., Lex A., Gehlenborg N. 2017. UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics*, 33: 2938 – 2940.

Csardi G., Nepusz T. 2006. The igraph software package for complex network research. *InterJournal, Complex Sy*, 1965: http://igraph.org (last accessed 30 August 2020).

Diepenbroek M., Glöckner F., Grobe P., Güntsch A., Huber R., König-Ries B., Kostadinov I. *et al.* 2014. Towards an integrated biodiversity and ecological research data management and archiving platform: the German Federation for the Curation of Biological Data (GFBio). *In Informatik 2014 – Big Data Komplexität Meistern. GI-Edition: Lecture Notes in Informatics (LNI) – Proceedings*, Vol. 232, Ed. by Plödereder E, Grunske L, Schneider E, Ull D. Köllen Verlag, Bonn. pp. 1711 – 1724.

Dray S., Dufour A. B. 2007. The ade4 package: implementing the duality diagram for ecologists. *Journal of Statistical Software*, 22: 1 – 20.

Dunne J. A., Williams R. J., Martinez N. D. 2002. Network structure and biodiversity loss in food webs: robustness increases with connectance. *Ecology Letters*, 5: 558 – 567.

DWD Climate Data Center (CDC). 2019. Daily station observations of sunshine duration in hours for Germany, version v19.3. http://www.dwd.de/cdc (last accessed 24 August 2020).

Estrada E. 2007. Food webs robustness to biodiversity loss: the roles of connectance, expansibility and degree distribution. *Journal of Theoretical Biology*, 244: 296 – 307.

Fadeev E., Salter I., Schourup-Kristensen V., Nöthig E.-M., Metfies K., Engel A., Piontek J. *et al.* 2018. Microbial communities in the east and west fram strait during sea ice melting season. *Frontiers in Marine Science*, 5: 1 – 21.

Faust K., Sathirapongsasuti J. F., Izard J., Segata N., Gevers D., Raes J., Huttenhower C. 2012. Microbial co-occurrence relationships in the Human Microbiome. *PLoS Computational Biology*, 8: e1002606.

Giner C. R., Balagué V., Krabberød A. K., Ferrera I., Reñé A., Garcés E., Gasol J. M. *et al.* 2019. Quantifying long-term recurrence in planktonic microbial eukaryotes. *Molecular Ecology*, 28: 923 – 935.

Grasshoff K. 1976. *Methods of Seawater Analysis*. Verlag Chemie, Weinheim and New York. 317 pp.

Guillou L., Bachar D., Audic S., Bass D., Berney C., Bittner L., Boutte C. *et al.* 2013. The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. *Nucleic Acids Research*, 41: D597 – D604.

Hernández-Ruiz M., Barber-Lluch E., Prieto A., Álvarez-Salgado X. A., Logares R., Teira E. 2018. Seasonal succession of small planktonic eukaryotes inhabiting surface waters of a coastal upwelling system. *Environmental Microbiology*, 20: 2955 – 2973.

Hickel W., Mangelsdorf P., Berg J. 1993. The human impact in the German Bight: eutrophication during three decades (1962-1991). *Helgoländer Meeresuntersuchungen*, 47: 243 – 263.

Hickel W. 1998. Temporal variability of micro- and nanoplankton in the German Bight in relation to hydrographic structure and nutrient changes. *ICES Journal of Marine Science*, 55: 600 – 609.

Hirai J., Kuriyama M., Ichikawa T., Hidaka K., Tsuda A. 2015. A metagenetic approach for revealing community structure of marine planktonic copepods. *Molecular Ecology Resources*, 15: 68 – 80.

Hirai J., Tsuda A. 2015. Metagenetic community analysis of epipelagic planktonic copepods in the tropical and subtropical Pacific. *Marine Ecology Progress Series*, 534: 65 – 78.

Jeong H. 1994. Predation by the heterotrophic dinoflagellate *Protoperidinium* cf. *divergens* on copepod eggs and early naupliar stages. *Marine Ecology Progress Series*, 114: 203 – 208.

Jeong H., Yoo Y., Kim S., Kang N. 2004. Feeding by the heterotrophic dinoflagellate *Protoperidinium bipes* on the diatom *Skeletonema costatum*. *Aquatic Microbial Ecology*, 36: 171 – 179.

Käse L., Kraberg A. C., Metfies K., Neuhaus S., Sprong P. A. A., Fuchs B. M., Boersma M. *et al.* 2020. Rapid succession drives spring community dynamics of small protists at Helgoland Roads, North Sea. *Journal of Plankton Research*, 42: 305 – 319.

Käse L., Metfies K., Neuhaus S., Boersma M., Wiltshire K. H., Kraberg A. C. 2021. Host-parasitoid associations in marine planktonic time series: can metabarcoding help reveal them? *PLoS One*, 16: e0244817.

Kenitz K. M., Visser A. W., Mariani P., Andersen K. H. 2017. Seasonal succession in zooplankton feeding traits reveals trophic trait coupling. *Limnology and Oceanography*, 62: 1184 – 1197.

Kim J., Jeong H. 2004. Feeding by the heterotrophic dinoflagellates *Gyrodinium dominans* and *G. spirale* on the red-tide dinoflagellate *Prorocentrum minimum*. *Marine Ecology Progress Series*, 280: 85 – 94.

Kindt R., Coe R. 2005. Tree diversity analysis. A manual and software for common statistical methods for ecological and biodiversity studies. World Agroforestry Centre (ICRAF), Nairobi (Kenya).

Knefelkamp B. 2009. *An Investigation of the Eukaryotic Picoplankton Community in the German North Sea by Different Methods*. University Bremen, Bremen.

Kurtz Z. D., Müller C. L., Miraldi E. R., Littman D. R., Blaser M. J., Bonneau R. A. 2015. Sparse and compositionally robust inference of microbial ecological networks. *PLOS Computational Biology*, 11: e1004226.

Lauritano C., Borra M., Carotenuto Y., Biffali E., Miralto A., Procaccini G., Ianora A. 2011. First molecular evidence of diatom effects in the copepod *Calanus helgolandicus*. *Journal of Experimental Marine Biology and Ecology*, 404: 79 – 86.

Lindeque P. K., Parry H. E., Harmer R. A., Somerfield P. J., Atkinson A. 2013. Next generation sequencing reveals the hidden diversity of zooplankton assemblages. *PLoS One*, 8: e81327.

Löder M. G. J., Meunier C., Wiltshire K. H., Boersma M., Aberle N. 2011. The role of ciliates, heterotrophic dinoflagellates and copepods in structuring spring plankton communities at Helgoland Roads, North Sea. *Marine Biology*, 158: 1551 – 1580.

Löder M. G. J., Kraberg A. C., Aberle N., Peters S., Wiltshire K. H. 2012. Dinoflagellates and ciliates at Helgoland Roads, North Sea. *Helgoland Marine Research*, 66: 11 – 23.

Lucas J., Wichels A., Teeling H., Chafee M., Scharfe M., Gerdts G. 2015. Annual dynamics of North Sea bacterioplankton: seasonal variability superimposes short-term variation. *FEMS Microbiology Ecology*, 91: fiv099.

Mahé F., Rognes T., Quince C., de Vargas C., 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., Dunthorn M. 2015. Swarm v2: highly-scalable and high-resolution amplicon clustering. *PeerJ*, 3: e1420.

McMurdie P. J., Holmes S. 2013. phyloseq: an R Package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One*, 8: e61217.

Medlin L. K., Metfies K., Mehl H., Wiltshire K., Valentin K. 2006. Picoeukaryotic plankton diversity at the Helgoland time series site as assessed by three molecular methods. *Microbial Ecology*, 52: 53 – 71.

Menden-Deuer S., Lessard E., Satterberg J., Grünbaum D. 2005. Growth rates and starvation survival of three species of the pallium-feeding, thecate dinoflagellate genus *Protoperidinium*. *Aquatic Microbial Ecology*, 41: 145 – 152.

Metfies K., Gescher C., Frickenhaus S., Niestroy R., Wichels A., Gerdts G., Knefelkamp B. *et al.* 2010. Contribution of the class Cryptophyceae to phytoplankton structure in the German bight. *Journal of Phycology*, 46: 1152 – 1160.

Mieruch S., Freund J. A., Feudel U., Boersma M., Janisch S., Wiltshire K. H. 2010. A new method of describing phytoplankton blooms: examples from Helgoland Roads. *Journal of Marine Systems*, 79: 36 – 43.

Nejstgaard J. C., Naustvoll L. J., Sazhin A. 2001. Correcting for underestimation of microzooplankton grazing in bottle incubation experiments with mesozooplankton. *Marine Ecology Progress Series*, 221: 59 – 75.

Nelsen M. P., Plata E. R., Andrew C. J., Lücking R., Lumbsch H. T. 2011. Phylogenetic diversity of trentepohlialean algae associated with lichen-forming fungi. *Journal of Phycology*, 47: 282 – 290.

Oksanen J., Blanchet F. G., Friendly M., Kindt R., Legendre P., Mcglinn D., Minchin P. R. *et al.* 2019. vegan: Community Ecology Package. https://cran.r-project.org/web/packages/vegan/vegan.pdf (last accessed 28 August 2020).

Paffenhöfer G. A. 1970. Cultivation of *Calanus helgolandicus* under controlled conditions. *Helgoländer Wissenschaftliche Meeresuntersuchungen*, 20: 346 – 359.

Paffenhöfer G. A. 1971. Grazing and ingestion rates of nauplii, copepodids and adults of the marine planktonic copepod *Calanus helgolandicus*. *Marine Biology*, 11: 286 – 298.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.r-project.org/ (last accessed 28 August 2020).

Rachik S., Christaki U., Li L. L., Genitsaris S., Breton E., Monchy S. 2018. Diversity and potential activity patterns of planktonic eukaryotic microbes in a mesoeutrophic coastal area (eastern English Channel). *PLoS One*, 13: e0196987.

Rohart F., Gautier B., Singh A., Lê Cao K. A. 2017. mixOmics: an R package for 'omics feature selection and multiple data integration. *PLoS Computational Biology*, 13: e1005752 – 19.

Saage A., Vadstein O., Sommer U. 2009. Feeding behaviour of adult *Centropages hamatus* (Copepoda, Calanoida): functional response and selective feeding experiments. *Journal of Sea Research*, 62: 16 – 21.

Sapp M., Wichels A., Wiltshire K. H., Gerdts G. 2007. Bacterial community dynamics during the winter-spring transition in the North Sea. *FEMS Microbiology Ecology*, 59: 622 – 637.

Scharfe M., Wiltshire K. H. 2019. Modeling of intra-annual abundance distributions: constancy and variation in the phenology of marine phytoplankton species over five decades at Helgoland Roads (North Sea). *Ecological Modelling*, 404: 46 – 60.

Schlüter M. H., Kraberg A., Wiltshire K. H. 2012. Long-term changes in the seasonality of selected diatoms related to grazers and environmental conditions. *Journal of Sea Research*, 67: 91 – 97.

Schnack S. 1979. Feeding of *Calanus helgolandicus* on phytoplankton mixtures. *Marine Ecology Progress Series*, 1: 41 – 47.

Sherr E. B., Sherr B. F. 1987. High rates of consumption of bacteria by pelagic ciliates. *Nature*, 325: 710 – 711.

Sommer S. A., Van Woudenberg L., Lenz P. H., Cepeda G., Goetze E. 2017. Vertical gradients in species richness and community composition across the twilight zone in the North Pacific Subtropical Gyre. *Molecular Ecology*, 26: 6136 – 6156.

Sprong P. A. A., Fofonova V., Wiltshire K. H., Neuhaus S., Ludwichowski K. U., Käse L., Androsov A. *et al.* 2020. Spatial dynamics of eukaryotic microbial communities in the German Bight. *Journal of Sea Research*, 163: 101914.

Suzuki K., Nakamura Y., Hiromi J. 1999. Feeding by the small calanoid copepod *Paracalanus* sp. on heterotrophic dinoflagellates and ciliates. *Aquatic Microbial Ecology*, 17: 99 – 103.

Teeling H., Fuchs B. M., Bennke C. M., Krüger K., Chafee M., Kappelmann L., Reintjes G. *et al.* 2016. Recurring patterns in bacterioplankton dynamics during coastal spring algae blooms. *eLife*, 5: 1 – 31.

Thüs H., Muggia L., Pérez-Ortega S., Favero-Longo S. E., Joneson S., O'Brien H., Nelsen M. P. *et al.* 2011. Revisiting photobiont diversity in the lichen family Verrucariaceae (Ascomycota). *European Journal of Phycology*, 46: 399 – 415.

Turner J. 1984. Zooplankton feeding ecology: contents of fecal pellets of the copepods *Eucalanus pileatus* and *Paracalanus quasimodo* from continental shelf waters of the Gulf of Mexico. *Marine Ecology Progress Series*, 15: 27 – 46.

Uchida T., Kamiyama T., Matsuyama Y. 1997. Predation by a photosynthetic dinoflagellate *Gyrodinium instriatum* on loricated ciliates. *Journal of Plankton Research*, 19: 603 – 608.

Whittaker R. H. 1960. Vegetation of the Siskiyou mountains. Orgeon and California. *Ecological Monographs*, 30: 279 – 338.

Wiltshire K. H., Dürselen C.-D. 2004. Revision and quality analyses of the Helgoland Reede long-term phytoplankton data archive. *Helgoland Marine Research*, 58: 252 – 268.

Wiltshire K. H., Malzahn A. M., Wirtz K., Greve W., Janisch S., Mangelsdorf P., Manly B. F. J. *et al.* 2008. Resilience of North Sea phytoplankton spring bloom dynamics: an analysis of long-term data at Helgoland Roads. *Limnology and Oceanography*, 53: 1294 – 1302.

Wiltshire K. H., Kraberg A., Bartsch I., Boersma M., Franke H. D., Freund J., Gebühr C. *et al.* 2010. Helgoland roads, north sea: 45 years of change. *Estuaries and Coasts*, 33: 295 – 310.

Wiltshire K. H., Boersma M., Carstens K., Kraberg A. C., Peters S., Scharfe M. 2015. Control of phytoplankton in a shelf sea: determination of the main drivers based on the Helgoland Roads Time Series. *Journal of Sea Research*, 105: 42 – 52.

Yang J., Löder M. G. J., Wiltshire K. H. 2014. A survey of ciliates at the long-term sampling station 'Helgoland Roads'. North Sea. *Helgoland Marine Research*, 68: 313 – 327.

Yeh H. D., Questel J. M., Maas K. R., Bucklin A. 2020. Metabarcoding analysis of regional variation in gut contents of the copepod *Calanus finmarchicus* in the North Atlantic Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography*, 180: 104738.

Yilmaz P., Kottmann R., Field D., Knight R., Cole J. R., Amaral-Zettler L., Gilbert J. A. *et al.* 2011. Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIxS) specifications. *Nature Biotechnology*, 29: 415 – 420.

7 Publication IV

Publication IV, namely, “Conventional microscopy and metabarcoding of marine plankton in long-term ecological research – Are the results comparable?“, addressed the difficulties of comparing light microscopy with metabarcoding datasets. While there has been emphasis on how metabarcoding benefits plankton research, it is still not a well-established element of long-term monitoring. The main aim of **Publication IV** was to analyse the extent to which the different datasets are comparable and complementary (aim 2a of this thesis). Potential matches and mismatches were identified by comparing diversity, community patterns, and associations with environmental parameters (aims 1a and b of this thesis). Additional comparison was conducted at different taxonomic levels. Especially at higher taxonomic levels, the discrepancy between the datasets is very prominent, mainly because of the differences in taxa coverage. **Publication IV** also emphasizes the differences between the methods and the advantages and disadvantages of each (aim 2b of this thesis). The fact that both methods obtained valuable information on plankton diversity and dynamics underlines the potential and synergistic effects of both for marine plankton research.

Conventional microscopy and metabarcoding of marine plankton in long-term ecological research – Are the results comparable?

Laura Käse[1*], Alexandra C. Kraberg[2], Stefan Neuhaus[2], Maarten Boersma[1,3], Karen H. Wiltshire[1,4], Katja Metfies[2,5]

[1]Alfred-Wegener-Institut, Helmholtz-Zentrum für Polar- und Meeresforschung, Biologische Anstalt Helgoland, Helgoland, Schleswig-Holstein, 27498, Germany

[2]Alfred-Wegener-Institut, Helmholtz-Zentrum für Polar- und Meeresforschung, Bremerhaven, Bremen, 27570, Germany

[3]University of Bremen, Bremen, Bremen, 28359, Germany

[4]Alfred-Wegener-Institut, Helmholtz-Zentrum für Polar- und Meeresforschung, Wadden Sea Station, List auf Sylt, Schleswig-Holstein, 25992, Germany

[5]Helmholtz-Institut für Funktionelle Marine Biodiversität, Oldenburg, Lower Saxony, 26129, Germany

This publication is in preparation for submission.

Sequence data have been deposited in the European Nucleotide Archive under accession number PRJEB37135 (https://www.ebi.ac.uk/ena/data/view/PRJEB37135).

Additionally, the metabarcoding dataset was published on PANGAEA:

Käse, Laura; Kraberg, Alexandra C; Metfies, Katja; Neuhaus, Stefan; Sprong, Pim; Fuchs, Bernhard M; Boersma, Maarten; Wiltshire, Karen Helen (2020): Eukaryotic microbial community at the LTER site Helgoland Roads from March 2016 to March 2019. PANGAEA, https://doi.org/10.1594/PANGAEA.921026.

Further supplementary data (text, tables and figures) can be found in the appendix of this thesis.

Abstract

Metabarcoding as a new evolving method is suitable for plankton time series research that can capture small sized eukaryotic plankton and composition of the whole plankton community. However implementation of new methods to ecological time series always need to be evaluated carefully and critically by comparison to the conventional method in use. Here, we compared plankton time series of three years originating from conventional counts as well as metabarcoding. Comparison on higher trophic levels, such as diatoms or dinoflagellates revealed weak correlating data sets. This is mostly due to several taxa not detected by the different methodologies. On genus level, good correlations could be found. Community patterns were differing, but associations to environmental parameters stayed similar. A great seasonal effect on the communities was visible in all datasets. It was possible to show the influence of small sized microbial eukaryotes to the plankton community patterns by metabarcoding. Despite existing drawbacks in methodology, which might be reduced in the future, metabarcoding is suitable for implementation in time series. Metabarcoding furthermore can provide valuable new information, but should only be used as a complementary method to conventional microscopical time series.

Introduction

In order to understand the functioning of marine ecosystems and their relationships with environmental parameters, it is important to study the composition of marine communities. A variety of hydrographic and physical factors affect the composition of plankton and other marine communities (Prairie *et al.*, 2012; Kröncke *et al.*, 2019). To understand the entire system, knowledge of all parameters over long periods of time is very important to separate inherent system variability from long-term drivers of community/ecosystem dynamics. Here, Long-Term Ecological Research (LTER) sites provide very helpful insights, which are distributed worldwide not only in the marine environment but also in regards to terrestrial and freshwater environments (Dirnböck *et al.*, 2019).

There exist several coastal and marine LTER sites, that have been implemented around the world, some even decades ago, covering different parts of the ocean (Muelbert *et al.*, 2019). They are ranging from single point measurements up to programmes that cover bigger areas, and differ in time intervals, parameters measured and methodology. Both, manual and autonomous long-term measurements generate a very diverse data set with regard to a wide variety of parameters (Muelbert *et al.*, 2019).

One example is the LTER at Helgoland Roads in the German North Sea (Wiltshire *et al.*, 2010). This unique long-term observation site between the main island and the dune island

of Helgoland is influenced by different water masses. These include freshwater influence from the surrounding rivers Elbe, Eider and Weser and advection of water masses from the English Channel. This results in the area around Helgoland being a transitional zone in terms of hydrography. In addition, the strong currents cause a thorough mixing of the whole water column (Hickel, 1972; Callies and Scharfe, 2015). Due to these unique hydrographic conditions, variability in plankton is very high.

The Helgoland Roads time series includes monitoring of several abiotic parameters such as temperature, salinity, inorganic nutrients as well as alkalinity and Secchi depth. Phytoplankton is counted week-daily via light microscopy since 1962 and zooplankton is identified several times per week since 1974 (Wiltshire *et al.*, 2010). As microscopic analysis only reveals a fraction of the whole community. Due to size limitations, interactions of small protists are barely investigated. The importance of the microbial loop to higher trophic levels has been demonstrated in the past (Azam *et al.*, 1983; Sherr and Sherr, 2002), which makes inclusion of these small size fraction necessary for complete food web analysis. Therefore, these conventional long-term observations need to be supported by complementary case studies using different methods to shed light on the black boxes, and to evaluate the potential and feasibility of implementing new methods into existing LTERs.

Generally, morphological surveys using microscopic counts have several advantages and disadvantages compared to molecular methods (see also Table 1). Besides the size limitation, light microscopy is very time consuming, requires good taxonomic knowledge (Not *et al.*, 2004; Stern *et al.*, 2018) and only a small fraction of sample can be investigated. However, compared to new molecular methods, light microscopy is easy to handle at low costs. At the same time, it can provide information on plankton species diversity including identification of different life stages, or even potential parasitic infections. However, electron microscopy (e.g. transmission electron microscopy (TEM) or scanning electron microscope) is needed for clear identification of parasitic or epiphytical taxa and can not be distinguished by light microscopy alone. Depending on the focus of plankton studies, a combination of conventional methods and molecular methods such as metabarcoding could give detailed insights in the plankton community.

So far, a lack in integration of molecular methods with conventional microscopy monitoring was reported in long-term research (Stern *et al.*, 2018). However, the use of eDNA and metabarcoding approaches, sampling and handling procedures have the potential to provide automated and standardized workflows (Zimmermann *et al.*, 2015). However, by covering the whole plankton community and giving insights to unknown size fractions, the approach also faces several challenges. These range from methodological drawbacks and

insufficient taxonomic information leading to poor quality of reference data bases (McManus and Katz, 2009). Differences in copy numbers of marker genes between species, for example, can influence the results massively and complicate abundance measurements and the comparability, not only within a single dataset but also to other datasets (Zhu *et al.*, 2005; Medinger *et al.*, 2010; Thomas *et al.*, 2016).

In general, studies have compared their metabarcoding results with datasets based on different methods such as microscopy or datasets based on mock samples. These are mainly focusing on different groups of plankton in several instances and environments (Abad *et al.*, 2016; Djurhuus *et al.*, 2018; Pitsch *et al.*, 2019; Kang *et al.*, 2021). While some studies suggest that especially for larger sized plankton the biomass correlates to the read abundances (Harvey *et al.*, 2017; Hirai *et al.*, 2017; Gran-Stadniczeñko *et al.*, 2019), other studies showed that cell counts can be correlated as well (Wollschläger *et al.*, 2014; Giner *et al.*, 2016; Bucklin *et al.*, 2019; Metfies *et al.*, 2020). Metabarcoding of mesozooplankton, which are mostly sampled using net samples, mostly resulted in a higher diversity compared to morphological counts (Doherty *et al.*, 2007; Lindeque *et al.*, 2013; Clarke *et al.*, 2017; Yang *et al.*, 2017).

Table 1: Differences, Bias and Potential Solutions in morphological (MO) and metabarcoding methods (MB)

	Bias	Resulting Bias	Potential solution(s)	Practical (Yes/No)
Sampling and Sample Volume	Different volumes (MO, MB)	Due to lower volume rare Taxa might not be included (MO)	Identical sampling and subsampling/replicates for methods Use identical samples for both methods	Yes (No)
Identification	Based on taxonomic expertise (MO) Dependent on quality of sequencing and Database (MB) Mismatch of taxonomy and differences in nomenclature (MO/MB)	Dependent on specific focus certain taxa might be favoured (MO/MB) Resolution of taxonomy (MO/MB)	Quality control/“Resampling“ by use of replicates Use of different databases Use of specific/different primers	Yes Yes Yes

Similar results were found for studies focusing on microzooplankton and phytoplankton diversity (Xiao *et al.*, 2014; Zimmermann *et al.*, 2015; Gran-Stadniczeñko *et al.*, 2017; Gran-Stadniczeñko *et al.*, 2019).

Similar to the studies cited above, molecular studies that have been carried out in addition to the conventional LTER at Helgoland Roads focused only on certain plankton groups (Gescher *et al.*, 2008; Knefelkamp, 2009; Metfies *et al.*, 2010, 2020; Sprong *et al.*, 2020), instead of providing a complete overview of the plankton community. A first comparison of occurrences of diatoms and dinoflagellates during a spring bloom, revealed differences between our metabarcoding results and the regular LTER counts (Käse *et al.*, 2020a), especially since metabarcoding revealed dinoflagellates as the main contributor to the plankton community. As a metabarcoding dataset that covers all levels of the planktonic food web over three years at Helgoland Roads is available (Käse *et al.*, 2021b), we have the unique opportunity to compare taxa of all trophic levels to the conventional LTER data. Therefore, the aim with this paper is to identify the extent to which different datasets based on microscopy and metabarcoding are comparable and complementary by comparison of different taxa that have been observed for several years and to examine what combination/taxonomic resolution of data will maximize the explanatory power in terms of detecting community patterns/change.

Material and Methods

General information on sampling procedure

Water surface samples were taken with a bucket at 1 m depth at the LTER sampling site (54°11’ N, 7°54’E) in work-daily intervals from March 2016 to March 2019 according to the LTER protocols (Hickel *et al.*, 1993; Wiltshire *et al.*, 2008, 2010). Aliquots of this water sample were used for the conventional phytoplankton LTER, for additional microscopic counts of certain taxa and for metabarcoding. Zooplankton LTER samples are taken with nets several times a week (Greve *et al.*, 2004).

Microscopic Counts

LTER phytoplankton counts were conducted work-daily as according to the LTER sampling protocol (Wiltshire *et al.*, 2010), while LTER zooplankton counts were conducted mainly three times a week (Greve *et al.*, 2004; Boersma *et al.*, 2015).

Additional to the LTER Counts of Phyto- and Zooplankton from March 2016 to March 2019, we used samples from December 6th 2016 to March 14th 2019 for additional microscopic counts. These microscopic counts were also based on Utermöhl inverted microscopy method (Utermöhl, 1958). 250 mL of seawater were fixed with 1.25 mL neutral Lugol solution and stored at room temperature and in the dark. After homogenization of the sample, an aliquot of 25 mL was filled in an Utermöhl chamber for sedimentation. After at least 24 hours the settled cells were counted at a magnification of 400 with an inverted

microscope (Axiovert 135, Zeiss, Germany). Defined taxa were quantified as detailed as possible using different groups distinguished by size classes and morphotypes as well as some unambiguously identifiable species. Counting was conducted in transects, a minimum of 50 individuals per group were needed to not continue counting this group in the follow up transects. If a minimum of 50 was not reached while counting these 5 transects, the complete chamber was counted for this group. The cells counted were documented photographically (AxioCam HRc & AxioVision V. 4.9.1, Zeiss, Germany).

Molecular analysis

General methods on sampling intervals and details, DNA extraction, MiSeq™ Illumina sequencing and Data processing have been described before (Käse *et al.*, 2021b, 2021a).

In short, we combined samples from three different sampling intervals, where 1 L of seawater was filtered each and all samples were stored at -20°C until DNA isolation. Due to different filtering techniques two different protocols were used for DNA extraction and sequentially filtered samples were pooled accordingly to achieve one sample per sampling date for all samples. A fragment (V4 region) of the 18S ribosomal (r) DNA was amplified using KAPA HiFi HotStartReadyMix (Kapa Biosystems, Inc., Roche, Germany) and the following primer set developed by Fadeev *et al.* (2018): 528iF (GCG GTA ATT CCA GCT CCA A) and 964iR (ACTTT CGT TCT TGA TYR R). About 43 million 2x300 bp paired-end sequences were produced using an Illumina MiSeq™ sequencer (Illumina, USA). After data processing, 21 million sequences remained and were clustered using s*warm* (version 2.2.2, Mahé *et al*. 2014, 2015). The Protist Ribosomal Reference database (PR2), version 4.11.1 (Guillou *et al.*, 2013) was used as reference. Sequence data for this study have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession number PRJEB37135, using the data brokerage service of the German Federation for Biological Data (GFBio, Diepenbroek *et al.*, 2014), in compliance with the Minimal Information about any (X) Sequence (MIxS) standard (Yilmaz *et al.*, 2011). Details of our pipeline are available on GitHub (https://github.com/PyoneerO/qzip) and the full table of operational taxonomic units (OTU, 280 samples with 59.284 OTUs) was archived in PANGAEA (Käse *et al.*, 2020b).

Dataset comparison and correlation analysis

Two different thresholds were used for the metabarcoding dataset for statistical analysis. One threshold was set at 0.001% of total reads (**205 reads in total**), resulting in 2790 OTUs, for the second threshold only OTUs with **at least 50 reads in total** were used (6108 OTUs in total). Reads of both metabarcoding datasets were then converted to relative abundances. The whole dataset was filtered for the respective taxa for diversity comparison, which was

found by microscopy counts, if the taxa was not found in the limited datasets. Identification by the database was accepted up to genus level, to verify the respective taxa BLAST alignment was used when identification by PR2 was not sufficient. BLAST alignment of known genera for alpha diversity was only conducted for the limited dataset (minimum 50 reads). Different lists of known taxa based on phytoplankton (Kraberg *et al.*, 2019), zooplankton (Harms, 1993; Greve *et al.*, 2004) were used. Datasets were checked for different spellings (e.g. *Phalacroma*/*Phalachroma*) and synonyms (e.g. *Tripos* and *Ceratium*) as they are used by the respective taxonomic databases (PR2, taxa lists). Additionally it needs to be mentioned that some taxa can not be identified sufficiently in the version of the PR2 database used (e.g. new taxa such as *Plagiolemma*). *I*dentification of certain taxonomic groups such as Ciliophora is also weak when using PR2 and BLAST.

All follow-up analyses, as described below, were conducted in R, version 4.0.2, (R Core Team, 2020). The significance level was set at $p<0.05$ for significance tests. We calculated two different coefficients using the cor.test function of the stats package for correlation analysis between the different datasets: Spearman ρ, Kendall τ. We chose these non-parametric methods, because both coefficients do not need the data to be normally distributed, they are more suited than the common Pearson r coefficient, with the Spearman correlation coefficient being a rank-based version of the Pearson's r. Negative values were not considered for interpretation as we expected the datasets to be positively correlated. We defined the positive range of the correlation coefficients with the following terms: Coefficients below 0.1 are considered to show no correlation between the datasets, coefficients above 0.1 and below 0.4 are weakly correlated, between 0.4 and 0.7 correlation was termed as good, above 0.7 we defined a strong correlation.

Analysis of community patterns on different taxonomic levels

We conducted two ordination analyses for examination and comparison of community patterns as they are displayed by the different datasets on different taxonomic levels: Non-metric multi dimensional scaling (NMDS) and Canonical correspondence analysis (CCA). Hereby the focus was on finding matching patterns in the LTER count datasets as previously reported in the metabarcoding data, such as the unique community composition in 2018 and the similar communities occurring during seasons of different years (Käse *et al.*, 2021b).

The real counts were combined on respective levels to create the count datasets: phytoplankton LTER (phylum, to taxa as named by counts dataset, 6 levels), zooplankton LTER (phylum to taxa as named by dataset, 5 levels). The relative abundances of the sequencing datasets were also combined (phylum to OTU level, 6 levels). If no level had been assigned, taxa were combined under "other". Then we conducted NMDS analysis by

using the metaMDS function of the vegan package (Oksanen *et al.*, 2019). We used the Bray Curtis distance and the data were autotransformed (Square root transformation, Wisconsin double standardization). Iterations (maxit) were set to 1000, maximum numbers of random starts in search of stable solution (trymax) was set to 250. Only convergent solutions with sufficient data were plotted with ggplot2 (Wickham, 2016). Datasets were combined on nearly all levels (see NMDS analysis for details), only the zooplankton dataset could not be analysed on phylum level (not enough taxa). Several samples of the counts and sequencing datasets were removed for CCA (Suppl. Table 1), due to missing data in the environmental parameters.

The parameters included information on season, temperature, salinity, Secchi depth, tide and sunshine duration as well as silicate, nitrate, ammonium and phosphate concentrations. All metabarcoding and count datasets were Hellinger transformed using the decostand function of the vegan package. We then conducted a “Constrained Ordination Model”, where the environmental parameters were added sequentially. The model was based on an analysis of variance (ANOVA)-like permutation test for CCA to assess the significance of the constraining factors, by testing for single term additions (Oksanen *et al.*, 2019). Maximum permutations were set to 1000. The single parameters were added to the model sequentially after they were chosen by their significance ($p<0.05$). If several significances were found in the same step, the lowest Akaike information criterion (AIC) value was chosen to minimize the information loss (Akaike, 1974). Furthermore, we conducted additional ANOVA to check for significances of the whole model (Permutation test for cca under reduced model) and of the single environmental parameters added to the model (Terms added sequentially (first to last)).

Results

Alpha diversity at genus level and general overview

In general metabarcoding revealed more genera than are known or regularly counted, especially for taxa that do not belong to Bacillariophyta, Dinoflagellata or Crustacea (Table 2). However, the phytoplankton LTER counts reveal more taxa on species level and size classes for most diatoms, dinoflagellates and other important taxa as the metabarcoding results are not reliable for most taxa on this level.

Table 2: Number of genera found from March 2016 to March 2019, Metabarcoding include all genera cut to the respective threshold, genera listed here include the ones identified by PR2, further BLAST identification is only mentioned in the text; LTER numbers only include taxa that are being counted regularly (other known taxa are only mentioned in the text).

	Metabarcoding (50 reads)	**Metabarcoding (0.001%; 205 reads)**	**"Phytoplankton LTER"**	**"Zooplankton LTER"**
Bacillariophyta	48 genera + unclassified	40 genera + indetermined	40 genera + indetermined	N/A
Dinoflagellata	44 genera + unclassified	36 genera + unclassified	18 genera + indetermined	only Noctiluca
Crustacea	26 genera + unclassified	21 genera + unclassified	N/A	7 genera + higher taxa & indetermined
Other taxa	413 genera + unclassified	278 genera + unclassified	9 genera + size groups of indetermined taxa	only higher taxa

The high number of OTUs that could not be identified to genus level furthermore hides more potential diversity. For example, 359 OTUs (threshold of 50 reads) in the dataset could not be assigned to any kingdom by PR2 (Eukaryota unclassified). PR2 identification on genus level was furthermore not possible for 3640 of 6108 OTUs (59.6 %, threshold of 50 reads). As the Zooplankton LTER for Crustacea comprises species or certain genera in groups, these taxa made up only a fraction of the genera found by metabarcoding.

The number of genera for Bacillariophyta (diatoms) was nearly similar for the metabarcoding dataset and the known genera in the LTER, however both methods also identified genera that were not identified by the other method (Figure 1). 27 Bacillariophyta genera were found by all methods (Figure 1, Suppl. Table 2). Two additional genera, that were present in the counts, were found in the dataset of the lower threshold. In total, 19 taxa were only found by metabarcoding, whereas 11 taxa were only found by counts (details on all genera see Suppl. Text and Suppl. Table 2).

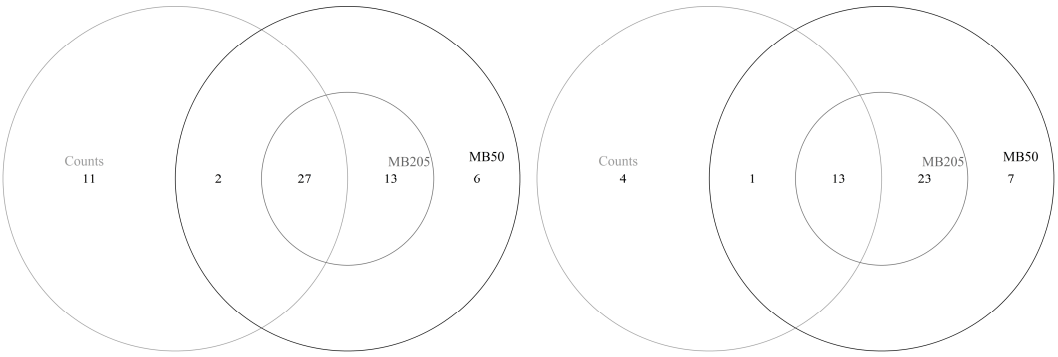


Figure 1: Venn diagram displaying shared genera by different methods for a) Bacillariophyta and b) Dinoflagellata; Counts = "Phytoplankton LTER", MB50 = metabarcoding results with a threshold of 50 reads, MB205 = metabarcoding results with a threshold of 205 reads.

When identifying Dinoflagellata, there also existed a bigger gap in known genera compared to the genera found by metabarcoding (Table 2). 13 genera belonging to Dinoflagellata were found by all methods (Figure 1, Suppl. Table 3). In total, 30 taxa were only found by metabarcoding, whereas 4 taxa were only found by counts (details on all general see Suppl. Text and Suppl. Table 3).

Out of the seven zooplankton genera (all copepods) that are counted at Helgoland Roads, two are combined into one genus complex (*Pseudo*-/*Paracalanus*), whereas further Crustacea as well as other zooplankton groups are counted on higher taxonomic levels. All of these seven genera were found in the metabarcoding dataset no matter which threshold was set (Table 2). Twelve other genera, that were found in at least one MB dataset, have been reported before, seven additional genera were found that have not been reported before at the LTER site (Suppl. Table 4).

Excluding Bacillariophyta, Dinoflagellata and Crustacea metabarcoding revealed 413 further genera belonging to 23 different phyla in total for a threshold of 50 reads and 278 genera of 20 different phyla for a threshold of 205 reads (Table 2 and Suppl. Table 5). Most genera counted at the LTER were found in the metabarcoding dataset at a threshold of 50 reads, however especially identification (PR2 and BLAST) of Ciliophora OTUs was sparse (see details in Suppl. Text).

Correlation of selected taxonomic groups and single genera

Diatoms in general showed no or only weak correlation between the LTER counts and metabarcoding dataset (Table 3), while Dinoflagellate and Crustacea datasets were weakly correlated. Correlation was better when comparing the LTER counts to Dinoflagellate sequences that did not include the parasitic Syndiniales than the comparison to the sequences including this group. The datasets of Cryptophyta & Katablepharidophyta combined showed contrasting values with no positive correlation. Correlation of different Coccolithophorida datasets was either weak or good depending on the correlation parameter.

Table 3: Correlation parameters of datasets of different groups of taxa; LTER = regular LTER microscopic counts, amc = additional microscopic counts for this study, seq = relative abundances of the metabarcoding datasets with threshold of 50 and 205 reads, no = number of compared observations, t = threshold of sequencing datasets, n OTUs = number of OTUs depending on the threshold, significance levels: $p < 0.001$ (***), $p < 0.01$ (**) and $p < 0.05$ (*).

Taxa	Comparison	no	t	n OTUs	Spearman ρ	Kendall τ
Diatoms	LTER – seq	280	50	596	0.130*	0.096*
			205	209	0.139*	0.010*
Dinoflagellates	LTER – seq[1]	280	50	818	0.189**	0.132**
			205	442	0.183**	0.128**
	LTER – seq[2]	280	50	488	0.351***	0.239***
			205	236	0.345***	0.234***
Crustacea	LTER – seq	53	50	996	0.302*	0.206*
			205	339	0.313*	0.213*
Cryptophyta & Katable-pharidophyta	amc – seq	224	50	36	-0.251***	-0.171***
			205	25	-0.250***	-0.172***
Coccolithophorida	LTER – amc	229	n.a.	n.a.	0.323***	0.235***
	LTER – seq	280	50	43	0.252***	0.182***
			205	27	0.253***	0.183***
	amc – seq[3]	224	50	43	0.407***	0.281***
			205	27	0.405***	0.279***
	LTER – seq[4]	280	50	9	0.284***	0.206***
			205	6	0.283***	0.204***
	amc – seq[4]	224	50	9	0.452***	0.315***
			205	6	0.453***	0.316***

[1] all Dinoflagellata, [2] Dinoflagellata without Syndiniales, [3] including three genera (*Braarudosphaera*, *Coccolithus*, *Emiliania*) and all unclassified Prymnesiophyceae, [4] including three genera, unclassified Noelaerhabdaceae and Syracosphaerales

Both *Phaeocystis* counts datasets revealed, that several peaks in abundance, as shown by the metabarcoding results, were not detected (Figure 2). All datasets showed high peaks in May 2018, in April 2016 higher peaks were only detected in the metabarcoding dataset. Whereas the two microscopic datasets for *Phaeocystis* showed a good correlation (Table

4), Kendall and Spearman correlation parameters were negative, when comparing the metabarcoding datasets with the counts. Similarly, comparison between *Pseudo-nitzschia* datasets revealed strong or good correlation between the microscopic counts (Table 4). Correlations to the metabarcoding dataset were good or weak depending on the correlation parameter. Especially high peaks that were detected in the metabarcoding dataset during spring 2016 (Figure 2) were mainly caused by *P*. *americana*, which is morphologically different from other *Pseudo-nitzschia* species and often found epiphytically. Other high peaks during the following years were mostly also detected by the microscopic counts. Correlations that were calculated for other exemplary diatom genera respective species, were varying greatly as well (Table 4). A good correlation between the LTER microscopic counts and the sequences was found for *Thalassionema*, while comparison of *Thalassiosira* datasets resulted in weak or no correlation.

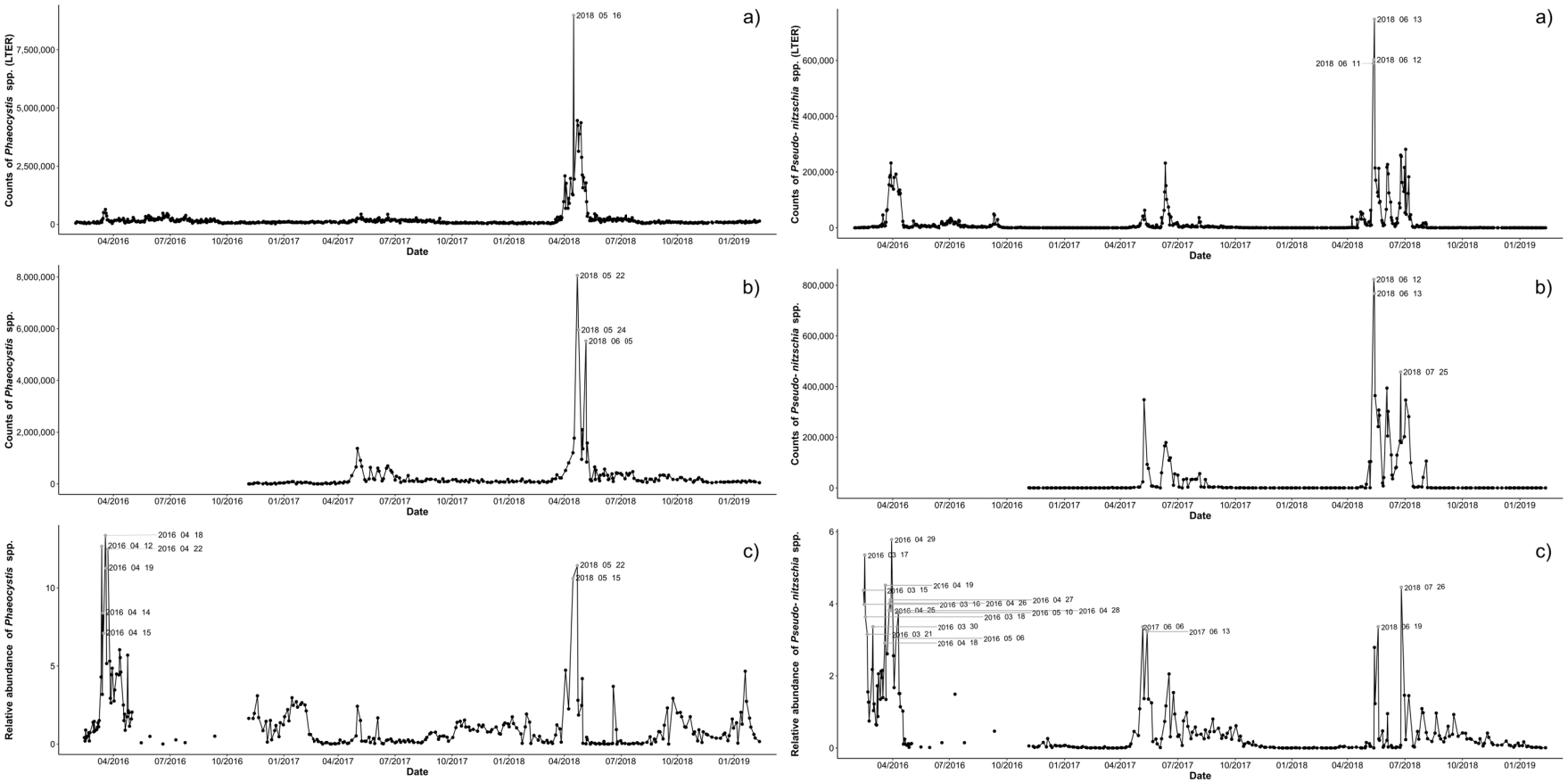


Figure 2: Comparison of *Phaeocystis* (left) and *Pseudo-nitzschia* (right) datasets, a) and c) from March 2016 to March 2019, b) from December 2016 to March 2019, Peaks with heights corresponding to at least 50% of the respective method's maximum are marked, relative abundances of the metabarcoding dataset are based on the 205 reads threshold.

Table 4: Correlation parameters of datasets of different genera; LTER = regular LTER microscopic counts, amc = additional microscopic counts for this study, seq = relative abundances of the metabarcoding datasets with threshold of 50 and 205 reads, no = number of compared observations, t = threshold of sequencing datasets, n OTUs = number of OTUs depending on the threshold, significance levels: p < 0.001 (***), p < 0.01 (**) and p < 0.05 (*).

Taxa	Comparison	no	t	n OTUs	Spearman ρ	Kendall τ
Phaeocystis	LTER – amc	229	n.a.	n.a.	0.551***	0.392***
	LTER – seq	280	50	8	-0.078	-0.057
			205	5	-0.082	-0.060
	amc – seq	224	50	8	-0.147*	-0.105*
			205	5	-0.150*	-0.108*
Pseudo- nitzschia	LTER – amc	229	n.a.	n.a.	0.801***	0.693***
	LTER – seq	280	50	6	0.626***	0.473***
			205	5	0.624***	0.472***
	amc – seq	224	50	6	0.477***	0.359***
			205	5	0.476***	0.359***
Thalassionema	LTER – seq	280	50	2	0.551***	0.439***
			205	2	0.550***	0.439***
Thalassiosira	LTER – seq[1]	280	50	68	0.192**	0.136**
			205	21	0.200***	0.142***
	LTER – seq[2]	280	50	292	0.119*	0.081
			205	51	0.152*	0.107*
Chaetoceros	LTER – seq[1]	280	50	38	0.245***	0.178***
			205	29	0.237***	0.173***
	LTER – seq[2]	280	50	47	0.522***	0.393***
			205	35	0.543***	0.371***
Chaetoceros socialis	LTER – seq[3]	280	50	1	0.316***	0.242***
			205	1	0.586***	0.495***
	amc – seq[3]	224	50	1	0.467***	0.346***
			205	1	0.222***	0.173**
	LTER[4] – amc	229	n.a.	n.a.	0.426***	0.381***
	LTER[5] – amc	229	n.a.	n.a.	0.824***	0.738***

[1] PR2, [2] PR2+BLAST, [3] OTU only identified by BLAST, [4] only *C. socialis*, [5] *C. socialis* and *Chaetoceros* sp. (< 10 µm),[6] only "Phytoplankton LTER"

Table 4: Correlation parameters of datasets of different genera (continued).

Taxa	Comparison	no	t	n OTUs	Spearman ρ	Kendall τ
Akashiwo	LTER – seq	280	50	4	0.576***	0.519***
			205	1	0.576***	0.519***
Tripos	LTER – seq	280	50	9	0.472***	0.358***
			205	5	0.471***	0.358***
Dinophysis	LTER – seq	280	50	1	0.548***	0.460***
			205	1	0.548***	0.460***
Gyrodinium	LTER – seq	280	50	27	0.338***	0.244***
			205	16	0.337***	0.243***
Diplopsalis	LTER – seq	280	50	1	0.230***	0.201***
Noctiluca	LTER[6] – amc	229	n.a.	n.a.	0.680***	0.641***
	LTER[6] – seq	280	50	12	0.469***	0.376***
			205	1	0.468***	0.375***
	amc – seq	224	50	12	0.465***	0.375***
			205	1	0.464***	0.375***
Oithona	LTER – seq	53	50	17	0.488***	0.346***
			205	4	0.485***	0.343***

[1] PR2, [2] PR2+BLAST, [3] OTU only identified by BLAST, [4] only *C. socialis*, [5] *C. socialis* and *Chaetoceros* sp. (< 10 µm),[6] only "Phytoplankton LTER"

If looking on genus level, *Chaetoceros* LTER counts were mostly weakly or even good correlated (Table 4). Hereby, a better correlation coefficient was found when the metabarcoding dataset did not only consist of sequences identified by PR2 but also by additional sequences as identified by BLAST analysis. On species level, a *Chaetoceros socialis* OTU (identified by BLAST alignment) showed contrasting correlation values to the LTER and additional counts of *C. socialis* depending on the threshold of the datasets. Additionally, we compared different LTER counts of *Chaetoceros* with the additional counting dataset. Hereby, the combination of *C. socialis* counts with Chaetoceros sp. counts with a size below 10 µm resulted in a better correlation to the additional counts than the correlation to only *C. socialis* counts of the regular LTER, indicating the difficulty of identifying these small sized diatoms.

In general, the correlation between different microscopic counts of *Noctiluca* were good, whereas correlation of these counts with the metabarcoding dataset were good when looking at Spearman but weak for Kendall (Table 4). It needs to be noted that other Noctilucales genera *Kofoidinium* and *Spatulodinium* (syn. *Gymnodinium*) were also found in the metabarcoding dataset, which might be included as *Noctiluca* in the microscopic counts as they are morphologically similar. The highest peak in 2017 was detected on the same day in all samples, high abundances in 2018 were also displayed by all three datasets (Figure 3).

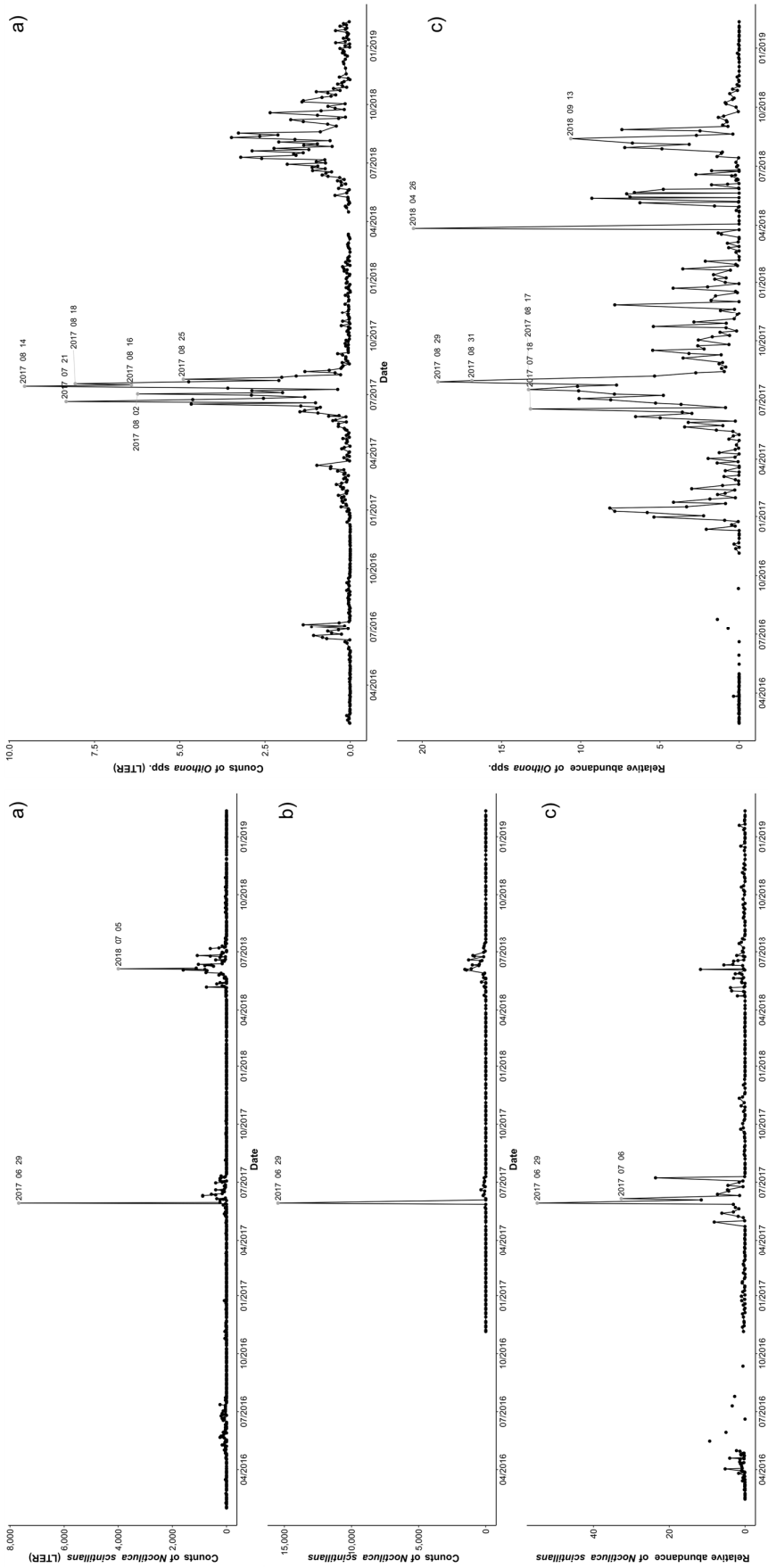


Figure 3: Comparison of *Noctiluca scintillans (left) and Oithona* spp. (right) datasets, a) and c) from March 2016 to March 2019, b) from December 2016 to March 2019 (only left), Peaks with heights corresponding to at least 50% of the respective method's maximum are marked, relative abundances of the metabarcoding dataset are based on the 205 reads threshold.

Results for other dinoflagellate genera that were only included in the LTER microscopic counts and the metabarcoding dataset were varying between different genera but also between different correlation parameters, with good (*Akashiwo* and *Dinophysis*) or weak (*Gyrodinium* and *Diplopsalis*) correlations being found (Table 4). Both, weak and good correlations, were found for *Tripos* (counted as *Ceratium* in the LTER) depending on the respective parameter.

In regards to the copepod Oithona weak or good correlations were found between the Zooplankton LTER counts and the metabarcoding dataset. However, it needs to be mentioned that less datapoints were included in the analysis compared to other taxa, as the sampling dates were not matching. Included were 32 samples from spring 2016 (March to May), as well as occasional samples from July, August and October 2016, May, June and December 2017, and May, June and July 2018. Comparison of peaks revealed matching phases in summer months (Figure 3), however several peaks during other seasons, which were found in the metabarcoding dataset, were not detected in the counts. The highest peak in the metabarcoding dataset was found in April 2018, however, no LTER counts were available during this time frame.

Additionally, copepod numbers of the 7 genera counted routinely at the LTER, were compared to the twelve highest peaks of the same copepod genera as depicted by the metabarcoding analysis (Figure 4). Similar to the individual result for the genus *Oithona*, high abundances were found during summer months in both time series. However, comparison throughout the time period reveals that the peaks of the respective genera do not match for the two datasets. Metabarcoding results furthermore revealed peaks in relative abundances during other seasons, indicating that these taxa are also present during other months, but might not be sampled in the LTER counts in such high amounts compared to adult state copepods.

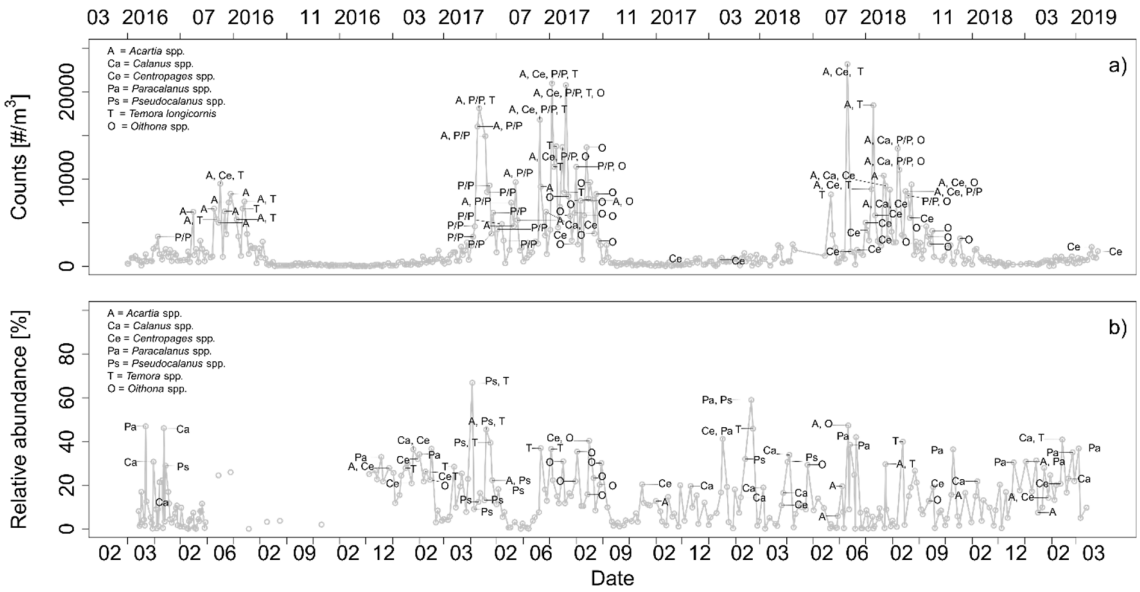


Figure 4: Comparison of a) microscopic counts of different copepod genera to b) relative abundances of the same copepod genera; for counts all peaks of more than 20% of the maximum value of each genus were marked; for relative abundances the twelve highest relative abundances for each genus were marked, relative abundances are based on the 50 reads threshold.

Analysis of community patterns on different taxonomic levels

Both sequencing datasets showed that spring and summer samples of 2018 were vastly differing from other seasons or years on all taxonomic levels. All levels reached convergence with stress being lower in the dataset with the 205 read threshold (Table 5). Stress was increasing each time from genus to phylum level, only NMDS on OTU level had a higher stress value compared to some higher levels.

Table 5: NMDS input and statistics, number of taxa included in the analysis on the respective taxonomic level, information on convergence, number of runs needed to reach convergence and stress values, Last taxa = Last assigned taxa, *warning: insufficient data.

Dataset	Taxonomic Level	Number of Taxa	Convergence reached yes/no	Stopped at run x (min=20)	Stress
"Phytoplankton LTER"	Last taxa	173	yes	26	0.2102474
	Genus	68	no	NA	NA
	Family	25	no	NA	NA
	Order	16	no	NA	NA
	Class	11	no	NA	NA
	Phylum	7	no	NA	NA
"Zooplankton LTER	Last taxa	43	yes	80	0.2282288
	Family	13	yes	61	0.1831734
	Order	8	yes	81	0.1538554
	Class	17	yes	35	0.2390478
	Phylum	2	yes	136	3.22384e-16*
Metabarcoding (50 reads)	OTU	6108	yes	20	0.1675477
	Genus	985	yes	20	0.1300917
	Family	518	yes	20	0.1435393
	Order	259	yes	20	0.1555015
	Class	138	yes	27	0.1800682
	Phylum	36	yes	125	0.1909234
Metabarcoding (0.001%; 205 reads)	OTU	2790	yes	20	0.14473
	Genus	734	yes	20	0.1256866
	Family	420	yes	20	0.1392937
	Order	212	yes	22	0.1536132
	Class	115	yes	20	0.1707951
	Phylum	30	yes	74	0.1785245

No convergence was reached for most levels using the phytoplankton LTER dataset. On taxa level (level, where each taxa was last assigned by the LTER dataset) the samples of spring and summer 2018 did not show a different pattern than other samples. While the NMDS analysis reached convergence for the zooplankton LTER on every level, there was also no difference in the samples from 2018 compared to other samples. In general, stress of the count datasets was higher compared to the sequencing datasets. Overall, similar patterns in seasonality, with seasons clustering together in the different years could be observed in the sequencing datasets and the phytoplankton count dataset (Figure 5, all NMDS plots of the metabarcoding datasets and the zooplankton LTER can be found in Suppl. Figure 1 - 3).

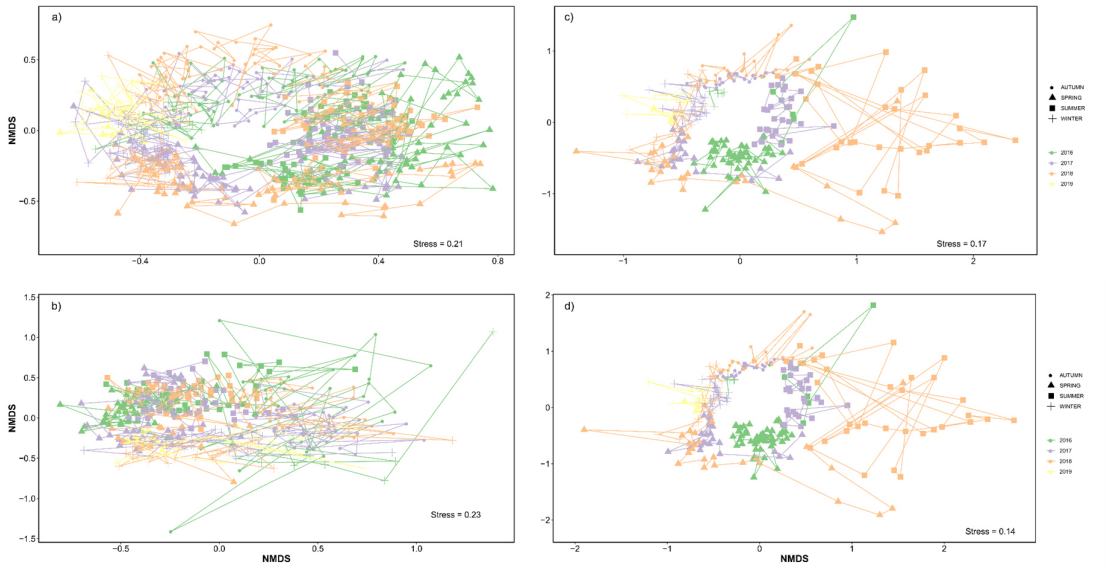


Figure 5: NMDS plots of the a) phytoplankton LTER dataset and b) zooplankton LTER dataset, both on last assigned taxa level, and c) metabarcoding dataset (50 reads) and d) metabarcoding dataset (0.001%, 205 reads), both on OTU level, seasons are displayed by different shapes, years are color coded.

The results of the different CCA models showed a high agreement on significant parameters for both metabarcoding datasets and the LTER datasets (Suppl. Table 1). At all times, seasons were the first parameter added to the model, followed by either temperature or silicate. The parameters tide and phosphate never were a significant variable in the CCAs based on the metabarcoding datasets. Phosphate was found to be significant once on genus level for the phytoplankton LTER. Inertia explained in full space was varying from 17.25 to 21.97 % for the metabarcoding dataset (50 reads), from 19.76 to 21.87 % for the metabarcoding dataset (205 reads), from 23.53 to 36.58 % in the phytoplankton LTER dataset and from 24.40 to 32.64 % in the zooplankton dataset, respectively. While highest values were reached on family level for the metabarcoding datasets, the phytoplankton LTER was best explained on phylum level, the zooplankton LTER on order level. In restricted space highest inertia was explained on phylum level for both metabarcoding and the phytoplankton LTER dataset.

In contrast to the NMDS, CCA of two taxonomic levels of the phytoplankton LTER (class and phylum level) also revealed a unique pattern in spring 2018 similarly to the metabarcoding datasets, which also displayed this in the CCA (Figure 6, all CCA plots can be found in Suppl. Figure 4 to 7). Additionally, at different levels the CCA of the phytoplankton LTER revealed a high variation in 2016, e.g. in spring (Suppl. Figure 6 a to d) or summer and autumn (Suppl. Figure 6 c and d). Hereby, it needs to be noted, that summer and autumn 2016 were only sampled sporadically for the metabarcoding dataset.

Especially, winter samples of the phytoplankton LTER dataset were tightly clustered. The zooplankton LTER showed no familiar patterns to the other datasets and significant variables were varying depending on the taxonomic level.

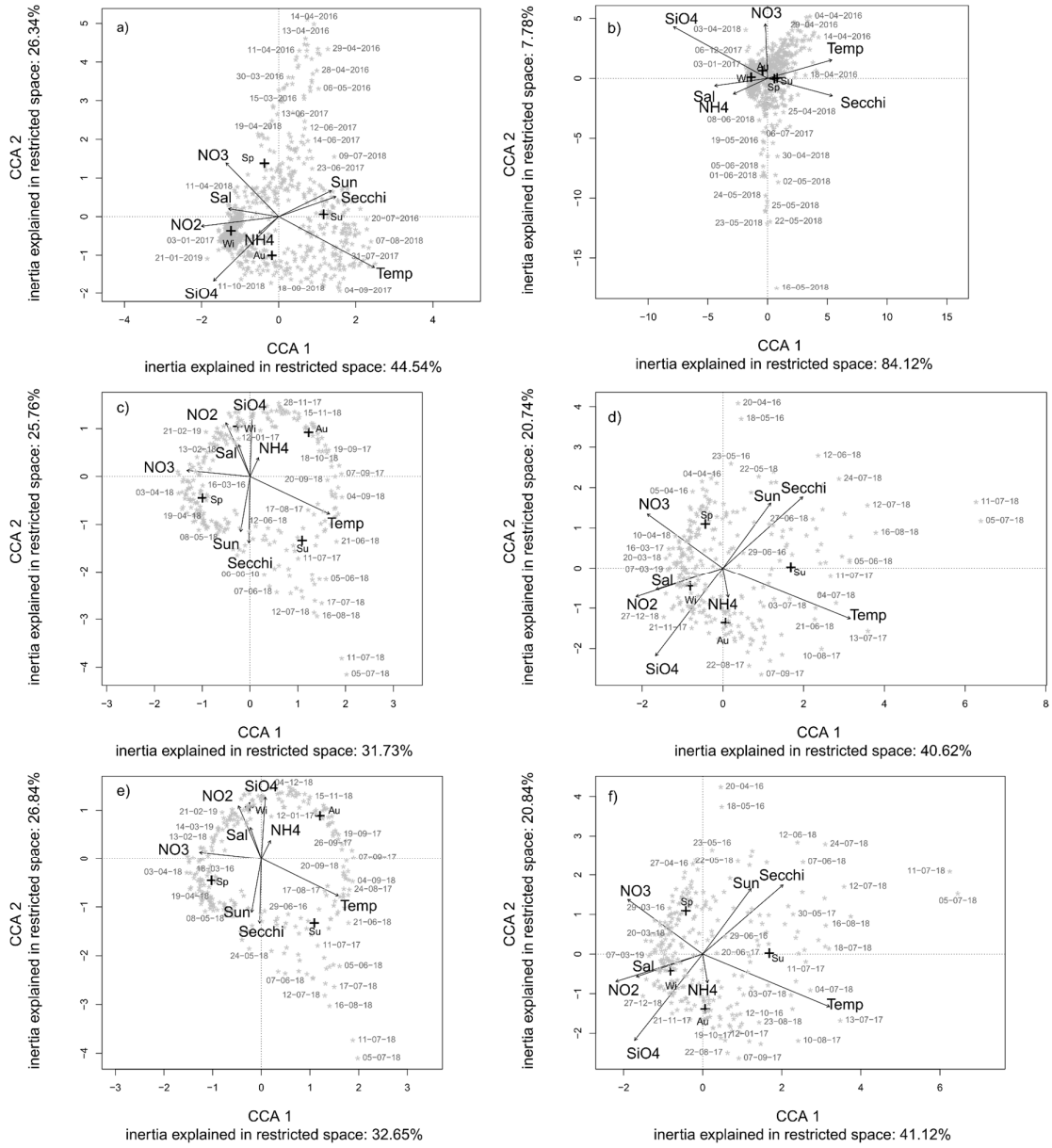


Figure 6: CCA plots of the a) and b) phytoplankton LTER dataset on a) taxa and b) class level, c) and d) metabarcoding dataset (50 reads), e) and f) metabarcoding dataset (0.001%, 205 reads), c) and e) on OTU and d) and f) on class level, samples are plotted as grey asterisks with partially displayed sampling dates, including significant parameters in black (if applicable): seasons (Sp = spring, Su = summer, Au = autumn, Wi = winter), Temperature (Temp), salinity (Sal), Secchi depth (Secchi), sunshine duration (Sun), silicate (SiO4), nitrate (NO3), nitrite (NO2), ammonium (NH4), phosphate (PO4); additional information on total inertia (full space) and inertia in restricted space can be found in Suppl. Table 1.

Discussion

So far, by using metabarcoding, we were able to shed light on a wide diversity of eukaryotic plankton of all size classes (Käse *et al.*, 2020a, 2021b), including diverse parasitic plankton that is not usually covered in conventional plankton research (Käse *et al.*, 2021a). Furthermore, we identified steady background communities consisting of plankton of all trophic levels that are consistent throughout certain seasons (Käse *et al.*, 2020a) or even throughout several years and visualised the high connectivity of the plankton network (Käse *et al.*, 2021b).

The high diversity that is depicted by this metabarcoding dataset covers nearly all genera that are found or known by the LTER data. At the same time it provides a higher diversity at genus level, especially for taxa that are not easily recognised by conventional monitoring methods, which is in accordance to other metabarcoding studies (e.g. Lindeque *et al.*, 2013; Zimmermann *et al.*, 2015; Gran-Stadniczeñko *et al.*, 2019). Besides the higher diversity in diatoms, especially dinoflagellates were more diverse than they are observed by microscopy. Additionally, more than 400 other genera throughout all size classes were detected. However, as the metabarcoding approach used is not suitable for strict species identification, morphological counts have an advantage for identification of distinct species at least if taxa are large enough to be recognised. As indicated earlier there exist certain drawbacks in methodologies. These include for example a lower sampling volume, but also not included taxa or not identified taxa due to size limitations or different morphology (Table 1). Potential solutions to tackle these issues might be identical sampling, use of replicates or even identical samples. For instance, samples used in morphological counts might be saved and later used for metabarcoding, however to our knowledge approaches like this have not been practised so far. Only recently a first study was conducted, where formalin fixed sampled were used for metabarcoding with generally promising results (Shiozaki *et al.*, 2021).

In regards to phytoplankton analysis from a raw water sample, the volume of the water sample that is used for metabarcoding analysis is higher. However sampling of zooplankton by the use of net samples results in even higher sampling volumes. Even though there is always the possibility to sample certain and especially rare taxa only by chance, a higher volume does not automatically translate in a higher diversity. In general, the needed sampling volume to cover the plankton community including rare species is also depending on the environment (Stern *et al.*, 2018). Additionally, whereas some taxa consisted of only one OTU, other taxa included several OTUs on genus level. This was the case for *Pseudocalanus* and *Calanus*, for example, which both consisted of more OTUs than other

copepods. The genus *Calanus* is suspected to include two different species (*Calanus finmarchicus* and *Calanus helgolandicus*) in the North Sea and at Helgoland (Beaugrand *et al.*, 2003; Boersma *et al.*, 2015), but the PR2 database did not identify any OTU as *C. finmarchicus.* As identification on species level is not necessarily correct due to restraints in the databases and chosen regions, it is difficult to say how many species were included in these genera. However, the high amount of OTUs hints at a further hidden diversity in certain taxa, which might be verified by metabarcoding based on specific primers with a better resolution on species level. Contrastingly, Bachy *et al.* (2013) stated that OTUs numbers, also for rare OTUs, are massively overestimated and may include large portions of artifacts. Whereas Zhan *et al.* (2014) showed, that depending on the intensity of potential artifact removal the influence on elimination of rare taxa is high. By using integrative approaches, needed thresholds can be altered depending on species coverage to balance the potential influence of amplification biases. In our case, *Thalassionema* as a common taxa according to the LTER counts, for example, was sequenced in very low abundances and would have been removed at a more conservative threshold.

In terms of comparability, both approaches have their advantages and disadvantages (see also Table 1). Besides the advantage of species identification, morphological counts also detected some genera that were not found in the dataset in this form. However, due to the high number of unidentified sequences, discrepancies in sequencing and taxonomy, it is possible that these genera were detected but not identified. One problem is the high amount of synonyms, which are not matching or updated in different databases and datasets, which might lead to believe that there exists a higher diversity on genus or species level. Another problem is a deficient phylogenetic resolution and quality in sequences, which might cause wrong annotations in databases and datasets. At the same time new sequences are found and implemented in the databases regularly, e.g. new *Chaetoceros* sequences have been implemented in the PR2 database in version 4.12.0 based on the study of Gaonkar *et al.* (2018). Other diatom species, that have not been observed via microscopy, can be identified with metabarcoding. This includes especially diatoms with small cell diameters, which are therefore difficult to identify using LM or to be detected in general (Arsenieff *et al.*, 2020). For example, Rynearson *et al.* (2020) could report new species of *Thalassiosira* in a metabarcoding study based at the Narragansett Bay Long-Term Plankton Time Series station.

Regarding long-term coverage, correlations of the microscopic counts and the metabarcoding results of distinct taxa were varying greatly. Datasets of certain taxonomic groups, such as diatoms, dinoflagellates or crustaceans were mostly not that well correlated, which might be explained by the higher diversity in the metabarcoding datasets, If

comparing only datasets of morphologic counts, these were already showing differing strength in correlation. However, in nearly all instances the microscopic datasets were better correlated to each other than to the metabarcoding datasets. For instance, datasets of the diatom *Thalassiosira* and *Phaeocystis*, which belongs to the Prymnesiophyceae, were not well correlated. Especially for genera such as *Phaeocystis*, which occur in great colonies, conducting of morphological counts is very challenging. Especially, the low sample volume that is used for the microscopic approach can cause biases if the water sample is not mixed thoroughly. In contrast to our weak long-term correlations, better short-termed correlations, that were based on automated sampling metabarcoding, were found at Helgoland for certain genera, including *Phaeocystis* (Metfies *et al.*, 2020). Other single genera were showing good or strong correlations in between the different datasets. Datasets showed good or strong correlation for the diatom *Pseudo-nitzschia*. However, visualising of the abundances throughout the three years revealed that an especially high peak in spring 2016, was not detected in the counts. As one of the OTUs was identified as P. *americana*, it is likely that it was not identified in the counts as it is morphologically different from other *Pseudo-nitzschia* genera and was only recorded at Helgoland a few years ago using TEM (Bresnan *et al.*, 2015).

Comparison of peaks in copepod abundances showed no good congruence between the counts and the metabarcoding results. However, the different sampling protocols and timing, can be a cause of these mismatches. Still, suggestions of relative abundances being related to biomass (Clarke *et al.*, 2017; Harvey *et al.*, 2017; Hirai *et al.*, 2017; Gran-Stadniczeñko *et al.*, 2019) might not be feasible in this approach either. Whilst by using net samples, the whole organism of the zooplankton is subject to DNA extraction, we did not extract e.g. any adult copepod state, that would have been visible on our filter. In general, there are still several issues, that make it questionable how the detection of taxa can be translated in or correlated to biomass and abundances (Bucklin *et al.*, 2016), as not only there is a difference in detection of uni- or multicellular components, but also in the amount of copy numbers per cell that vary in all taxa (Prokopowich *et al.*, 2003; Gong *et al.*, 2013; Fu and Gong, 2017).

Identification of similar patterns in the metabarcoding datasets was possible on any level, in both NMDS and CCA, and seasonal patterns were also observed in certain LTER datasets. However, outliers or anormal patterns found by metabarcoding, when comparing all three years, rarely fit to the LTER datasets. At the same time, most environmental parameters were found to be significantly associated to the community structure in the CCA of all datasets. Because an increase in stress on higher levels (class and phylum) was found in the NMDS analysis, we recommend to either use genus level (lowest stress values) or

order level (similar stress as OTU) to detect community patterns. A higher threshold to decrease the size of the metabarcoding dataset was sufficient to detect these patterns. If looking at significantly associated environmental parameters, only phylum level was less practical compared to the other taxonomic levels. In terms of identification of unique patterns, a lower sampling frequency (2-3 times a week) than the LTER counts (5 times a week) was sufficient. In general, we can recommend two scenarios depending on the focus of the study. A higher threshold and higher taxonomic levels than OTU or genus, e.g. order level, are sufficient for recognition of general patterns in the community. However, especially for identification of rare taxa a low threshold and a fine scaled identification of taxa is necessary.

The named drawbacks in combination with different foci and time scales complicate the recognition of patterns, especially if these might be caused by taxa that are not considered or identifiable by certain methods. The unique communities in 2018 found in the metabarcoding datasets were, for example, affected by Prymnesiophyceae such as *Emiliania* and *Phaeocystis* and by certain fungi, macroalgae and picoplanktonic green algae (Käse *et al.*, 2021b). Out of these taxa only Phaeocystis is part of the regular LTER phytoplankton counts on genus level, which explains why different patterns could be identified only by metabarcoding. This example shows, that despite all existing drawbacks in metabarcoding, it complements conventional microscopic data in long-term observations.

Conclusion

This metabarcoding study provides new insights into the diversity of plankton at Helgoland Roads and demonstrates the importance of integrative approaches in order to understand plankton communities. By implementing newly evolving methods into LTER research, currently ignored parts of the plankton community can augment the knowledge on ecosystems and their dynamics. At the same time, it is necessary to consider the current drawbacks of these methods. In terms of metabarcoding, the choice of methodology regarding primers, gene regions and databases influence the results massively. Here, we need to emphasize the high amount of unknown taxa and therefore lack of information in databases. Increasing availability of sequences in databases in combination with more in-depth studies by use of various primers focusing on different genes, are needed to increase the congruence between morphological and metabarcoding results. Also the lack of information on life cycle stages and actual cell abundances or biomass limits the comparability of the datasets. Conventional morphological approaches are therefore still needed to capture these life stage information, which are essential for understanding of ecosystems. Nevertheless, especially if databases are sufficient, the plankton diversity can

be covered by one sample in a fast and reliable way and metabarcoding reveals valuable information on hidden diversity, community patterns and food web connections even on higher taxonomic levels. Therefore, metabarcoding is a good additional source of information on plankton diversity and understanding of ecosystems.

References

Abad, D., Albaina, A., Aguirre, M., Laza-Martínez, A., Uriarte, I., Iriarte, A., Villate, F., *et al*. 2016. Is metabarcoding suitable for estuarine plankton monitoring? A comparative study with microscopy. Marine Biology, 163: 1–13.

Akaike, H. 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control, 19: 716–723.

Arsenieff, L., Le Gall, F., Rigaut-Jalabert, F., Mahé, F., Sarno, D., Gouhier, L., Baudoux, A. C., *et al*. 2020. Diversity and dynamics of relevant nanoplanktonic diatoms in the Western English Channel. ISME Journal, 14: 1966–1981.

Azam, F., Fenchel, T., Field, J. G., Gray, J. S., Meyer-Reil, L. A., and Thingstad, F. 1983. The Ecological Role of Water-Column Microbes in the Sea. Marine Ecology Progress Series, 10: 257–263.

Bachy, C., Dolan, J. R., López-García, P., Deschamps, P., and Moreira, D. 2013. 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. The ISME Journal, 7: 244–255.

Beaugrand, G., Brander, K. M., Alistair Lindley, J., Souissi, S., and Reid, P. C. 2003. Plankton effect on cod recruitment in the North Sea. Nature, 426: 661–664.

Boersma, M., Wiltshire, K. H., Kong, S. M., Greve, W., and Renz, J. 2015. Long-term change in the copepod community in the southern German Bight. Journal of Sea Research, 101: 41–50.

Bresnan, E., Kraberg, A., Fraser, S., Brown, L., Hughes, S., and Wiltshire, K. H. 2015. Diversity and seasonality of *Pseudo-nitzschia* (Peragallo) at two North Sea time-series monitoring sites. Helgoland Marine Research, 69: 193–204.

Bucklin, A., Lindeque, P. K., Rodriguez-Ezpeleta, N., Albaina, A., and Lehtiniemi, M. 2016. Metabarcoding of marine zooplankton: Prospects, progress and pitfalls. Journal of Plankton Research, 38: 393–400.

Bucklin, A., Yeh, H. D., Questel, J. M., Richardson, D. E., Reese, B., Copley, N. J., and Wiebe, P. H. 2019. Time-series metabarcoding analysis of zooplankton diversity of the NW Atlantic continental shelf. ICES Journal of Marine Science, 76: 1162–1176.

Callies, U., and Scharfe, M. 2015. Mean spring conditions at Helgoland Roads, North Sea: Graphical modeling of the influence of hydro-climatic forcing and Elbe River discharge. Journal of Sea Research, 101: 1–11.

Clarke, L. J., Beard, J. M., Swadling, K. M., and Deagle, B. E. 2017. Effect of marker choice and thermal cycling protocol on zooplankton DNA metabarcoding studies. Ecology and Evolution, 7: 873–883.

Diepenbroek, M., Glöckner, F., Grobe, P., Güntsch, A., Huber, R., König-Ries, B., Kostadinov, I., *et al*. 2014. Towards an Integrated Biodiversity and Ecological Research Data Management and Archiving Platform: The German Federation for the Curation of Biological Data (GFBio). In: Plödereder E, Grunske L, Schneider E, Ull D, editors. Informatik 2014 – Big Data Komplexität meistern. GI-Edition: Lecture Notes in Informatics (LNI) – Proceedings., Vol. 232, Bonn: Köllen Verlag; 2014, pp. 1711–1724.

Dirnböck, T., Haase, P., Mirtl, M., Pauw, J., and Templer, P. H. 2019. Contemporary International Long-Term Ecological Research (ILTER)—from biogeosciences to socio-ecology and biodiversity research. Regional Environmental Change, 19: 309–311.

Djurhuus, A., Pitz, K., Sawaya, N. A., Rojas-Márquez, J., Michaud, B., Montes, E., Muller-Karger, F., *et al*. 2018. Evaluation of marine zooplankton community structure through environmental DNA metabarcoding. Limnology and Oceanography: Methods, 16: 209–221.

Doherty, M., Costas, B., McManus, G., and Katz, L. 2007. Culture-independent assessment of planktonic ciliate diversity in coastal northwest Atlantic waters. Aquatic Microbial Ecology, 48: 141–154.

Fadeev, E., Salter, I., Schourup-Kristensen, V., Nöthig, E.-M., Metfies, K., Engel, A., Piontek, J., *et al.* 2018. Microbial Communities in the East and West Fram Strait During Sea Ice Melting Season. Frontiers in Marine Science, 5: 1–21.

Fu, R., and Gong, J. 2017. Single Cell Analysis Linking Ribosomal (r)DNA and rRNA Copy Numbers to Cell Size and Growth Rate Provides Insights into Molecular Protistan Ecology. Journal of Eukaryotic Microbiology, 64: 885–896.

Gaonkar, C. C., Piredda, R., Minucci, C., Mann, D. G., Montresor, M., Sarno, D., and Kooistra, W. H. C. F. 2018. Annotated 18S and 28S rDNA reference sequences of taxa in the planktonic diatom family Chaetocerotaceae. PLoS ONE, 13: 1–20.

Gescher, C., Metfies, K., Frickenhaus, S., Knefelkamp, B., Wiltshire, K. H., and Medlin, L. K. 2008. Feasibility of assessing the community composition of prasinophytes at the Helgoland roads sampling site with a DNA microarray. Applied and Environmental Microbiology, 74: 5305–5316.

Giner, C. R., Forn, I., Romac, S., Logares, R., de Vargas, C., and Massana, R. 2016. Environmental sequencing provides reasonable estimates of the relative abundance of specific picoeukaryotes. Applied and Environmental Microbiology, 82: 4757–4766.

Gong, J., Dong, J., Liu, X., and Massana, R. 2013. Extremely High Copy Numbers and Polymorphisms of the rDNA Operon Estimated from Single Cell Analysis of Oligotrich and Peritrich Ciliates. Protist, 164: 369–379.

Gran-Stadniczeñko, S., Šupraha, L., Egge, E. D., and Edvardsen, B. 2017. Haptophyte Diversity and Vertical Distribution Explored by 18S and 28S Ribosomal RNA Gene Metabarcoding and Scanning Electron Microscopy. Journal of Eukaryotic Microbiology, 64: 514–532.

Gran-Stadniczeñko, S., Egge, E., Hostyeva, V., Logares, R., Eikrem, W., and Edvardsen, B. 2019. Protist Diversity and Seasonal Dynamics in Skagerrak Plankton Communities as Revealed by Metabarcoding and Microscopy. Journal of Eukaryotic Microbiology, 66: 494–513.

Greve, W., Reiners, F., Nast, J., and Hoffmann, S. 2004. Helgoland Roads meso- and macrozooplankton time-series 1974 to 2004: Lessons from 30 years of single spot, high frequency sampling at the only off-shore island of the North Sea. Helgoland Marine Research, 58: 274–288.

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., Boutte, C., *et al*. 2013. The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. Nucleic Acids Research, 41: D597–D604.

Harms, J. 1993. Check list of species (algae, invertebrates and vertebrates) found in the vicinity of the island of helgoland (North Sea, German Bight) - a review of recent records. Helgoländer Meeresuntersuchungen, 47: 1–34.

Harvey, J. B. J., Johnson, S. B., Fisher, J. L., Peterson, W. T., and Vrijenhoek, R. C. 2017. Comparison of morphological and next generation DNA sequencing methods for assessing zooplankton assemblages. Journal of Experimental Marine Biology and Ecology, 487: 113–126.

Hickel, W. 1972. Kurzzeitige Veränderungen hydrographischer Faktoren und der Sestonkomponenten in driftenden Wassermassen in der Helgoländer Bucht. Helgoländer Wissenschaftliche Meeresuntersuchungen, 23: 383–392.

Hickel, W., Mangelsdorf, P., and Berg, J. 1993. The human impact in the German Bight: Eutrophication during three decades (1962-1991). Helgoländer Meeresuntersuchungen, 47: 243–263.

Hirai, J., Nagai, S., and Hidaka, K. 2017. Evaluation of metagenetic community analysis of planktonic copepods using Illumina MiSeq: Comparisons with morphological classification and metagenetic analysis using Roche 454. PLOS ONE, 12: e0181452.

Kang, W., Anslan, S., Börner, N., Schwarz, A., Schmidt, R., Künzel, S., Rioual, P., *et al*. 2021. Diatom metabarcoding and microscopic analyses from sediment samples at Lake Nam Co, Tibet: The effect of sample-size and bioinformatics on the identified communities. Ecological Indicators, 121.

Käse, L., Kraberg, A. C., Metfies, K., Neuhaus, S., Sprong, P. A. A., Fuchs, B. M., Boersma, M., *et al*. 2020a. Rapid succession drives spring community dynamics of small protists at Helgoland Roads, North Sea. Journal of Plankton Research, 42: 305–319.

Käse, L., Kraberg, A. C., Metfies, K., Neuhaus, S., Sprong, P., Fuchs, B. M., Boersma, M., *et al*. 2020b. Eukaryotic microbial community at the LTER site Helgoland Roads from March 2016 to March 2019. PANGAEA. https://doi.org/10.1594/PANGAEA.921026.

Käse, L., Metfies, K., Neuhaus, S., Boersma, M., Wiltshire, K. H., and Kraberg, A. C. 2021a. Host-parasitoid associations in marine planktonic time series: Can metabarcoding help reveal them? PLOS ONE, 16: e0244817.

Käse, L., Metfies, K., Kraberg, A. C., Neuhaus, S., Meunier, C. L., Wiltshire, K. H., and Boersma, M. 2021b. Metabarcoding analysis suggests that flexible food web interactions in the eukaryotic plankton community are more common than specific predator–prey relationships at Helgoland Roads, North Sea. ICES Journal of Marine Science, 78: 3372–3386.

Knefelkamp, B. 2009. An Investigation of the Eukaryotic Picoplankton Community in the German North Sea by Different Methods. University Bremen, Bremen.

Kraberg, A., Kieb, U., Peters, S., and Wiltshire, K. H. 2019. An updated phytoplankton check-list for the Helgoland Roads time series station with eleven new records of diatoms and dinoflagellates. Helgoland Marine Research, 73: 9.

Kröncke, I., Neumann, H., Dippner, J. W., Holbrook, S., Lamy, T., Miller, R., Padedda, B. M., *et al*. 2019. Comparison of biological and ecological long-term trends related to northern hemisphere climate in different marine ecosystems. Nature Conservation, 34: 311–341.

Lindeque, P. K., Parry, H. E., Harmer, R. A., Somerfield, P. J., and Atkinson, A. 2013. Next Generation Sequencing Reveals the Hidden Diversity of Zooplankton Assemblages. PLoS ONE, 8: e81327.

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, 3: e1420.

McManus, G. B., and Katz, L. A. 2009. Molecular and morphological methods for identifying plankton: what makes a successful marriage? Journal of Plankton Research, 31: 1119–1129.

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 microorganisms. Molecular Ecology, 19: 32–40.

Metfies, K., Gescher, C., Frickenhaus, S., Niestroy, R., Wichels, A., Gerdts, G., Knefelkamp, B., *et al*. 2010. Contribution of the class cryptophyceae to phytoplankton structure in the German bight. Journal of Phycology, 46: 1152–1160.

Metfies, K., Hessel, J., Klenk, R., Petersen, W., Wiltshire, K. H., and Kraberg, A. 2020. Uncovering the intricacies of microbial community dynamics at Helgoland Roads at the end of a spring bloom using automated sampling and 18S meta-barcoding. PLOS ONE, 15: e0233921.

Muelbert, J. H., Nidzieko, N. J., Acosta, A. T. R., Beaulieu, S. E., Bernardino, A. F., Boikova, E., Bornman, T. G., *et al*. 2019. ILTER – The International Long-Term Ecological Research Network as a Platform for Global Coastal and Ocean Observation. Frontiers in Marine Science, 6: 1–14.

Not, F., Latasa, M., Marie, D., Cariou, T., Vaulot, D., and Simon, N. 2004. A Single Species, *Micromonas pusilla* (Prasinophyceae), Dominates the Eukaryotic Picoplankton in the Western English Channel. Applied and Environmental Microbiology, 70: 4064–4072.

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., Minchin, P. R., *et al*. 2019. vegan: Community Ecology Package. https://cran.r-project.org/web/packages/vegan/vegan.pdf.

Pitsch, G., Bruni, E. P., Forster, D., Qu, Z., Sonntag, B., Stoeck, T., and Posch, T. 2019. Seasonality of planktonic freshwater ciliates: Are analyses based on V9 regions of the 18S rRNA gene correlated with morphospecies counts? Frontiers in Microbiology, 10: 1–15.

Prairie, J. C., Sutherland, K. R., Nickols, K. J., and Kaltenberg, A. M. 2012. Biophysical interactions in the plankton: A cross-scale review. Limnology and Oceanography: Fluids and Environments, 2: 121–145.

Prokopowich, C. D., Gregory, T. R., and Crease, T. J. 2003. The correlation between rDNA copy number and genome size in eukaryotes. Genome, 46: 48–50.

R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.r-project.org/.

Rynearson, T. A., Flickinger, S. A., and Fontaine, D. N. 2020. Metabarcoding Reveals Temporal Patterns of Community Composition and Realized Thermal Niches of *Thalassiosira* spp. (Bacillariophyceae) from the Narragansett Bay Long-Term Plankton Time Series. Biology, 9.

Sherr, E. B., and Sherr, B. F. 2002. Significance of predation by protists in aquatic microbial food webs. Antonie van Leeuwenhoek, 81: 293–308.

Shiozaki, T., Itoh, F., Hirose, Y., Onodera, J., Kuwata, A., and Harada, N. 2021. A DNA metabarcoding approach for recovering plankton communities from archived samples fixed in formalin. PLoS ONE, 16: 1–15.

Sprong, P. A. A., Fofonova, V., Wiltshire, K. H., Neuhaus, S., Ludwichowski, K. U., Käse, L., Androsov, A., *et al*. 2020. Spatial dynamics of eukaryotic microbial communities in the German Bight. Journal of Sea Research, 163: 101914.

Stern, R., Kraberg, A., Bresnan, E., Kooistra, W. H. C. F., Lovejoy, C., Montresor, M., Morán, X. A. G., *et al*. 2018. Molecular analyses of protists in long-term observation programmes—current status and future perspectives. Journal of Plankton Research, 40: 519–536.

Thomas, A. C., Deagle, B. E., Eveson, J. P., Harsch, C. H., and Trites, A. W. 2016. Quantitative DNA metabarcoding: Improved estimates of species proportional biomass using correction factors derived from control material. Molecular Ecology Resources, 16: 714–726.

Utermöhl, H. 1958. Zur Vervollkommnung der quantitativen Phytoplankton-Methodik. Internationale Vereinigung für Theoretische und Angewandte Limnologie (Mitteilung No. 9), 9: 1–38.

Wickham, H. 2016. ggplot2: Elegant Graphics for Data Analysis. Use R! Springer International Publishing. 260 pp. https://ggplot2.tidyverse.org.

Wiltshire, K. H., Malzahn, A. M., Wirtz, K., Greve, W., Janisch, S., Mangelsdorf, P., Manly, B. F. J., *et al*. 2008. Resilience of North Sea phytoplankton spring bloom dynamics: An analysis of long-term data at Helgoland Roads. Limnology and Oceanography, 53: 1294–1302.

Wiltshire, K. H., Kraberg, A., Bartsch, I., Boersma, M., Franke, H. D., Freund, J., Gebühr, C., *et al*. 2010. Helgoland roads, North Sea: 45 years of change. Estuaries and Coasts, 33: 295–310.

Wollschläger, J., Nicolaus, A., Wiltshire, K. H., and Metfies, K. 2014. Assessment of North Sea phytoplankton via molecular sensing: A method evaluation. Journal of Plankton Research, 36: 695–708.

Xiao, X., Sogge, H., Lagesen, K., Tooming-Klunderud, A., Jakobsen, K. S., and Rohrlack, T. 2014. Use of High Throughput Sequencing and Light Microscopy Show Contrasting Results in a Study of Phytoplankton Occurrence in a Freshwater Environment. PLoS ONE, 9: e106510.

Yang, J., Zhang, X., Xie, Y., Song, C., Zhang, Y., Yu, H., and Burton, G. A. 2017. Zooplankton community profiling in a eutrophic freshwater ecosystem-lake tai basin by DNA metabarcoding. Scientific Reports, 7: 1–11.

Yilmaz, P., Kottmann, R., Field, D., Knight, R., Cole, J. R., Amaral-Zettler, L., Gilbert, J. A., *et al*. 2011. Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIxS) specifications. Nature Biotechnology, 29: 415–420.

Zhan, A., Xiong, W., He, S., and MacIsaac, H. J. 2014. Influence of artifact removal on rare species recovery in natural complex communities using high-throughput sequencing. PLoS ONE, 9.

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 Microbiology Ecology, 52: 79–92.

Zimmermann, J., Glöckner, G., Jahn, R., Enke, N., and Gemeinholzer, B. 2015. Metabarcoding vs. morphological identification to assess diatom diversity in environmental studies. Molecular Ecology Resources, 15: 526–542.

8 Discussion

The LTER underway at Helgoland Roads monitors over 70 genera of phyto- and zooplankton (chapter 7). Many microphytoplankton are recorded there at species level. Furthermore, the phytoplankton checklist alone contains over 230 known taxa (diatoms and dinoflagellates) around the island of Helgoland (Kraberg *et al.*, 2019). Numerous additional taxa have been registered for zooplankton (Greve *et al.*, 2004). Different size groups of flagellates and coccolithophorids are regularly counted at Helgoland Roads (chapter 7). However, only a few small-sized eukaryotic microbes such as *Phaeocystis* are counted at genus level. Due to everyday logistical constraints, it has not been possible to have studies record all taxa in every sample, and so there have been studies of individual taxon groups over the past fifteen years, including studies of pico- and nanoplanktonic communities and individual taxa (Medlin *et al.*, 2006; Gescher *et al.*, 2008; Knefelkamp, 2009; Metfies *et al.*, 2010, 2020); bacteria (Sapp *et al.*, 2007; Lucas *et al.*, 2015, 2016; Teeling *et al.*, 2016; Chafee *et al.*, 2018); fungi (Banos *et al.*, 2020); and micro- and mesozooplankton (Löder *et al.*, 2012; Yang *et al.*, 2015, 2021). New species have regularly been observed and described by light and electron microscopy (Hoppenrath, 2004; Kraberg *et al.*, 2018, 2019) or by those methods coupled with molecular methods (Thines *et al.*, 2015; Buaya *et al.*, 2017). Despite the great plankton diversity already discovered around Helgoland, it is clear that covering the entire planktonic diversity in the course of the daily morphological time series is logistically and financially impossible. This is mostly because the Utermöhl technique as well as the optical limitation of light microscopy would make it unrealistically time consuming to perform observation and counting on a large enough sample (Utermöhl, 1958; Paxinos and Mitchell, 2000; Edler and Elbrächter, 2010).

This thesis used 18S metabarcoding for plankton analysis at the Helgoland Roads LTER site in the North Sea over a three-year period. The main aim was to analyse the diversity and general plankton community structure throughout the year of the typically understudied plankton compartments (aim 1 of this thesis). Chapter 4 focused on identifying the eukaryotic microbial community during spring. We also used metabarcoding to identify parasitoids and investigate potential host–parasitoid relationships in the plankton community at Helgoland Roads (chapter 5) over three years. Using the same dataset, we then focused on identifying predator-prey interactions and associations in the food web by conducting network analysis (chapter 6)**.** Chapters 4 through 7 also looked for associations with environmental parameters (chapter 4 to 7). Phyto- and zooplankton diversity and abundance is monitored several days a week at the Helgoland Roads LTER using

conventional microscopic analyses, and this unique high-frequency dataset allowed us to compare datasets at different taxonomic levels (aim 2 of this thesis, chapters 4 and 7).

What is the specific added value of using metabarcoding to study the ecology of the eukaryotic plankton community?

The following sections discuss the outcomes of aims 1a–d to emphasize the added value of using metabarcoding in plankton ecological research.

Greater detection of biodiversity and understudied plankton components

One of the first aims of this thesis was to evaluate in depth the diversity of the plankton community (aim 1a of this thesis). Pico- and nanoplankton are important components of the microbial food web (Azam *et al.*, 1983; Sherr and Sherr, 2002; Caron and Hu, 2019) but remain underrepresented in studies of phytoplankton diversity (Stern *et al.*, 2018). Plankton research also commonly ignores parasitoids, which can kill their respective hosts (Huxham *et al.*, 1995; Marcogliese, 2001). These underrepresented compartments of the plankton community may play an important role in the food web at Helgoland.

High plankton diversity was found at all trophic levels throughout the whole three-year study period (chapters 4 to 7); but while a total of over 530 genera were identified, more than half of the OTUs could not be identified at genus level (chapter 7). Chapter 5 discusses findings of high diversity among parasitoids (10 different phyla) with potential hosts throughout the food web. For details, see the discussion of aim 1c in the section entitled “Detection of parasitoids as understudied components of marine plankton” below.

Chapter 4 focuses on small-sized eukaryotic microbes as it examines the diversity of the plankton community during spring 2016. The sequence assemblage included various OTUs identified as taxa of these size classes. These findings thus expand the traditional view that diatoms dominate the spring community at Helgoland. While diatoms may constitute the bulk of the phytoplankton biomass under the microscope (Wiltshire *et al.*, 2015), the high diversity and abundance of other taxa found to be present on the basis of metabarcoding should not be ignored. Some of these are capable of blooming themselves (e.g. *Phaeocystis*, *Emiliania*) or of influencing the spring blooms of various diatoms, for example via parasitic infections (Metfies *et al.*, 2020) or grazing; and as past observations have shown (Leterme *et al.*, 2006; Boersma *et al.*, 2015) small environmental changes can affect how certain plankton contribute to the community composition at Helgoland and in the North Sea in general. The great diversity of the pico- and nanoplanktonic taxa indicates that they

cannot be treated as a single plankton compartment, because each can have a different impact on the food web.

The abundance of non-autotrophic taxa in our dataset is remarkable. So far, the abundance of so-called phytoflagellates as well as blooms of *Phaeocystis* and *Emiliania*, which were highly abundant in the metabarcoding dataset (chapters 4 to 7), parallel observations from the Western Channel during spring (Widdicombe *et al.*, 2010) or from Eastern Fram Strait during summer (Nöthig *et al.*, 2015). High abundances of *Phaeocystis* can heavily influence the food web dynamics at Helgoland. *Phaeocystis* is known to form gelatinous colonies, and it produces deterring chemicals, that make the colonies resistant to grazing by small-sized copepods. Larger copepods or microzooplankton such as ciliates and dinoflagellates, however, can feed on *Phaeocystis* (Hamm, 2000; Stelfox-Widdicombe *et al.*, 2004; Schoemann *et al.*, 2005). In regards to the parasitoid, heterotrophic and mixotrophic taxa that were found (chapters 4 to 7), other metabarcoding studies also revealed high abundances in different regions (Marquardt *et al.*, 2016; Piredda *et al.*, 2017; Gran-Stadniczeñko *et al.*, 2019). Heterotrophic pico- and nanoplankton included inter alia filter-feeding choanoflagellates (Dayel and King, 2014) and MAST, which both graze on bacteria and detritus and in turn serve as prey for larger consumers (Massana *et al.*, 2004, 2009; King, 2005; Logares *et al.*, 2012). MAST are furthermore known to include potential parasites and epiphytic taxa (Gómez *et al.*, 2011); however, as they are still poorly known as a group, there is a lack of knowledge about their interactions with other plankton groups, including prokaryotes (Lin *et al.*, 2021). Heterotrophic picoplankton such as the Picozoa, which were found in nearly every sample, are another example of potentially important components of the microbial eukaryotic plankton. Upon their detection, Picozoa were termed Picobiliphyta because they were assumed to be autotrophs; but they ultimately turned out to be heterotrophs (Moreira and López-García, 2014) and so played a different role in the food web. Cryptophyta, as a potentially mixotrophic food source for microzooplankton, had previously been identified as important and constant contributors to the pico- and nanoplankton community in the German Bight and at Helgoland (Metfies *et al.*, 2010; Medlin *et al.*, 2017), and they were found during spring 2016 (chapter 4) as well as during all three years (chapter 6). The high diversity of potential prey for microzooplankton emphasizes the need for a better understanding of this compartment of the food web. The presence of microzooplankton found at Helgoland was continuous and highly diverse (chapters 4 and 6), which supports the notion of the importance of these consumers in the marine realm (Landry and Calbet, 2004; Sherr and Sherr, 2007; Yang *et al.*, 2021). Influences on phytoplankton bloom formation and duration are to be expected (Montagnes and Lessard, 1999; Sherr and Sherr, 2002; Aberle *et al.*, 2007); using

microscopy, Löder *et al*. (2012) showed that dinoflagellates are an important contributor to the food web at Helgoland Roads. Yang *et al.* (2021) furthermore emphasized that dinoflagellates are a substantial grazer in late summer and autumn. This is reinforced by the metabarcoding results, where dinoflagellates of various trophic modes (mixotrophic, heterotrophic and parasitic) were found in highest abundances and great diversity.

Community structure and associations with environmental conditions

Generally, environmental parameters such as temperature and salinity influence the plankton communities on different spatial and temporal levels and plankton diversity and biomass are not steady throughout the year (Wiltshire *et al.*, 2010, 2015; Scharfe and Wiltshire, 2019).

Previously reported communities, which at all times have included a certain number of species, were observed in the marine environment and especially in the picoplanktonic fraction (Jacquet *et al.*, 2002; Zhu *et al.*, 2005; Knefelkamp, 2009). Here, it was possible to identify a community whose taxa are always present (with fluctuations in abundance) across all planktonic trophic levels. At all times during spring 2016, the community included the previously mentioned highly abundant pico- and nanoplanktonic taxa as well as some dinoflagellates and diatoms (chapter 4). By prolonging the metabarcoding dataset out to three years and including bigger organisms such as Metazoa, a community that includes the same 295 OTUs at all times occurs not only during spring but also continuously throughout the years (chapter 6). While this community included taxa of all trophic levels (such as crustacean taxa or heterotrophic and mixotrophic dinoflagellates), it also included small-sized taxa in the pico- and nanoplankton range. Due to the continuous presence of consumers (without knowledge of their life stages) in the metabarcoding dataset (chapters 6 and 7), zooplankton does not display the previously described progressions of peaks in phytoplankton followed by peaks in zooplankton abundance (Lewandowska *et al.*, 2015; Wiltshire *et al.*, 2015).

Sprong *et al.* (2020) – the abstract of this related publication is appended to this thesis – compared metabarcoding data from two coastal stations in the German Bight with the plankton community at Helgoland during spring 2016 (chapter 4). The plankton community in the two coastal areas was not dominated by dinoflagellates, but the plankton community at Helgoland, in contrast, was. Wollschläger *et al*. (2015) also reported a community dominated by dinoflagellates in offshore areas in the German Bight while coastal area communities were dominated by diatoms. Similar community structures – diatoms more abundant at the coast and dinoflagellates more abundant away from it – were also reported for the Dutch coast by Alvarez-Fernandez and Riegman (2014). The plankton community

at Helgoland faced a higher salinity and lower nutrient concentrations compared to the other stations (Sprong *et al.*, 2020). Salinity at Helgoland Roads has been increasing for several decades (Wiltshire *et al.*, 2010; Scharfe and Wiltshire, 2019), and further shifts in salinity ranges may therefore influence the composition of the plankton community.

Just from looking at the plankton communities at Helgoland, it is possible to describe several associations with environmental parameters. Chapter 4, which was based on a dataset that covered a single season, revealed temperature as the most important environmental parameter, followed by silicate concentrations, salinity, sunshine duration and tide. Especially temperature and salinity were linked to the beta diversity patterns found in the spring plankton community. Temperature is commonly one of the key parameters for marine life (Andersson *et al.*, 1994), and the rapidly warming oceans are affecting marine life tremendously (Hughes *et al.*, 2019). The effects of temperature have been seen all across the phytoplankton (Karentz and Smayda, 1984), and the use of metabarcoding has also detected its influence on individual diatom genera (Canesi and Rynearson, 2016; Rynearson *et al.*, 2020). Temperature has been named as a dominating factor for the diatom phytoplankton community at Helgoland Roads (Wiltshire *et al.*, 2015), and furthermore temperatures at Helgoland have been increasing rapidly (Wiltshire *et al.*, 2010; Scharfe and Wiltshire, 2019).

This thesis can only make statements about the seasonality of the plankton, because the durations of the studies in this thesis do not permit statements to be made about the influence of climate change on the plankton communities. Seasonality was found to exert a strong influence throughout the entire three years; indeed, the factor “season” exhibited the most important association with community structure during this period (chapters 5, 6 and 7). Other significant parameters included temperature, silicates, salinity, and sunshine duration. Additionally, our diversity indices suggested that the autumn community was most diverse, even more so than the spring community (chapter 6). Similar seasonal patterns or associations with the plankton community have been reported for other areas (Abad *et al.*, 2016; Berry *et al.*, 2019; Gran-Stadniczeñko *et al.*, 2019). Gran-Stadniczeñko *et al.* (2019) also reported a higher diversity in late summer and autumn, which they linked to the seasonal influence of the North Atlantic.

Detection of parasitoids as understudied components of marine plankton

Planktonic parasitoids, which often complete their life cycle in a single host (Rasconi *et al.*, 2011), might infect various other plankton or other organisms in the marine realm (Skovgaard, 2014). They kill their hosts as they complete their life cycles and can therefore greatly influence the plankton community, for instance by altering bloom duration (Park *et*

al., 2004; Chambouvet *et al.*, 2008; Peacock *et al.*, 2014). Tillmann *et al*. (1999) suggested there was competition between parasitic infections on phytoplankton and zooplankton in the North Frisian Wadden Sea because the parasitoids kill the cells, making them unavailable to higher trophic levels such as mesozooplankton. At Helgoland and generally in the North Sea, various plankton parasitoids have been also described using light microscopy, transmission electron microscopy (TEM), and scanning electron microscopy (SEM) (Drebes *et al.*, 1996; Kühn *et al.*, 1996; Schweikert and Schnepf, 1997; Thines *et al.*, 2015; Buaya *et al.*, 2017). Parasitoids of phyto- and zooplankton have typically been ignored or at least not identified in detail in regular plankton studies, including in the Helgoland Roads time series; this is because identification is very time-consuming, and their complex life cycles mean not all infection stages are easily recognizable if they are known at all.

Chapter 5 unravelled – for the first time in detail – the diversity of protistan parasitoids at Helgoland Roads using metabarcoding. A time series of freshwater parasites (Beng *et al.*, 2021) revealed a similar proportion of parasites in the plankton community, finding 14.4% amplicon sequence variants as parasites compared to this thesis's finding of 10.2% and 16.5%, respectively, depending on the dataset thresholds (chapter 5). A wide diversity of parasitoids of taxa belonging to ten different phyla was observed, which hints at infections throughout the food web. The relative abundance, richness, and community composition of parasitoids also changed both seasonally and annually (chapter 5). Some taxa occurred in high abundance either throughout the year or only during certain periods. While this thesis was unable to clearly correlate parasitoid occurrences to any particular season, the highest relative abundance of parasitoids occurred during summer. Morphological counts of coastal waters in the Mediterranean Sea have resulted in the same kind of finding, given that the potential zooplanktonic hosts were most abundant in summer (Skovgaard and Saiz, 2006); and the parasitoids known to appear in the North Sea or at Helgoland were also isolated and reported during summer and autumn (Drebes *et al.*, 1996; Kühn *et al.*, 1996; Schweikert and Schnepf, 1997; Thines *et al.*, 2015; Buaya *et al.*, 2017).

Five phyla included parasitoids with diatoms as known potential hosts (chapter 5). These included two phyla that were highly abundant and that were found with a high number of OTUs, namely Cercozoa (e.g. Cryomonadida) and Stramenopiles (e.g. Oomycetes). It has been suggested that one of these oomycete parasitoids, *Miracula helgolandica*, which has been described at Helgoland (Buaya *et al.*, 2017), possibly affected the termination of a diatom bloom of *Pseudo-nitzschia pungens* there in May 2016 (Metfies *et al.*, 2020). As chapter 5 discusses, co-occurrences of this host–parasitoid pairing were found in 2016, but an even higher abundance of the parasitoid was found in 2018. While both taxa occurred in similar temperature ranges and the diatom was present in all three years, the parasitoid

was not observed in 2017, which coincided with a previous period of lower salinity. It remains unclear why the parasitoid was absent during that year; one reason might have been the influence of different water masses, which might also explain the differences in salinity. Another reason might have been that the timing of the parasitoid life cycle and the short occurrence of *P. pungens* did not coincide, rendering infection impossible.

Known host–parasitoid systems were identified in two different cases (chapter 5) in which individual parasitoid taxa could be linked to potential hosts as described in the literature. The first case included alternating associations between peaks in potential host and parasitoid abundances that were identified for several known pairings (*Rhizosolenia imbricata*/*Olpidiopsis drebesii*, *Coscinodiscus* sp./*Lagenisma coscinodisci*, *Guinardia* sp./*Cryothecomonas aestivalis, Akashiwo* sp./*Parvilucifera prorocentri*, and *Eucampia zodiacus*/Cercozoa). The second example showed simultaneous rather than alternating occurrences, interpreted as a sign of current infection. This situation was identified for several pairings: *Rhizosolenia imbricata*/*Olpidiopsis drebesii*, *Pseudo-nitzschia pungens*/*Miracula helgolandica*, *Coscinodiscus* sp./*Lagenisma coscinodisci*, *Guinardia* sp./Cryomonadida and *Pirsonia* clade, *Eucampia zodiacus*/Cercozoa and Syndiniales genera/Crustacea & Tintinnida. However, mostly due to variability in parasitoid occurrence and abundance, it was not possible to identify potential new host–parasitoid systems (chapter 5). The co-occurrence network analysis discussed in chapter 6 also revealed various associations between parasitic and other planktonic taxa. Syndiniales, for example, were found in connection to other dinoflagellate genera, but also to Metazoa, Chlorophyta, Ciliophora, and Ochrophyta, including diatoms. So far, Syndiniales are known as potential parasitoids of several taxa such as Radiolaria, Dinoflagellata, Ciliates, Cnidaria, Chaetognatha, and Crustacea such as Copepoda and Amphipoda, as well as of fish eggs (Coats, 1999; Chambouvet *et al.*, 2008; Guillou *et al.*, 2010; Skovgaard, 2014; Lima-Mendez *et al.*, 2015; Cleary and Durbin, 2016). Anderson and Harvey (2020) also suggested that host–parasitoid infections of Syndiniales display flexible dynamics. The authors conducted co-occurrence network analysis and revealed several associations of Syndiniales and potential hosts suggested as concurrent infections. Therefore, it is highly likely that these parasitoids display flexible infection dynamics, which makes predicting their influence on food web dynamics even more complex. So far, diatoms have not been known to be susceptible to infection by Syndiniales; however, other studies have also found positive and negative associations between these two groups which hints at potential infections in cases of positive associations (Anderson and Harvey, 2020; Sassenhagen *et al.*, 2020; Vincent and Bowler, 2020; Zamora-Terol *et al.*, 2021). These Syndiniales, already found in high abundance during spring 2016 as discussed in chapter 4, were abundant at

all times, and they contributed the most OTUs (chapter 5). A recent metabarcoding study, which compared abundances of Syndiniales in water samples and net samples, has even revealed a higher relative abundance of Syndiniales in zooplanktonic organisms such as copepods than in water (Zamora-Terol *et al.*, 2021). The authors suggested two reasons why various associations of Syndiniales with the zooplankton had been detected: either an ongoing infection of zooplankton hosts, or indirect uptake by the zooplankton due to them having fed on other infected prey.

One of the three phyla found in chapter 5 that are known to infect dinoflagellates were Perkinsea. Parasitoids of the genus *Parvilucifera* are known to infect different dinoflagellate species (Park *et al.*, 2004). *Parvilucifera* do not actively select their host but infect potential hosts randomly instead (Alacid *et al.*, 2016). At the same time, in some cases, they seem to prefer some species over others, and they are able to infect dominating dinoflagellates in mixed dinoflagellate communities as well as influence the community structure (Alacid *et al.*, 2016) since they ultimately kill their hosts. Reñé *et al*. (2021) studied dinoflagellate bloom and no-bloom periods to explore the host–parasitoid interactions of different Perkinsea species. They suggested that even though Perkinsea have several potential hosts, only blooms of preferred hosts provide optimal conditions for the parasitoid to reproduce, which might be one of the reasons why Perkinsea were not as abundant as other parasitoids even though dinoflagellates, as potential hosts, were always present. The overall large number of potential interactions underscores the importance of considering parasitism in food webs.

Detection of food web connections

Understanding the complexity of the marine food web depends on discoveries of plankton interactions. Due to the great diversity of plankton, interactions include competition for resources, grazing pressure (predation), and infections by and of other planktonic species (Hutchinson, 1961; Rhodes *et al.*, 2008; Wiltshire *et al.*, 2010; D'Alelio *et al.*, 2016).

The high variability of and numerous potential connections between plankton taxa (chapter 6) reflect the complexity of the food web at Helgoland. If too many taxa are combined into trophic compartments, and if hard-to-detect interactions such as microbial cross-feeding, auxotrophy, and mixotrophy in the plankton food web are ignored, it is highly likely that food web dynamics and important processes within the plankton community will be overlooked (Millette *et al.*, 2018). In recent years, the number of studies that use metabarcoding data has increased; in general, they are applying network analysis and food web models to plankton data in order to investigate potential food web connections (D'Alelio *et al.*, 2016; Chafee *et al.*, 2018; Zamora-Terol *et al.*, 2020; Bolaños *et al.*, 2021; Kobari *et al.*, 2021;

Novotny *et al.*, 2021; Suter *et al.*, 2022). These studies have indicated that connections in the food web vary according to changes in the marine environment. Zamora-Terol *et al.* (2020) investigated food web interactions by metabarcoding water samples alongside copepod gut content. Copepods were adapting their preferred prey by season depending on which prey species were available. While some copepods fed on the most easily available prey species, others consumed highly diverse prey and might qualify as opportunistic feeders. Some prey was fed upon by more than one species while other prey was specific to a single species, which enables species to coexist even when competition is high (Zamora-Terol *et al.*, 2020). Network analysis of the metabarcoding data as discussed in chapter 6 showed similar results. The various copepod species showed highly diverse associations, including several with pico- or nanoplanktonic taxa. Additionally, some species exhibited a greater number of or more robust associations with specific taxa that are known to be prey, such as diatoms, dinoflagellates, or ciliates. It is also known that copepods may feed selectively depending on their life-cycle stage (Meunier *et al.*, 2016); however, the present metabarcoding approach is not capable of distinguishing life stages. The consumption of some taxa might be accidental or the result of opportunistic feeding on detritus and particulate matter (Zamora-Terol *et al.*, 2020). As discussed in chapter 6, while it was possible to show co-occurrences and potential associations of species, the proportional consumption of potential prey remains unclear since no gut content was analysed. This is also emphasized by proportional differences in the prey evident in water samples as compared to gut samples (Kobari *et al.*, 2021). Therefore, network analysis should be conducted to investigate potential species associations and further investigations should shed light on which ones are actually based on predator-prey dynamics.

Is the information obtained through metabarcoding comparable to that of conventional counting methods?

This section focuses on the comparability of metabarcoding datasets and those of conventional morphological counts. Morphological analysis such as light microscopy and metabarcoding each have several advantages and disadvantages (see also chapter 7, Table 1). While the choice of method ought to depend on the focus of any plankton ecological study, morphological and metabarcoding methods are both useful in assessing biodiversity in general.

Advantages of morphological analysis are its relatively low cost and ability to identify distinct species as well as their life stages and potential parasitic infections; but the method is time-consuming and based on only a small sample volume, so it disregards rare species and requires good taxonomic knowledge on the part of the investigator (Utermöhl, 1958; Not *et*

al., 2004; Edler and Elbrächter, 2010; Stern *et al.*, 2018). Including electron microscopy allows the identification and description of new species or certain particular taxa (Schweikert and Schnepf, 1997; Johnson and Martiny, 2015; Kraberg *et al.*, 2019). However, by incorporating different approaches into morphological counts it is also possible to increase the number of planktonic species the study can cover. It is possible to gain more detailed insights into the plankton community by increasing expertise on certain taxa or summarizing certain taxa as morphotypes; this has proved true for ciliates and dinoflagellates at Helgoland Roads (Löder *et al.*, 2012; Yang *et al.*, 2014). Similarly, the course of this study included testing the analysis of small-sized morphotypes (chapter 7); however, for most groups no clear correlation to metabarcoding results could be found.

Metabarcoding can be seen as a promising alternative way to determine plankton diversity using large sample volumes (several litres of water) that have the potential to capture rare species. Regardless of cell size, metabarcoding has no trouble identifying species; but it also comes with a range of drawbacks (Table 1 in chapter 3) that need to be carefully evaluated in choosing metabarcoding. These include among other things the inability to distinguish life stages (Bucklin *et al.*, 2016), biases in the molecular analysis, and insufficient resolution of some genes as a means of distinguishing taxa especially at species level (de Vargas *et al.*, 2015; Wangensteen *et al.*, 2018). Unidentified sequences and gaps in database coverage of some taxon groups impede identification of certain taxonomic groups for a large proportion of metabarcoding datasets; and this was also the case here (chapters 4 to 7). The utility of databases depends on good sequence quality (it is crucial to avoiding wrong annotations), and new sequences need to be implemented regularly (Guillou *et al.*, 2013; del Campo *et al.*, 2018; Gaonkar *et al.*, 2018). Metabarcoding datasets display the proportion of sequences found, and this can vary greatly depending on the number of gene copies found in individual cells in species and depending on what region of the genome is amplified (LaJeunesse *et al.*, 2005; Connolly *et al.*, 2008; Hong *et al.*, 2016), but they cannot provide exact information on actual cell abundances or biomass, which limits the comparability of their results to those of cell counting methods such as light microscopy.

Metabarcoding is still evolving; its methodological drawbacks are being eliminated and new procedures and pipelines are being developed. While currently more cost-intensive, new methodological developments in metabarcoding are helping to lower costs. This thesis has suggested some initial (potential) solutions to these drawbacks (chapter 3, Table 1 and chapter 7, Table 1). One way to overcome the main problems of methodological and organismal bias is to include mock samples and correction factors in the analysis. However, especially for eukaryotes but sometimes even for prokaryotes the development of correction factors has been limited too (Kembel *et al.*, 2012; Louca *et al.*, 2018; Gong and Marchetti,

2019; Starke *et al.*, 2021; Yarimizu *et al.*, 2021). A better alternative might be to target alternative gene regions that occur in one copy per cell. For example, Pierella Karlusich *et al.* (2022). introduced a new approach using the photosynthetic gene psbO; the gene appears mainly in one copy per cell and is present in photosynthetic eukaryotes and prokaryotes. Another solution might be to implement multi-marker approaches (Shi *et al.*, 2011; Stefanni *et al.*, 2018; McNichol *et al.*, 2021; Sildever *et al.*, 2021; Yeh *et al.*, 2021); these allow an even broader detection of diversity in the plankton community. Such approaches have been implemented in comparative studies before and have also been suggested for use in long-term monitoring (Stefanni *et al.*, 2018; Berry *et al.*, 2019; Sildever *et al.*, 2021). Metabarcoding still covers life stages that can escape detection through microscopy even if distinguishing individual stages is impossible. As described above, metabarcoding also enables identification of pico- and nanoplankton as well as regularly overlooked groups such as parasitoids.

A recent meta-analysis of aquatic community assessment studies that included both conventional and various metabarcoding approaches revealed that comparability depends on key groups (Keck *et al.*, 2022). The authors summarized that comparability suffices for studies of fish but not necessarily for studies of macroinvertebrates, microphytobenthos, or plankton. Keck *et al*. (2022) therefore stated that metabarcoding in these groups may provide complementary information beyond what conventional methods can obtain. As for plankton research, various studies have focused on comparing metabarcoding results with datasets gleaned from various methods including morphological identification by microscopy, flow cytometry, or even datasets on mock samples (Doherty *et al.*, 2007; Bachy *et al.*, 2013; Lindeque *et al.*, 2013; Wollschläger *et al.*, 2014; Xiao *et al.*, 2014; Zimmermann *et al.*, 2015; Abad *et al.*, 2016; Giner *et al.*, 2016; Harvey *et al.*, 2017; Hirai *et al.*, 2017; Yang *et al.*, 2017; Clarke *et al.*, 2017; Gran-Stadniczeñko *et al.*, 2017; Djurhuus *et al.*, 2018; Pitsch *et al.*, 2019; Bucklin *et al.*, 2019; Gran-Stadniczeñko *et al.*, 2019; Metfies *et al.*, 2020; Bailet *et al.*, 2020; Kang *et al.*, 2021; Semmouri *et al.*, 2021). In general, comparability depends on where the focus of the ecological assessment lay. The level comparability of datasets can be distinguished according to two particular foci: comparability of diversity and comparability of species abundance or biomass.

Generally, metabarcoding is known to show greater diversity than morphological analysis (Doherty *et al.*, 2007; Lindeque *et al.*, 2013; Xiao *et al.*, 2014; Zimmermann *et al.*, 2015; Clarke *et al.*, 2017; Gran-Stadniczeñko *et al.*, 2017; Yang *et al.*, 2017; Gran-Stadniczeñko *et al.*, 2019). Chapter 7 revealed a similar taxonomic richness for diatoms, which are well monitored at Helgoland (Wiltshire *et al.*, 2008, 2015; Gebühr *et al.*, 2009; Mieruch *et al.*, 2010; Freund *et al.*, 2012; Schlüter *et al.*, 2012; Kraberg *et al.*, 2019; Scharfe and Wiltshire,

2019); however, species composition varied depending on the dataset. Comparison of other taxa revealed greater diversity in the metabarcoding dataset than in the LTER datasets, but some taxa found in the LTER were not found in the metabarcoding dataset (chapter 7). Both methods also include high fractions of unidentified species. For the most part, their small size is what makes these species unidentifiable in morphological analysis; but metabarcoding also fails to identify them due to databases that either badly resolve or wholly lack the particular sequences. Similar discrepancies in common taxa have been reported in other areas including for freshwater and benthic taxa (Abad *et al.*, 2016; Kang *et al.*, 2021; Nistal-García *et al.*, 2021; Santi *et al.*, 2021).

The large differences in methodology limit the comparability of metabarcoding datasets with actual cell counts or biomass measurements. So far, significant positive correlations of biomass to read abundances have been found in studies using zooplankton net samples (Harvey *et al.*, 2017; Hirai *et al.*, 2017). Other studies have correlated cell counts of different plankton taxa to read abundances using various 18S regions (Wollschläger *et al.*, 2014; Giner *et al.*, 2016; Bucklin *et al.*, 2019; Metfies *et al.*, 2020). Significant correlations have been reported for different taxa such as choanoflagellates and the diatom *Chaetoceros* sp. (Metfies *et al.*, 2020), Calanoida, Gastropoda, and Chaetognatha (Bucklin *et al.*, 2019) as well as for *Minorisa minuta* (Rhizaria), Pelagophyceae, and *Micromonas* spp. (Chlorophyta) (Giner *et al.*, 2016). However, such correlations were not consistent throughout all taxa (Giner *et al.*, 2016; Bucklin *et al.*, 2019; Metfies *et al.*, 2020), and they depended on different gene regions (Giner *et al.*, 2016). Differences in diversity are especially capable of heavily influencing comparability in terms of abundances. General patterns in the plankton community and plankton succession might contrast between the methods (chapter 7), making comparison difficult or even impossible in some cases. This depends heavily on the much higher diversity in metabarcoding and on genera that are not observed by morphological analysis. Correlations between LTER and metabarcoding datasets over three years were missing or weak for diatoms, dinoflagellates, and Crustacea (chapter 7); these particular correlations are of morphological cell counts with relative abundances in the metabarcoding dataset. The weakness or absence of these correlations could be linked to the greater diversity captured for these groups, because the strength of a correlation might increase if certain taxa that were not included in morphological counts were removed from the metabarcoding dataset (e.g. Syndiniales in dinoflagellates).

Differences in abundances are especially salient for dinoflagellates. High dinoflagellate abundances were found at Helgoland throughout all three years (chapters 4 to 7) whereas stations closer to the coast did not display such a dominance (Sprong *et al.*, 2020). Such a dominance or a higher ratio of dinoflagellate reads in metabarcoding has been reported

elsewhere (LaJeunesse *et al.*, 2005; Massana, 2011; Taylor and Cunliffe, 2014; Massana *et al.*, 2015; Hong *et al.*, 2016; Chen *et al.*, 2021) and been linked mostly to high copy numbers in dinoflagellate taxa or the greater diversity revealed by metabarcoding. In terms of dinoflagellates, correlations depended on the genus (chapter 7). For *Akashiwo* for instance, correlations between methods were good; but for *Gyrodinium*, correlations were weak. There may be several reasons for these differences, including difficulties in morphological analysis of certain genera, generally large differences in copy numbers especially in dinoflagellates (Hong *et al.*, 2016), or different coverage of individual taxa in the databases.

When it comes to copepod abundances, it might not be feasible to compare morphological analysis of net samples with metabarcoding of environmental DNA in water samples. Copepod datasets matched poorly throughout all three years (chapter 7). Summer peaks of *Oithona* and other genera were captured; however, metabarcoding for three full years showed more peaks in abundance – and suggested different genera were more abundant – than what the LTER dataset showed and suggested. Studies focusing on the abundance or biomass of zooplankton communities might be more easily comparable if they are both based on the same sampling procedure (such as standardized net sampling). Previous studies based on zooplankton bulk samples have suggested a possible correlation between sequence abundances and plankton biomass (Hirai *et al.*, 2017; Gran-Stadniczeñko *et al.*, 2019) or abundances (Bucklin *et al.*, 2019). In the case of zooplankton, for instance, correlations between biomass and sequence abundances varied depending on taxa and sometimes were even gene-specific (Harvey *et al.*, 2017). Either way, one should expect there to be discrepancies in correlations, because each method has different drawbacks: copy number issues in metabarcoding versus undistinguishable life cycle stages in morphological analysis and need to be carefully discerned.

Overall, how well the datasets corresponded depended on taxonomic levels and the biodiversity of the plankton. If contrasting results are found, it adds weight to the question of whether the respective methods may mask potentially important details and other kinds of information. Implementing multi-marker approaches, which can reliably cover different compartments of the food web, seems especially capable of increasing our knowledge of all trophic levels in the plankton community. So to provide the best added value, it is reasonable to seek to develop techniques to integrate different datasets. Large, detailed datasets, such as are available at Helgoland Roads – including the dataset this thesis presents – have the potential to be very helpful in the future in this regard.

Is it feasible and scientifically reasonable to integrate metabarcoding into an existing long-term time series, and how can results of different methods be combined for the best added value in long-term time series?

Molecular methods have already been recognized as generally useful and beneficial (Stern *et al.*, 2018; Santoferrara, 2019).

While there are various reasons to implement metabarcoding in long-term series, metabarcoding and conventional morphological identification can produce non-compatible results, and therefore any implementation of metabarcoding requires careful evaluation. Given the speed at which metabarcoding is evolving and finding applications in plankton research, it could have a fundamental role to play in future long-term ecological research. Developing and employing automated systems in particular could simplify and accelerate the use of metabarcoding in plankton research. In the following, I will emphasize a few methodological drawbacks that future research should address and emphasize the vast advantages of putting metabarcoding to use in ecological plankton research.

One of the main reasons for implementing metabarcoding is the fact that it has the power to observe greater diversity, capturing not only small-sized taxa that cannot be identified with microscopy but also rare taxa which are more easily overlooked in the smaller sample volumes used for conventional microscopy. One caveat, however, is that a high fraction of the metabarcoding results cannot be identified due to insufficient information in databases, which hints at an even greater diversity of plankton communities that has yet to be revealed. Furthermore, knowledge about potential occurrences of specific taxa could potentially improve their detection even at routine microscopy analysis (Stern *et al.*, 2018).

Due to higher sample volumes potentially in excess of several litres of water, it is possible to confirm rare species or record new ones, which might be especially useful in cases of potentially harmful and toxic taxa. One litre of water enabled the identification of various parasitoids and small-sized eukaryotic microbes, providing new information on food web connections at our sampling station (chapters 4, 5, and 6). New patterns in the community were discovered that had been overlooked in the LTER, and occurrences at the genus level were found to partly correlate with LTER results (chapter 7). All this information is important for understanding the marine planktonic food web and its biodiversity as well as species interactions in order to comprehend how those interactions affect ecosystem functionality and services especially in times of global change.

Depending on one’s assessment of the importance of this new information, a variety of opportunities exist to implement the new methods. At Helgoland Roads, for example, where

water samples are taken five times a week, the additional volume needed to conduct metabarcoding analysis can be sampled easily. Filtration of the sample does not take long; but one alternative would be an automated sampling system such as the Automated Filtration System for Marine Microbes (Metfies *et al.*, 2020). Another would be to use the same sample as that used for microscopy (chapter 7 Table 1). However, so far there has only been a preliminary trial of metabarcoding after formalin fixation (Shiozaki *et al.*, 2021), so this alternative should be evaluated further. Instead of the morphological expertise needed for conventional methods, expertise in molecular biology would be required for further processing of metabarcoding samples. On the contrary, several samples could be processed at once. One general feature of this would be a slightly delayed analysis, which, if intended for purposes that require timely detection such as to detect harmful blooms, would be a significant drawback; however, microscopic analysis is also time-consuming, and so the time gap between the two methods might be small.

Especially since the value of time series increases if the available data span at least a few decades (Rebstock, 2002; Walther *et al.*, 2002), it is crucial not to delay implementing promising new methods such as metabarcoding. Even when full processing of samples may not be possible due to time and cost restraints in an existing time series, creating environmental biobanks or collecting filters of water samples can be helpful for future analysis in long-term monitoring (Jarman *et al.*, 2018). Additionally, archiving raw data in sequence data repositories (as is already being done) carries the potential to support reanalysis of data later on once drawbacks such as insufficient databases are resolved.

Generally, it would be naïve to think that metabarcoding can replace conventional methods, because conventional methods provide a vast amount of information that metabarcoding cannot reliably cover. Information on life stages, ongoing parasite infections, and individual species still requires microscopy techniques. Where light microscopy yields too little, other microscopy methods such as TEM and SEM can provide detailed cell structure analysis and further identification at species level. These methods are still common in describing new species (Chrétiennot-Dinet *et al.*, 1995; Heimdal, 1997; Johnson and Martiny, 2015).

While information on community diversity and dynamics might be available at higher taxonomic levels, these levels in the metabarcoding dataset cover too many species that conventional light microscopy does not. Comparisons should thus be conducted instead at the genus or species level if the species level could be recognized through the respective metabarcoding approach. Some barcoding regions are not suitable for species recognition and should be reserved for community analysis. In other cases, species-level information might be gained through greater efforts at barcoding species identified taxonomically

through conventional methods; nor should the complementary information on other taxa obtained through metabarcoding be ignored. As for sampling frequency, a simultaneous sampling would result in the best comparability of datasets. As for the additional workload in relation to the added value of newly obtainable information, a reduced sampling might still be enough to capture the various communities of small-sized eukaryotic microbes over the years. For example, in spring 2016 mostly five samples were gathered per week (chapter 4), whereas sampling frequency fell to two to three samples per week for the rest of the sampling period (chapters 5 to 7). However, using genera that both methods reliably capture the comparability of the datasets can be verified regularly. Implementation of multi-marker approaches (which might also include information on prokaryotes) or the development of correction factors might increase both the level of information on food web components and interactions as well as the comparability of datasets gathered through different methods. In general, one must also ensure that different metabarcoding datasets can be compared, especially if the workflow changes. As bioinformatic pipelines and databases are evolving and get updated constantly, metadata from the metabarcoding analysis could also be used for re-analysis, which could lower the high fraction of unidentified diversity even further, improving our ability to explore interactions between various compartments of the marine food web.

9 Conclusion and outlook

Overall, this thesis was able to show how metabarcoding can access the wide range of heretofore hidden information in the planktonic food web. A sample size of one litre of water contained evidence of eukaryotic taxa throughout all trophic levels. The applied method displayed a complex and highly variable marine food web with a great variety of potential species interactions and a high degree of parasitism. If sampling frequencies are high enough, it is also possible to identify communities made up of consistently present taxa. Metabarcoding as part of an ecological time series may help identify global change in plankton communities on a comprehensive scale. It is already possible to identify temporal patterns in these communities at higher taxonomic levels, which can simplify the analysis; and identifications of this nature can be conducted even if databases are still lacking in information about lower taxonomic levels and a sizeable share of the communities still remains unidentified.

The method is useful especially for identifying small eukaryotic microbes, and improvements to databases will enhance knowledge about the diversity of these small organisms even more. The extent to which these taxa form and influence the planktonic food web still needs to be investigated; in particular, the high diversity and frequent occurrence of parasitoids in the food web needs further study as their effect on the plankton community can be considerable. Strong but variable connections in the food web as represented in metabarcoding datasets can be visualized through network analysis, for example, which can help to untangle complex interactions in the food web. Future investigations ought to include surveys of the various associations and co-occurrences that turn up in the food web over the years. Such surveys might guide experimental work on predator-prey interactions. What is more, multi-marker approaches or connections to prokaryotic datasets where possible should result in a wealth of discoverable information, further expanding our knowledge of various planktonic interactions whose existence and importance would be neglected if the different compartments were examined independently of each other.

Since identifying important players in the marine food web remains a challenging task, metabarcoding methods would be a welcome addition to long-term observations. However, one main issue remains the difficulty of merging different datasets, and there is no guaranteed solution. Therefore, the different methodologies might focus on different aspects of future research. While microscopy is essential for the identification and quantification of distinct plankton species, metabarcoding might be used to enhance knowledge about whole communities; but only by effectively combining the results of the

conventional methods with those of metabarcoding will it be possible to profit from the added value of emerging methods; information provided by emerging methods might contrast with traditional views in plankton research and broaden its focus. For example, metabarcoding identified a surprisingly high abundance of non-autotrophic taxa during spring. Besides conventional light microscopy, other methods such as flow-through imaging or general flow cytometry could be coupled with an automated sampling station that took samples for use in barcoding approaches and may even enhance database quality. If feasible, and if time and cost constraints do not limit the time series, the potential synergistic effects of the methods and the datasets they require are impossible to overlook; but the inaccuracies of all methods need to be addressed to ensure the quality of any such time series.

Besides the difficulties in comparing results between different methods used in a single long-term time series, there is also a problem of comparability with other time series. It might be helpful to increase the use of standardized methods. By implementing new methods into existing time series or by starting time series from scratch, methods such as metabarcoding could be used in an internationally standardized manner. This would be a first step toward easier comparison of datasets from several time series stations and could provide a more global view of changes in plankton communities and dynamics in a changing world.

If metabarcoding is implemented in a time series, it may be advisable to use several primer sets of different regions for the same samples to broaden the spectrum of identification to include both prokaryotic and eukaryotic taxa and to minimize the bias of a single primer set. However, in order to minimize personnel costs and workload and to simplify the analysis, the easiest way would be to automize some or even most processes. While separate automatic sampling would impede the comparability of a conventional microscopy dataset with the metabarcoding dataset, it could improve comparability between different metabarcoding datasets due to a standardized sampling without human inputs.

Metabarcoding can provide information on the diversity of small-sized microbial eukaryotes including planktonic parasitoids as well as on regularly observed plankton. New insights into the complex planktonic food web at Helgoland were also obtained; this food web turns out to be much more diverse and flexible than known previously. This previously missing information raises new questions. It can also be implemented into ecological models to get a more life-like picture and increase the reliability and accuracy of these models. The new insights into the plankton community that this thesis presents emphasize the need for ongoing, in-depth analysis of plankton communities and their impact on marine life using a combination of molecular and conventional methods.

10 References

Abad, D., Albaina, A., Aguirre, M., Laza-Martínez, A., Uriarte, I., Iriarte, A., Villate, F., *et al.* 2016. Is metabarcoding suitable for estuarine plankton monitoring? A comparative study with microscopy. Marine Biology, 163: 1–13.

Aberle, N., Lengfellner, K., and Sommer, U. 2007. Spring bloom succession, grazing impact and herbivore selectivity of ciliate communities in response to winter warming. Oecologia, 150: 668–681.

Acinas, S. G., Sarma-Rupavtarm, R., Klepac-Ceraj, V., and Polz, M. F. 2005. PCR-Induced Sequence Artifacts and Bias: Insights from Comparison of Two 16S rRNA Clone Libraries Constructed from the Same Sample. Applied and Environmental Microbiology, 71: 8966–8969.

Adl, S. M., Bass, D., Lane, C. E., Lukeš, J., Schoch, C. L., Smirnov, A., Agatha, S., *et al.* 2019. Revisions to the Classification, Nomenclature, and Diversity of Eukaryotes. Journal of Eukaryotic Microbiology, 66: 4–119.

Alacid, E., Park, M. G., Turon, M., Petrou, K., and Garcés, E. 2016. A Game of Russian Roulette for a Generalist Dinoflagellate Parasitoid: Host Susceptibility Is the Key to Success. Frontiers in Microbiology, 7: 1–13.

Alvarez-Fernandez, S., and Riegman, R. 2014. Chlorophyll in North Sea coastal and offshore waters does not reflect long term trends of phytoplankton biomass. Journal of Sea Research, 91: 35–44.

Andersen, P., and Throndsen, J. 1995. Estimating cell numbers. *In* Manual on harmful marine microalgae. IOC Manuals and Guides, pp. 63–80. Ed. by G. M. Hallegraeff, D. M. Anderson, and A. D. Cembella. UNESCO Publishing.

Anderson, S. R., and Harvey, E. L. 2020. Temporal Variability and Ecological Interactions of Parasitic Marine Syndiniales in Coastal Protist Communities. mSphere, 5: 1–16.

Andersson, A., Haecky, P., and Hagström, Å. 1994. Effect of temperature and light on the growth of micro- nano- and pico-plankton: impact on algal succession. Marine Biology, 120: 511–520.

Arranz, V., Pearman, W. S., Aguirre, J. D., and Liggins, L. 2020. MARES, a replicable pipeline and curated reference database for marine eukaryote metabarcoding. Scientific Data, 7: 1–8.

Arsenieff, L., Simon, N., Rigaut-Jalabert, F., Le Gall, F., Chaffron, S., Corre, E., Com, E., *et al.* 2019. First Viruses Infecting the Marine Diatom *Guinardia delicatula*. Frontiers in Microbiology, 9.

Azam, F., Fenchel, T., Field, J. G., Gray, J. S., Meyer-Reil, L. A., and Thingstad, F. 1983. The Ecological Role of Water-Column Microbes in the Sea. Marine Ecology Progress Series, 10: 257–263.

Bachy, C., Dolan, J. R., López-García, P., Deschamps, P., and Moreira, D. 2013. 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. The ISME Journal, 7: 244–255.

Bailet, B., Apothéloz-Perret-Gentil, L., Baričević, A., Chonova, T., Franc, A., Frigerio, J.-M., Kelly, M., *et al.* 2020. Diatom DNA metabarcoding for ecological assessment: Comparison among bioinformatics pipelines used in six European countries reveals the need for standardization. Science of The Total Environment, 745: 140948.

Banos, S., Gysi, D. M., Richter-Heitmann, T., Glöckner, F. O., Boersma, M., Wiltshire, K. H., Gerdts, G., *et al.* 2020. Seasonal Dynamics of Pelagic Mycoplanktonic Communities: Interplay of Taxon Abundance, Temporal Occurrence, and Biotic Interactions. Frontiers in Microbiology, 11: 1–17.

Becker, S., Böger, P., Oehlmann, R., and Ernst, A. 2000. PCR Bias in Ecological Analysis: a Case Study for Quantitative Taq Nuclease Assays in Analyses of Microbial Communities. Applied and Environmental Microbiology, 66: 4945–4953.

Beng, K. C., Wolinska, J., Funke, E., Van den Wyngaert, S., Gsell, A. S., and Monaghan, M. T. 2021. Temporal dynamics of freshwater planktonic parasites inferred using a DNA metabarcoding time-series. Parasitology, 148: 1602–1611.

Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., and Sayers, E. W. 2013. GenBank. Nucleic Acids Research, 41: 36–42.

Berry, T. E., Saunders, B. J., Coghlan, M. L., Stat, M., Jarman, S., Richardson, A. J., Davies, C. H., *et al.* 2019. Marine environmental DNA biomonitoring reveals seasonal patterns in biodiversity and identifies ecosystem responses to anomalous climatic events. PLoS Genetics, 15: 1–19.

Boersma, M., Wiltshire, K. H., Kong, S. M., Greve, W., and Renz, J. 2015. Long-term change in the copepod community in the southern German Bight. Journal of Sea Research, 101: 41–50.

Bograd, S. J., Checkley, D. A., and Wooster, W. S. 2003. CalCOFI: a half century of physical, chemical, and biological research in the California Current System. Deep Sea Research Part II: Topical Studies in Oceanography, 50: 2349–2353.

Bolaños, L. M., Choi, C. J., Worden, A. Z., Baetge, N., Carlson, C. A., and Giovannoni, S. 2021. Seasonality of the Microbial Community Composition in the North Atlantic. Frontiers in Marine Science, 8: 1–16.

Bolger, A. M., Lohse, M., and Usadel, B. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics, 30: 2114–2120.

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., Alexander, H., *et al.* 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nature Biotechnology, 37: 852–857.

Brandt, M. I., Trouche, B., Quintric, L., Günther, B., Wincker, P., Poulain, J., and Arnaud-Haond, S. 2021. Bioinformatic pipelines combining denoising and clustering tools allow for more comprehensive prokaryotic and eukaryotic metabarcoding. Molecular Ecology Resources, 21: 1904–1921.

Brockmann, U. H., Laane, R. W. P. M., and Postma, J. 1990. Cycling of nutrient elements in the North Sea. Netherlands Journal of Sea Research, 26: 239–264.

Buaya, A. T., Ploch, S., Hanic, L., Nam, B., Nigrelli, L., Kraberg, A., and Thines, M. 2017. Phylogeny of *Miracula helgolandica* gen. et sp. nov. and *Olpidiopsis drebesii* sp. nov., two basal oomycete parasitoids of marine diatoms, with notes on the taxonomy of *Ectrogella*-like species. Mycological Progress, 16: 1041–1050.

Bucklin, A., Lindeque, P. K., Rodriguez-Ezpeleta, N., Albaina, A., and Lehtiniemi, M. 2016. Metabarcoding of marine zooplankton: Prospects, progress and pitfalls. Journal of Plankton Research, 38: 393–400.

Bucklin, A., Yeh, H. D., Questel, J. M., Richardson, D. E., Reese, B., Copley, N. J., and Wiebe, P. H. 2019. Time-series metabarcoding analysis of zooplankton diversity of the NW Atlantic continental shelf. ICES Journal of Marine Science, 76: 1162–1176.

Burkill, P. H., Edwards, E. S., and Sleight, M. A. 1995. Microzooplankton and their role in controlling phytoplankton growth in the marginal ice zone of the Bellingshausen Sea. Deep Sea Research Part II: Topical Studies in Oceanography, 42: 1277–1290.

Calbet, A., and Landry, M. R. 2004. Phytoplankton growth, microzooplankton grazing, and carbon cycling in marine systems. Limnology and Oceanography, 49: 51–57.

Calbet, A., and Saiz, E. 2005. The ciliate-copepod link in marine ecosystems. Aquatic Microbial Ecology, 38: 157–167.

Calbet, A., Trepat, I., Almeda, R., Saló, V., Saiz, E., Movilla, J. I., Alcaraz, M., *et al.* 2008. Impact of micro- and nanograzers on phytoplankton assessed by standard and size-fractionated dilution grazing experiments. Aquatic Microbial Ecology, 50: 145–156.

Calbet, A. 2008. The trophic roles of microzooplankton in marine systems. ICES Journal of Marine Science, 65: 325–331.

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and Holmes, S. P. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nature Methods, 13: 581–583.

Callahan, B. J., McMurdie, P. J., and Holmes, S. P. 2017. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. ISME Journal, 11: 2639–2643.

Callies, U., and Scharfe, M. 2015. Mean spring conditions at Helgoland Roads, North Sea: Graphical modeling of the influence of hydro-climatic forcing and Elbe River discharge. Journal of Sea Research, 101: 1–11.

Callies, U., Gaslikova, L., Kapitza, H., and Scharfe, M. 2017. German Bight residual current variability on a daily basis: principal components of multi-decadal barotropic simulations. Geo-Marine Letters, 37: 151–162.

Canesi, K. L., and Rynearson, T. A. 2016. Temporal variation of *Skeletonema* community composition from a long-term time series in Narragansett Bay identified using high-throughput DNA sequencing. Marine Ecology Progress Series, 556: 1–16.

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

Caron, D. A. 2016. Mixotrophy stirs up our understanding of marine food webs. Proceedings of the National Academy of Sciences, 113: 2806–2808.

Caron, D. A., and Hu, S. K. 2019. Are We Overestimating Protistan Diversity in Nature? Trends in Microbiology, 27: 197–205.

Cermeño, P., Teixeira, I. G., Branco, M., Figueiras, F. G., and Marañón, E. 2014. Sampling the limits of species richness in marine phytoplankton communities. Journal of Plankton Research, 36: 1135–1139.

Chafee, M., Fernàndez-Guerra, A., Buttigieg, P. L., Gerdts, G., Eren, A. M., Teeling, H., and Amann, R. I. 2018. Recurrent patterns of microdiversity in a temperate coastal marine environment. The ISME Journal, 12: 237–252.

Chambouvet, A., Morin, P., Marie, D., and Guillou, L. 2008. Control of Toxic Marine Dinoflagellate Blooms by Serial Parasitic Killers. Science, 322: 1254–1257.

Chen, T., Zhang, Y., Song, S., Liu, Y., Sun, X., and Li, C. 2021. Diversity and seasonal variation of marine phytoplankton in Jiaozhou Bay, China revealed by morphological observation and metabarcoding. Journal of Oceanology and Limnology.

Chivers, W. J., Walne, A. W., and Hays, G. C. 2017. Mismatch between marine plankton range movements and the velocity of climate change. Nature Communications, 8: 14434.

Choi, J. W., and Stoecker, D. K. 1989. Effects of Fixation on Cell Volume of Marine Planktonic Protozoa. Applied and Environmental Microbiology, 55: 1761–1765.

Chrétiennot-Dinet, M.-J., Courties, C., Vaquer, A., Neveux, J., Claustre, H., Lautier, J., Machado, M. C., *et al.* 1995. A new marine picoeucaryote: *Ostreococcus tauri* gen. et sp. nov. (Chlorophyta, Prasinophyceae). Phycologia, 34: 285–292.

Clarke, L. J., Beard, J. M., Swadling, K. M., and Deagle, B. E. 2017. Effect of marker choice and thermal cycling protocol on zooplankton DNA metabarcoding studies. Ecology and Evolution, 7: 873–883.

Clayton, S., Alexander, H., Graff, J. R., Poulton, N. J., Thompson, L. R., Benway, H., Boss, E., *et al.* 2022. Bio-GO-SHIP: The Time Is Right to Establish Global Repeat Sections of Ocean Biology. Frontiers in Marine Science, 8: 767443.

Cleary, A. C., and Durbin, E. G. 2016. Unexpected prevalence of parasite 18S rDNA sequences in winter among Antarctic marine protists. Journal of Plankton Research, 38: 401–417.

Coats, D. W. 1999. Parasitic Life Styles of Marine Dinoflagellates. The Journal of Eukaryotic Microbiology, 46: 402–409.

Compson, Z. G., McClenaghan, B., Singer, G. A. C., Fahner, N. A., and Hajibabaei, M. 2020. Metabarcoding From Microbes to Mammals: Comprehensive Bioassessment on a Global Scale. Frontiers in Ecology and Evolution, 8: 1–19.

Connolly, J. A., Oliver, M. J., Beaulieu, J. M., Knight, C. A., Tomanek, L., and Moline, M. A. 2008. Correlated evolution of genome size and cell volume in diatoms

(Bacillariophyceae). Journal of Phycology, 44: 124–131.

D'Alelio, D., Libralato, S., Wyatt, T., and Ribera d'Alcalà, M. 2016. Ecological-network models link diversity, structure and function in the plankton food-web. Scientific Reports, 6: 21806.

Dayel, M. J., and King, N. 2014. Prey capture and phagocytosis in the choanoflagellate *Salpingoeca rosetta*. PLoS ONE, 9: 1–6.

de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., Lara, E., *et al.* 2015. Eukaryotic plankton diversity in the sunlit ocean. Science, 348: 1261605–1/11.

del Campo, J., Kolisko, M., Boscaro, V., Santoferrara, L. F., Nenarokov, S., Massana, R., Guillou, L., *et al.* 2018. EukRef: Phylogenetic curation of ribosomal RNA to enhance understanding of eukaryotic diversity and distribution. PLoS Biology, 16: 1–14.

Djurhuus, A., Pitz, K., Sawaya, N. A., Rojas-Márquez, J., Michaud, B., Montes, E., Muller-Karger, F., *et al.* 2018. Evaluation of marine zooplankton community structure through environmental DNA metabarcoding. Limnology and Oceanography: Methods, 16: 209–221.

Doherty, M., Costas, B., McManus, G., and Katz, L. 2007. Culture-independent assessment of planktonic ciliate diversity in coastal northwest Atlantic waters. Aquatic Microbial Ecology, 48: 141–154.

Drebes, G. 1966. Ein parasitischer Phycomycet (Lagenidiales) in *Coscinodiscus*. Helgoländer Wissenschaftliche Meeresuntersuchungen, 13: 426–435.

Drebes, G., Kühn, S. F., Gmelch, A., and Schnepf, E. 1996. *Cryothecomonas aestivalis* sp. nov., a colourless nanoflagellate feeding on the marine centric diatom *Guinardia delicatula* (Cleve) Hasle. Helgoländer Meeresuntersuchungen, 50: 497–515.

Duarte, C. M., Cebrián, J., and Marbà, N. 1992. Uncertainty of detecting sea change. Nature, 356: 190–190.

Ducklow, H. W., Steinberg, D. K., and Buesseler, K. O. 2001. Upper ocean carbon export and the biological pump. Oceanography, 14: 50–58.

Ducklow, H. W., Doney, S. C., and Steinberg, D. K. 2009. Contributions of long-term research and time-series observations to marine ecology and biogeochemistry. Annual Review of Marine Science, 1: 279–302.

Ducrotoy, J.-P., Elliott, M., and de Jonge, V. N. 2000. The North Sea. Marine Pollution Bulletin, 41: 5–23.

Durand, M. D., Olson, R. J., and Chisholm, S. W. 2001. Phytoplankton population dynamics at the Bermuda Atlantic Time-series station in the Sargasso Sea. Deep-Sea Research Part II: Topical Studies in Oceanography, 48: 1983–2003.

Ebenezer, V., Medlin, L. K., and Ki, J. S. 2012. Molecular Detection, Quantification, and Diversity Evaluation of Microalgae. Marine Biotechnology, 14: 129–142.

Edgar, R. C. 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics, 26: 2460–2461.

Edler, L., and Elbrächter, M. 2010. The Utermöhl method for quantitative phytoplankton analysis. *In* Microscopic and molecular methods for quantitative phytoplankton analysis (IOC Manuals and Guides, no. 55.), pp. 13–20. Ed. by B. Karlson, C. Cusack, and E. Bresnan. UNESCO Publishing.

Edwards, E. S., and Burkill, P. H. 1995. Abundance, biomass and distribution of microzooplankton in the Irish Sea. Journal of Plankton Research, 17: 771–782.

Edwards, K. F. 2019. Mixotrophy in nanoflagellates across environmental gradients in the ocean. Proceedings of the National Academy of Sciences, 116: 6211–6220.

Emeis, K. C., van Beusekom, J., Callies, U., Ebinghaus, R., Kannen, A., Kraus, G., Kröncke, I., *et al.* 2015. The North Sea - A shelf sea in the Anthropocene. Journal of Marine Systems, 141: 18–33.

Estaki, M., Jiang, L., Bokulich, N. A., McDonald, D., González, A., Kosciolek, T., Martino, C., *et al.* 2020. QIIME 2 Enables Comprehensive End-to-End Analysis of Diverse Microbiome Data and Comparative Studies with Publicly Available Data. Current Protocols in Bioinformatics, 70: 1–46.

Field, C. B., Behrenfeld, M. J., Randerson, J. T., and Falkowski, P. 1998. Primary Production of the Biosphere: Integrating Terrestrial and Oceanic Components. Science, 281: 237–240.

Figueiras, F. G., Arbones, B., Castro, C. G., Froján, M., and Teixeira, I. G. 2020. About Pigmented Nanoflagellates and the Importance of Mixotrophy in a Coastal Upwelling System. Frontiers in Marine Science, 7.

Flynn, K. J., Mitra, A., Anestis, K., Anschütz, A. A., Calbet, A., Ferreira, G. D., Gypens, N., *et al.* 2019. Mixotrophic protists and a new paradigm for marine ecology: where does plankton research go now? Journal of Plankton Research, 41: 375–391.

Fogg, G. E. 1995. Some comments on picoplankton and its importance in the pelagic

ecosystem. Aquatic Microbial Ecology, 9: 33–39.

Fonda Umani, S., Tirelli, V., Beran, A., and Guardiani, B. 2005. Relationships between microzooplankton and mesozooplankton: competition versus predation on natural assemblages of the Gulf of Trieste (northern Adriatic Sea). Journal of Plankton Research, 27: 973–986.

Freund, J. A., Grüner, N., Brüse, S., and Wiltshire, K. H. 2012. Changes in the phytoplankton community at Helgoland, North Sea: Lessons from single spot time series analyses. Marine Biology, 159: 2561–2571.

Fu, L., Niu, B., Zhu, Z., Wu, S., and Li, W. 2012. CD-HIT: Accelerated for clustering the next-generation sequencing data. Bioinformatics, 28: 3150–3152.

Gaonkar, C. C., Piredda, R., Minucci, C., Mann, D. G., Montresor, M., Sarno, D., and Kooistra, W. H. C. F. 2018. Annotated 18S and 28S rDNA reference sequences of taxa in the planktonic diatom family Chaetocerotaceae. PLoS ONE, 13: 1–20.

Gebühr, C., Wiltshire, K. H., Aberle, N., van Beusekom, J. E. E., and Gerdts, G. 2009. Influence of nutrients, temperature, light and salinity on the occurrence of *Paralia sulcata* at Helgoland Roads, North Sea. Aquatic Biology, 7: 185–197.

Gentsch, E., Kreibich, T., Hagen, W., and Niehoff, B. 2009. Dietary shifts in the copepod *Temora longicornis* during spring: Evidence from stable isotope signatures, fatty acid biomarkers and feeding experiments. Journal of Plankton Research, 31: 45–60.

Gescher, C., Metfies, K., Frickenhaus, S., Knefelkamp, B., Wiltshire, K. H., and Medlin, L. K. 2008. Feasibility of assessing the community composition of prasinophytes at the Helgoland roads sampling site with a DNA microarray. Applied and Environmental Microbiology, 74: 5305–5316.

Gifford, D. J. 1991. The Protozoan-Metazoan Trophic Link In Pelagic Ecosystems. The Journal of Protozoology, 38: 81–86.

Giner, C. R., Forn, I., Romac, S., Logares, R., de Vargas, C., and Massana, R. 2016. Environmental sequencing provides reasonable estimates of the relative abundance of specific picoeukaryotes. Applied and Environmental Microbiology, 82: 4757–4766.

Gómez, F., Moreira, D., Benzerara, K., and López-García, P. 2011. *Solenicola setigera* is the first characterized member of the abundant and cosmopolitan uncultured marine stramenopile group MAST-3. Environmental Microbiology, 13: 193–202.

Gong, W., and Marchetti, A. 2019. Estimation of 18S gene copy number in marine

eukaryotic plankton using a next-generation sequencing approach. Frontiers in Marine Science, 6: 1–5.

Gower, J., King, S., and Goncalves, P. 2008. Global monitoring of plankton blooms using MERIS MCI. International Journal of Remote Sensing, 29: 6209–6216.

Gran-Stadniczeñko, S., Šupraha, L., Egge, E. D., and Edvardsen, B. 2017. Haptophyte Diversity and Vertical Distribution Explored by 18S and 28S Ribosomal RNA Gene Metabarcoding and Scanning Electron Microscopy. Journal of Eukaryotic Microbiology, 64: 514–532.

Gran-Stadniczeñko, S., Egge, E., Hostyeva, V., Logares, R., Eikrem, W., and Edvardsen, B. 2019. Protist Diversity and Seasonal Dynamics in Skagerrak Plankton Communities as Revealed by Metabarcoding and Microscopy. Journal of Eukaryotic Microbiology, 66: 494–513.

Grasshoff, K. 1976. Methods of Seawater Analysis. Verlag Chemie, Weinheim and New York. 317 pp.

Greve, W., Reiners, F., Nast, J., and Hoffmann, S. 2004. Helgoland Roads meso- and macrozooplankton time-series 1974 to 2004: Lessons from 30 years of single spot, high frequency sampling at the only off-shore island of the North Sea. Helgoland Marine Research, 58: 274–288.

Grob, C., Ulloa, O., Claustre, H., Huot, Y., Alarcón, G., and Marie, D. 2007. Contribution of picoplankton to the total particulate organic carbon concentration in the eastern South Pacific. Biogeosciences, 4: 837–852.

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

Guillou, L., Alves-de-Souza, C., Siano Dr, R., and González, H. 2010. The ecological significance of small, eukaryotic parasites in marine ecosystems. Microbiology Today, 37: 92–95.

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., Boutte, C., *et al.* 2013. The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. Nucleic Acids Research, 41: D597–D604.

Hällfors, G., Melvasalo, T., Niemi, Å., and Viljamaa, H. 1979. Effect of different fixatives and

preservatives on phytoplankton counts. Publ. Wat. Res. Inst., 34: 25–34.

Halpern, B. S., Walbridge, S., Selkoe, K. A., Kappel, C. V., Micheli, F., D'Agrosa, C., Bruno, J. F., *et al.* 2008. A Global Map of Human Impact on Marine Ecosystems. Science, 319: 948–952.

Hamm, C. E. 2000. Architecture, ecology and biogeochemistry of *Phaeocystis* colonies. Journal of Sea Research, 43: 307–315.

Hansen, B., Bjornsen, P. K., and Hansen, P. J. 1994. The size ratio between planktonic predators and their prey. Limnology and Oceanography, 39: 395–403.

Hansen, P. J., and Calado, A. J. 1999. Phagotrophic Mechanisms and Prey Selection in Free-living Dinoflagellates. The Journal of Eukaryotic Microbiology, 46: 382–389.

Harris, R. 2010. The L4 time-series: The first 20 years. Journal of Plankton Research, 32: 577–583.

Harvey, J. B. J., Johnson, S. B., Fisher, J. L., Peterson, W. T., and Vrijenhoek, R. C. 2017. Comparison of morphological and next generation DNA sequencing methods for assessing zooplankton assemblages. Journal of Experimental Marine Biology and Ecology, 487: 113–126.

Hays, G. C., Richardson, A. J., and Robinson, C. 2005. Climate change and marine plankton. Trends in Ecology and Evolution, 20: 337–344.

Heimdal, B. R. 1997. Chapter 6: Modern Coccolithophorids. *In* Identifying Marine Phytoplankton, pp. 731–831. Ed. by C. R. Tomas. Academic Press, Inc.

Hensen, V. 1887. Ueber die Bestimmung des Planktons oder des im Meere treibenden Materials an Pflanzen und Thieren. Bericht der Komission zur wissenschaftlichen Untersuchung der deutschen Meere, 5: 1–109.

Hensen, V. 1890. Einige Ergebnisse der Plankton-Expedition der Humboldt-Stiftung. Naturwissenschaftliche Rundschau 5: 318–320.

Hensen, V. 1895. Methodik der Untersuchungen. Lipsius & Tischer.

Henson, S. A. 2014. Slow science: the value of long ocean biogeochemistry records. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 372: 20130334.

Hickel, W. 1998. Temporal variability of micro- and nanoplankton in the German Bight in relation to hydrographic structure and nutrient changes. ICES Journal of Marine

Science, 55: 600–609.

Hirai, J., Nagai, S., and Hidaka, K. 2017. Evaluation of metagenetic community analysis of planktonic copepods using Illumina MiSeq: Comparisons with morphological classification and metagenetic analysis using Roche 454. PLoS ONE, 12: e0181452.

Hong, H.-H., Lee, H.-G., Jo, J., Kim, H. M., Kim, S.-M., Park, J. Y., Jeon, C. B., *et al.* 2016. The exceptionally large genome of the harmful red tide dinoflagellate *Cochlodinium polykrikoides* Margalef (Dinophyceae): determination by flow cytometry. ALGAE, 31: 373–378.

Hoppenrath, M. 2004. A revised checklist of planktonic diatoms and dinoflagellates from Helgoland (North Sea, German Bight). Helgoland Marine Research, 58: 243–251.

Hughes, A. R., Hanley, T. C., Moore, A. F. P., Ramsay-Newton, C., Zerebecki, R. A., and Sotka, E. E. 2019. Predicting the sensitivity of marine populations to rising temperatures. Frontiers in Ecology and the Environment, 17: 17–24.

Hughes, B. B., Beas-Luna, R., Barner, A. K., Brewitt, K., Brumbaugh, D. R., Cerny-Chipman, E. B., Close, S. L., *et al.* 2017. Long-Term studies contribute disproportionately to ecology and policy. BioScience, 67: 271–278.

Hutchinson, G. E. 1961. The Paradox of the Plankton. The American Naturalist, 95: 137–145.

Huxham, M., Raffaelli, D., and Pike, A. 1995. Parasites and Food Web Patterns. The Journal of Animal Ecology, 64: 168.

Irigoien, X., Meyer, B., Harris, R., and Harbour, D. 2004. Using HPLC pigment analysis to investigate phytoplankton taxonomy: The importance of knowing your species. Helgoland Marine Research, 58: 77–82.

Jacquet, S., Prieur, L., Avois-Jacquet, C., Lennon, J. F., and Vaulot, D. 2002. Short-timescale variability of picophytoplankton abundance and cellular parameters in surface waters of the Alboran Sea (western Mediterranean). Journal of Plankton Research, 24: 635–651.

Jarman, S. N., Berry, O., and Bunce, M. 2018. The value of environmental DNA biobanking for long-term biomonitoring. Nature Ecology and Evolution, 2: 1192–1193.

Jeong, H. 1994. Predation by the heterotrophic dinoflagellate *Protoperidinium* cf. *divergens* on copepod eggs and early naupliar stages. Marine Ecology Progress Series, 114: 203–208.

John, E. H., Batten, S. D., Harris, R. P., and Hays, G. C. 2001. Comparison between zooplankton data collected by the Continuous Plankton Recorder survey in the English Channel and by WP-2 nets at station L4, Plymouth (UK). Journal of Sea Research, 46: 223–232.

Johnson, Z. I., and Martiny, A. C. 2015. Techniques for Quantifying Phytoplankton Biodiversity. Annual Review of Marine Science, 7: 299–324.

Jones, R. H., and Flynn, K. J. 2005. Nutritional status and diet composition affect the value of diatoms as copepod prey. Science, 307: 1457–1459.

Juranek, L. W., White, A. E., Dugenne, M., Henderikx Freitas, F., Dutkiewicz, S., Ribalet, F., Ferrón, S., *et al.* 2020. The Importance of the Phytoplankton “Middle Class” to Ocean Net Community Production. Global Biogeochemical Cycles, 34.

Kang, W., Anslan, S., Börner, N., Schwarz, A., Schmidt, R., Künzel, S., Rioual, P., *et al.* 2021. Diatom metabarcoding and microscopic analyses from sediment samples at Lake Nam Co, Tibet: The effect of sample-size and bioinformatics on the identified communities. Ecological Indicators, 121.

Karentz, D., and Smayda, T. 1984. Temperature and seasonal occurrence patterns of 30 dominant phytoplankton species in Narragansett Bay over a 22-year period (1959-1980). Marine Ecology Progress Series, 18: 277–293.

Karl, D. M., Bidigare, R. R., and Letelier, R. M. 2001. Long-term changes in plankton community structure and productivity in the North Pacific Subtropical Gyre: The domain shift hypothesis. Deep Sea Research Part II: Topical Studies in Oceanography, 48: 1449–1470.

Karl, D. M., and Church, M. J. 2014. Microbial oceanography and the Hawaii Ocean Time-series programme. Nature Reviews Microbiology, 12: 699–713.

Keck, F., Blackman, R. C., Bossart, R., Brantschen, J., Couton, M., Hürlemann, S., Kirschner, D., *et al.* 2022. Meta-analysis shows both congruence and complementarity of DNA and eDNA metabarcoding to traditional methods for biological community assessment. Molecular Ecology: 1–16.

Kembel, S. W., Wu, M., Eisen, J. A., and Green, J. L. 2012. Incorporating 16S Gene Copy Number Information Improves Estimates of Microbial Diversity and Abundance. PLoS Computational Biology, 8: 16–18.

King, N. 2005. Choanoflagellates. Current biology : CB, 15: 113–114.

Knefelkamp, B. 2009. An Investigation of the Eukaryotic Picoplankton Community in the German North Sea by Different Methods. University Bremen, Bremen.

Kobari, T., Tokumo, Y., Sato, I., Kume, G., and Hirai, J. 2021. Metabarcoding analysis of trophic sources and linkages in the plankton community of the Kuroshio and neighboring waters. Scientific Reports, 11: 1–12.

Kraberg, A., Kieb, U., Peters, S., and Wiltshire, K. H. 2019. An updated phytoplankton check-list for the Helgoland Roads time series station with eleven new records of diatoms and dinoflagellates. Helgoland Marine Research, 73: 9.

Kraberg, A. C., Widdicombe, C. E., Beckett, R., Rick, J., Rooks, P., and Van Wezel, R. 2018. Further records of a new diatom species in the English Channel and North Sea: The importance of image-referenced data. Marine Biodiversity Records, 11: 1–9.

Kuebbing, S. E., Reimer, A. P., Rosenthal, S. A., Feinberg, G., Leiserowitz, A., Lau, J. A., and Bradford, M. A. 2018. Long-term research in ecology and evolution: a survey of challenges and opportunities. Ecological Monographs, 88: 245–258.

Kühn, S. F., Drebes, G., and Schnepf, E. 1996. Five new species of the nanoflagellate *Pirsonia* in the German Bight, North Sea, feeding on planktic diatoms. Helgoländer Meeresuntersuchungen, 50: 205–222.

Lafferty, K. D., Dobson, A. P., and Kuris, A. M. 2006. Parasites dominate food web links. Proceedings of the National Academy of Sciences, 103: 11211–11216.

Lafferty, K. D., Allesina, S., Arim, M., Briggs, C. J., De Leo, G., Dobson, A. P., Dunne, J. A., *et al.* 2008. Parasites in food webs: The ultimate missing links. Ecology Letters, 11: 533–546.

LaJeunesse, T. C., Lambert, G., Andersen, R. A., Coffroth, M. A., and Galbraith, D. W. 2005. *Symbiodinium* (Pyrrhophyta) genome sizes (DNA content) are smallest among dinoflagellates. Journal of Phycology, 41: 880–886.

Landry, M. R., Al-Mutairi, H., Selph, K. E., Christensen, S., and Nunnery, S. 2001. Seasonal patterns of mesozooplankton abundance and biomass at Station ALOHA. Deep Sea Research Part II: Topical Studies in Oceanography, 48: 2037–2061.

Landry, M. R., and Calbet, A. 2004. Microzooplankton production in the oceans. ICES Journal of Marine Science, 61: 501–507.

Legendre, L., and Le Fèvre, J. 1995. Microbial food webs and the export of biogenic carbon in oceans. Aquatic Microbial Ecology, 9: 69–77.

Letelier, R. M., Bidigare, R. R., Hebel, D. V., Ondrusek, M., Winn, C. D., and Karl, D. M. 1993. Temporal variability of phytoplankton community structure based on pigment analysis. Limnology and Oceanography, 38: 1420–1437.

Leterme, S., Seuront, L., and Edwards, M. 2006. Differential contribution of diatoms and dinoflagellates to phytoplankton biomass in the NE Atlantic Ocean and the North Sea. Marine Ecology Progress Series, 312: 57–65.

Lewandowska, A., and Sommer, U. 2010. Climate change and the spring bloom: A mesocosm study on the influence of light and temperature on phytoplankton and mesozooplankton. Marine Ecology Progress Series, 405: 101–111.

Lewandowska, A. M., Striebel, M., Feudel, U., Hillebrand, H., and Sommer, U. 2015. The importance of phytoplankton trait variability in spring bloom formation. ICES Journal of Marine Science, 72: 1908–1915.

Lima-Mendez, G., Faust, K., Henry, N., Decelle, J., Colin, S., Carcillo, F., Chaffron, S., *et al.* 2015. Determinants of community structure in the global plankton interactome. Science, 348: 1262073–1262073.

Lin, Y.-C., Chin, C.-P., Yang, J. W., Chiang, K.-P., Hsieh, C., Gong, G.-C., Shih, C.-Y., *et al.* 2021. How Communities of Marine Stramenopiles Varied with Environmental and Biological Variables in the Subtropical Northwestern Pacific Ocean. Microbial Ecology.

Lindeque, P. K., Parry, H. E., Harmer, R. A., Somerfield, P. J., and Atkinson, A. 2013. Next Generation Sequencing Reveals the Hidden Diversity of Zooplankton Assemblages. PLoS ONE, 8: e81327.

Löder, M. G. J., Kraberg, A. C., Aberle, N., Peters, S., and Wiltshire, K. H. 2012. Dinoflagellates and ciliates at Helgoland Roads, North Sea. Helgoland Marine Research, 66: 11–23.

Löder, M. G. J. J., Meunier, C., Wiltshire, K. H., Boersma, M., and Aberle, N. 2011. The role of ciliates, heterotrophic dinoflagellates and copepods in structuring spring plankton communities at Helgoland Roads, North Sea. Marine Biology, 158: 1551–1580.

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 Journal, 6: 1823–1833.

Logares, R., Audic, S., Bass, D., Bittner, L., Boutte, C., Christen, R., Claverie, J. M., *et al.* 2014. Patterns of rare and abundant marine microbial eukaryotes. Current Biology, 24: 813–821.

Louca, S., Doebeli, M., and Parfrey, L. W. 2018. Correcting for 16S rRNA gene copy numbers in microbiome surveys remains an unsolved problem. Microbiome, 6: 1–12.

Lucas, J., Wichels, A., Teeling, H., Chafee, M., Scharfe, M., and Gerdts, G. 2015. Annual dynamics of North Sea bacterioplankton: Seasonal variability superimposes short-term variation. FEMS Microbiology Ecology, 91.

Lucas, J., Koester, I., Wichels, A., Niggemann, J., Dittmar, T., Callies, U., Wiltshire, K. H., *et al.* 2016. Short-Term Dynamics of North Sea Bacterioplankton-Dissolved Organic Matter Coherence on Molecular Level. Frontiers in Microbiology, 7: 1–14.

Lund, J. W. G. 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.

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., Mayor, J., Bunge, J., Chi, J., Siemensmeyer, T., Stoeck, T., Wahl, B., *et al.* 2015a. Comparing high-throughput platforms for sequencing the V4 region of SSU-rDNA in environmental microbial eukaryotic diversity surveys. Journal of Eukaryotic Microbiology, 62: 338–345.

Mahé, F., Rognes, T., Quince, C., de Vargas, C., and Dunthorn, M. 2015b. Swarm v2: highly-scalable and high-resolution amplicon clustering. PeerJ, 3: e1420.

Marcogliese, D. J. 2001. Pursuing parasites up the food chain: Implications of food web structure and function on parasite communities in aquatic systems. Acta Parasitologica, 46: 82–93.

Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., Berka, J., *et al.* 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature, 437: 376–380.

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

Marquardt, M., Vader, A., Stübner, E. I., Reigstad, M., and Gabrielsen, T. M. 2016. Strong Seasonality of Marine Microbial Eukaryotes in a High-Arctic Fjord (Isfjorden, in West Spitsbergen, Norway). Applied and Environmental Microbiology, 82: 1868–1880.

Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing

reads. EMBnet.journal, 17: 10.

Massana, R., Castresana, J., Balagué, V., Guillou, L., Romari, K., Groisillier, A., Valentin, K., *et al.* 2004. Phylogenetic and Ecological Analysis of Novel Marine Stramenopiles. Applied and Environmental Microbiology, 70: 3528–3534.

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 Journal, 3: 588–595.

Massana, R. 2011. Eukaryotic Picoplankton in Surface Oceans. Annual Review of Microbiology, 65: 91–110.

Massana, R., Del Campo, J., Sieracki, M. E., Audic, S., and Logares, R. 2014. Exploring the uncultured microeukaryote majority in the oceans: Reevaluation of ribogroups within stramenopiles. ISME Journal, 8: 854–866.

Massana, R., Gobet, A., Audic, S., Bass, D., Bittner, L., Boutte, C., Chambouvet, A., *et al.* 2015. Marine protist diversity in European coastal waters and sediments as revealed by high-throughput sequencing. Environmental microbiology, 17: 4035–4049.

McNichol, J., Berube, P. M., Biller, S. J., and Fuhrman, J. A. 2021. Evaluating and Improving Small Subunit rRNA PCR Primer Coverage for Bacteria, Archaea, and Eukaryotes Using Metagenomes from Global Ocean Surveys. mSystems, 6.

McQuatters-Gollop, A., Edwards, M., Helaouët, P., Johns, D. G., Owens, N. J. P., Raitsos, D. E., Schroeder, D., *et al.* 2015. The Continuous Plankton Recorder survey: How can long-term phytoplankton datasets contribute to the assessment of Good Environmental Status? Estuarine, Coastal and Shelf Science, 162: 88–97.

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 microorganisms. Molecular Ecology, 19: 32–40.

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. Microbial Ecology, 52: 53–71.

Medlin, L. K., and Kooistra, W. H. C. F. 2010. Methods to estimate the diversity in the marine photosynthetic protist community with illustrations from case studies: A review. Diversity, 2: 973–1014.

Medlin, L. K., Piwosz, K., and Metfies, K. 2017. Uncovering hidden biodiversity in the cryptophyta : clone library studies at the Helgoland time series site in the southern German bight identifies the cryptophycean clade potentially responsible for the majority of its genetic diversity during the spring. Vie et Milieu - Life and Environment, 67: 27–32.

Menden-Deuer, S., Lessard, E., and Satterberg, J. 2001. Effect of preservation on dinoflagellate and diatom cell volume, and consequences for carbon biomass predictions. Marine Ecology Progress Series, 222: 41–50.

Metfies, K., Gescher, C., Frickenhaus, S., Niestroy, R., Wichels, A., Gerdts, G., Knefelkamp, B., *et al.* 2010. Contribution of the class cryptophyceae to phytoplankton structure in the German bight. Journal of Phycology, 46: 1152–1160.

Metfies, K., Hessel, J., Klenk, R., Petersen, W., Wiltshire, K. H., and Kraberg, A. 2020. Uncovering the intricacies of microbial community dynamics at Helgoland Roads at the end of a spring bloom using automated sampling and 18S meta-barcoding. PLoS ONE, 15: e0233921.

Meunier, C. L., Boersma, M., Wiltshire, K. H., and Malzahn, A. M. 2016. Zooplankton eat what they need: Copepod selective feeding and potential consequences for marine systems. Oikos, 125: 50–58.

Mieruch, S., Freund, J. A., Feudel, U., Boersma, M., Janisch, S., and Wiltshire, K. H. 2010. A new method of describing phytoplankton blooms: Examples from Helgoland Roads. Journal of Marine Systems, 79: 36–43.

Millette, N. C., Grosse, J., Johnson, W. M., Jungbluth, M. J., and Suter, E. A. 2018. Hidden in plain sight: The importance of cryptic interactions in marine plankton. Limnology and Oceanography Letters, 3: 341–356.

Mitra, A., Flynn, K. J., Burkholder, J. M., Berge, T., Calbet, A., Raven, J. A., Granéli, E., *et al.* 2014. The role of mixotrophic protists in the biological carbon pump. Biogeosciences, 11: 995–1005.

Montagnes, D. J. S., and Lessard, E. J. 1999. Population dynamics of the marine planktonic ciliate *Strombidinopsis multiauris*: Its potential to control phytoplankton blooms. Aquatic Microbial Ecology, 20: 167–181.

Moreira, D., and López-García, P. 2014. The rise and fall of picobiliphytes: How assumed autotrophs turned out to be heterotrophs. BioEssays, 36: 468–474.

Nejstgaard, J. C., Naustvoll, L. J., and Sazhin, A. 2001. Correcting for underestimation of

microzooplankton grazing in bottle incubation experiments with mesozooplankton. Marine Ecology Progress Series, 221: 59–75.

Nistal-García, A., García-García, P., García-Girón, J., Borrego-Ramos, M., Blanco, S., and Bécares, E. 2021. DNA metabarcoding and morphological methods show complementary patterns in the metacommunity organization of lentic epiphytic diatoms. Science of The Total Environment, 786: 147410.

Not, F., Latasa, M., Marie, D., Cariou, T., Vaulot, D., and Simon, N. 2004. A Single Species, *Micromonas pusilla* (Prasinophyceae), Dominates the Eukaryotic Picoplankton in the Western English Channel. Applied and Environmental Microbiology, 70: 4064–4072.

Not, F., Latasa, M., Scharek, R., Viprey, M., Karleskind, P., Balagué, V., Ontoria-Oviedo, I., *et al.* 2008. Protistan assemblages across the Indian Ocean, with a specific emphasis on the picoeukaryotes. Deep-Sea Research Part I: Oceanographic Research Papers, 55: 1456–1473.

Nöthig, E.-M., Bracher, A., Engel, A., Metfies, K., Niehoff, B., Peeken, I., Bauerfeind, E., *et al.* 2015. Summertime plankton ecology in Fram Strait—a compilation of long- and short-term observations. Polar Research, 34: 23349.

Novotny, A., Zamora-Terol, S., and Winder, M. 2021. DNA metabarcoding reveals trophic niche diversity of micro and mesozooplankton species. Proceedings of the Royal Society B: Biological Sciences, 288.

Otero-Ferrer, J. L., Cermeño, P., Bode, A., Fernández-Castro, B., Gasol, J. M., Morán, X. A. G., Marañon, E., *et al.* 2018. Factors controlling the community structure of picoplankton in contrasting marine environments. Biogeosciences, 15: 6199–6220.

Park, M. G., Yih, W., and Coats, D. W. 2004. Parasites and Phytoplankton, with Special Emphasis on Dinoflagellate Infections. The Journal of Eukaryotic Microbiology, 51: 145–155.

Parr, T. W., Ferretti, M., Simpson, I. C., Forsius, M., and Kovács-Láng, E. 2002. Towards a long-term integrated monitoring programme in Europe: Network design in theory and practice. Environmental Monitoring and Assessment, 78: 253–290.

Paxinos, R., and Mitchell, J. G. 2000. A rapid Utermöhl method for estimating algal numbers. Journal of Plankton Research, 22: 2255–2262.

Peacock, E., Olson, R., and Sosik, H. 2014. Parasitic infection of the diatom *Guinardia delicatula*, a recurrent and ecologically important phenomenon on the New England Shelf. Marine Ecology Progress Series, 503: 1–10.

Petersen, W. 2014. FerryBox systems : State-of-the-art in Europe and future development. Journal of Marine Systems, 140: 4–12.

Pierella Karlusich, J. J., Pelletier, E., Zinger, L., Lombard, F., Zingone, A., Colin, S., Gasol, J. M., *et al.* 2022. A robust approach to estimate relative phytoplankton cell abundances from metagenomes. Molecular Ecology Resources: 1–25. doi: 10.1111/1755-0998.13592.

Piredda, R., Tomasino, M. P., D'Erchia, A. M., Manzari, C., Pesole, G., Montresor, M., Kooistra, W. H. C. F., *et al.* 2017. Diversity and temporal patterns of planktonic protist assemblages at a Mediterranean Long Term Ecological Research site. FEMS Microbiology Ecology, 93.

Pitsch, G., Bruni, E. P., Forster, D., Qu, Z., Sonntag, B., Stoeck, T., and Posch, T. 2019. Seasonality of planktonic freshwater ciliates: Are analyses based on V9 regions of the 18S rRNA gene correlated with morphospecies counts? Frontiers in Microbiology, 10: 1–15.

Pomeroy, L. R. 1974. The ocean's food web, a changing paradigm. BioScience, 24: 499–504.

Pruesse, E., Quast, C., Knittel, K., Fuchs, B. M., Ludwig, W., Peplies, J., and Glöckner, F. O. 2007. SILVA: A comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Research, 35: 7188–7196.

Pugnetti, A., Acri, F., Bernardi Aubry, F., Camatti, E., Cecere, E., Facca, C., Franzoi, P., *et al.* 2013. The Italian long-term ecosystem research (LTER-Italy) network: Results, opportunities, and challenges for coastal transitional ecosystems. Transitional Waters Bulletin, 7: 43–63.

Purcell-Meyerink, D., Fortune, J., and Butler, E. C. V. 2017. Productivity dominated by picoplankton in a macro-tidal tropical estuary, Darwin Harbour. New Zealand Journal of Botany, 55: 47–63.

Raabe, T., and Wiltshire, K. H. 2009. Quality control and analyses of the long-term nutrient data from Helgoland Roads, North Sea. Journal of Sea Research, 61: 3–16.

Rasconi, S., Jobard, M., and Sime-Ngando, T. 2011. Parasitic fungi of phytoplankton: Ecological roles and implications for microbial food webs. Aquatic Microbial Ecology, 62: 123–137.

Raven, J. A. 1998. The Twelfth Tansley Lecture, Small is Beautiful: The Picophytoplankton.

Functional Ecology, 12: 503–513.

Ray, R., Haas, L., and Sieracki, M. 1989. Autotrophic picoplankton dynamics in a Chesapeake Bay sub-estuary. Marine Ecology Progress Series, 52: 273–285.

Rebstock, G. A. 2002. Climatic regime shifts and decadal-scale variability in calanoid copepod populations off southern California. Global Change Biology, 8: 71–89.

Reid, P. C., Lancelot, C., Gieskes, W. W. C., Hagmeier, E., and Weichart, G. 1990. Phytoplankton of the North Sea and its dynamics: A review. Netherlands Journal of Sea Research, 26: 295–331.

Reid, P. C., Colebrook, J. M., Matthews, J. B. L., and Aiken, J. 2003. The Continuous Plankton Recorder: concepts and history, from Plankton Indicator to undulating recorders. Progress in Oceanography, 58: 117–173.

Reñé, A., Timoneda, N., Sampedro, N., Alacid, E., Gallisai, R., Gordi, J., Fernández-Valero, A. D., *et al.* 2021. Host preferences of coexisting Perkinsea parasitoids during coastal dinoflagellate blooms. Molecular Ecology, 30: 2417–2433.

Rhodes, C. J., Truscott, J. E., and Martin, A. P. 2008. Viral infection as a regulator of oceanic phytoplankton populations. Journal of Marine Systems, 74: 216–226.

Richardson, A. J., John, E. H., Irigoien, X., Harris, R. P., and Hays, G. C. 2004. How well does the Continuous Plankton Recorder (CPR) sample zooplankton? A comparison with the Longhurst Hardy Plankton Recorder (LHPR) in the northeast Atlantic. Deep Sea Research Part I: Oceanographic Research Papers, 51: 1283–1294.

Riegman, R., Kuipers, B. R., Noordeloos, A. A. M., and Witte, H. J. 1993. Size-differential control of phytoplankton and the structure of plankton communities. Netherlands Journal of Sea Research, 31: 255–265.

Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahé, F. 2016. VSEARCH: a versatile open source tool for metagenomics. PeerJ, 4: e2584.

Rothberg, J. M., and Leamon, J. H. 2008. The development and impact of 454 sequencing. Nature Biotechnology, 26: 1117–1124.

Roura, Á., Strugnell, J. M., Guerra, Á., González, Á. F., and Richardson, A. J. 2018. Small copepods could channel missing carbon through metazoan predation. Ecology and Evolution, 8: 10868–10878.

Ruppert, K. M., Kline, R. J., and Rahman, M. S. 2019. Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods,

monitoring, and applications of global eDNA. Global Ecology and Conservation, 17: e00547.

Rynearson, T. A., Flickinger, S. A., and Fontaine, D. N. 2020. Metabarcoding Reveals Temporal Patterns of Community Composition and Realized Thermal Niches of *Thalassiosira* spp. (Bacillariophyceae) from the Narragansett Bay Long-Term Plankton Time Series. Biology, 9.

Saage, A., Vadstein, O., and Sommer, U. 2009. Feeding behaviour of adult *Centropages hamatus* (Copepoda, Calanoida): Functional response and selective feeding experiments. Journal of Sea Research, 62: 16–21.

Sanders, R., Berninger, U., Lim, E., Kemp, P., and Caron, D. 2000. Heterotrophic and mixotrophic nanoplankton predation on picoplankton in the Sargasso Sea and on Georges Bank. Marine Ecology Progress Series, 192: 103–118.

Sanders, R. W. 1991. Mixotrophic Protists In Marine and Freshwater Ecosystems. The Journal of Protozoology, 38: 76–81.

Santi, I., Kasapidis, P., Karakassis, I., and Pitta, P. 2021. A comparison of DNA metabarcoding and microscopy methodologies for the study of aquatic microbial eukaryotes. Diversity, 13: 1–12.

Santoferrara, L. F. 2019. Current practice in plankton metabarcoding: Optimization and error management. Journal of Plankton Research, 41: 571–582.

Sapp, M., Wichels, A., Wiltshire, K. H., and Gerdts, G. 2007. Bacterial community dynamics during the winter-spring transition in the North Sea. FEMS Microbiology Ecology, 59: 622–637.

Sarker, S., and Wiltshire, K. H. 2017. Phytoplankton carrying capacity: Is this a viable concept for coastal seas? Ocean & Coastal Management, 148: 1–8.

Sarker, S., Feudel, U., Meunier, C. L., Lemke, P., Dutta, P. S., and Wiltshire, K. H. 2018. To share or not to share ? Phytoplankton species coexistence puzzle in a competition model incorporating multiple resource-limitation and synthesizing unit concepts. Ecological Modelling, 383: 150–159.

Sarker, S. 2018. What explains phytoplankton dynamics ? An analysis of the Helgoland Roads Time Series data sets. Jacobs University, Bremen.

Sassenhagen, I., Irion, S., Jardillier, L., Moreira, D., and Christaki, U. 2020. Protist Interactions and Community Structure During Early Autumn in the Kerguelen Region

(Southern Ocean). Protist, 171: 125709.

Scharfe, M., and Wiltshire, K. H. 2019. Modeling of intra-annual abundance distributions: Constancy and variation in the phenology of marine phytoplankton species over five decades at Helgoland Roads (North Sea). Ecological Modelling, 404: 46–60.

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., Lesniewski, R. A., *et al.* 2009. Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. Applied and Environmental Microbiology, 75: 7537–7541.

Schlüter, M. H., Kraberg, A., and Wiltshire, K. H. 2012. Long-term changes in the seasonality of selected diatoms related to grazers and environmental conditions. Journal of Sea Research, 67: 91–97.

Schoemann, V., Becquevort, S., Stefels, J., Rousseau, V., and Lancelot, C. 2005. *Phaeocystis* blooms in the global ocean and their controlling mechanisms: A review. Journal of Sea Research, 53: 43–66.

Schweikert, M., and Schnepf, E. 1997. Light and electron microscopical observations on *Pirsonia punctigerae* spec. nov., a nanoflagellate feeding on the marine centric diatom *Thalassiosira punctigera*. European Journal of Protistology, 33: 168–177.

Semmouri, I., De Schamphelaere, K. A. C., Willemse, S., Vandegehuchte, M. B., Janssen, C. R., and Asselman, J. 2021. Metabarcoding reveals hidden species and improves identification of marine zooplankton communities in the North Sea. ICES Journal of Marine Science, 78: 3411–3427.

Sherr, E., and Sherr, B. 1988. Role of microbes in pelagic food webs: A revised concept. Limnology and Oceanography, 33: 1225–1227.

Sherr, E. B., and Sherr, B. F. 2002. Significance of predation by protists in aquatic microbial food webs. Antonie van Leeuwenhoek, 81: 293–308.

Sherr, E. B., and Sherr, B. F. 2007. Heterotrophic dinoflagellates: A significant component of microzooplankton biomass and major grazers of diatoms in the sea. Marine Ecology Progress Series, 352: 187–197.

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.

Shiozaki, T., Itoh, F., Hirose, Y., Onodera, J., Kuwata, A., and Harada, N. 2021. A DNA

metabarcoding approach for recovering plankton communities from archived samples fixed in formalin. PLoS ONE, 16: 1–15.

Sieburth, J. M., Smetacek, V., and Lenz, J. 1978. Pelagic ecosystem structure: Heterotrophic compartments of the plankton and their relationship to plankton size fractions. Limnology and Oceanography, 23: 1256–1263.

Sildever, S., Laas, P., Kolesova, N., Lips, I., Lips, U., and Nagai, S. 2021. Plankton biodiversity and species co-occurrence based on environmental DNA – a multiple marker study. Metabarcoding and Metagenomics, 5: 175–197.

Simon, N., Cras, A.-L., Foulon, E., and Lemée, R. 2009. Diversity and evolution of marine phytoplankton. Comptes Rendus Biologies, 332: 159–170.

Skovgaard, A., and Saiz, E. 2006. Seasonal occurrence and role of protistan parasites in coastal marine zooplankton. Marine Ecology Progress Series, 327: 37–49.

Skovgaard, A. 2014. Dirty tricks in the plankton: Diversity and role of marine parasitic protists. Acta Protozoologica, 53: 51–62.

Šlapeta, J., López-García, P., and Moreira, D. 2006. Global Dispersal and Ancient Cryptic Species in the Smallest Marine Eukaryotes. Molecular Biology and Evolution, 23: 23–29.

Smayda, T. J. 1998. Patterns of variability characterizing marine phytoplankton, with examples from Narragansett Bay. ICES Journal of Marine Science, 55: 562–573.

Soares, M. C. S., Lobão, L. M., Vidal, L. O., Noyma, N. P., Barros, N. O., Cardoso, S. J., and Roland, F. 2011. Light Microscopy in Aquatic Ecology: Methods for Plankton Communities Studies. *In* Methods in Molecular Biology, pp. 215–227. Ed. by H. Chiarini-Garcia and R. C. N. Melo. Springer Science+Business Media, LLC 2011.

Sogin, M. L., Morrison, H. G., Huber, J. A., Welch, D. M., Huse, S. M., Neal, P. R., Arrieta, J. M., *et al.* 2006. Microbial diversity in the deep sea and the underexplored 'rare biosphere'. Proceedings of the National Academy of Sciences, 103: 12115–12120.

Soltwedel, T., Bauerfeind, E., Bergmann, M., Budaeva, N., Hoste, E., Jaeckisch, N., von Juterzenka, K., *et al.* 2005. HAUSGARTEN: Multidisciplinary Investigations at a Deep-Sea, Long-Term Observatory in the Arctic Ocean. Oceanography, 18: 46–61.

Soltwedel, T., Bauerfeind, E., Bergmann, M., Bracher, A., Budaeva, N., Busch, K., Cherkasheva, A., *et al.* 2016. Natural variability or anthropogenically-induced variation? Insights from 15 years of multidisciplinary observations at the arctic marine LTER site

HAUSGARTEN. Ecological Indicators, 65: 89–102.

Sommer, U., Stibor, H., Katechakis, A., Sommer, F., and Hansen, T. 2002. Pelagic food web configurations at different levels of nutrient richness and their implications for the ratio fish production:primary production. Hydrobiologia, 484: 11–20.

Sprong, P. A. A., Fofonova, V., Wiltshire, K. H., Neuhaus, S., Ludwichowski, K. U., Käse, L., Androsov, A., *et al.* 2020. Spatial dynamics of eukaryotic microbial communities in the German Bight. Journal of Sea Research, 163: 101914.

Starke, R., Pylro, V. S., and Morais, D. K. 2021. 16S rRNA Gene Copy Number Normalization Does Not Provide More Reliable Conclusions in Metataxonomic Surveys. Microbial Ecology, 81: 535–539.

Stefanni, S., Stanković, D., Borme, D., de Olazabal, A., Juretić, T., Pallavicini, A., and Tirelli, V. 2018. Multi-marker metabarcoding approach to study mesozooplankton at basin scale. Scientific Reports, 8: 1–13.

Stefels, J., Steinke, M., Turner, S., Malin, G., and Belviso, S. 2007. Environmental constraints on the production and removal of the climatically active gas dimethylsulphide (DMS) and implications for ecosystem modelling. Biogeochemistry, 83: 245–275.

Stelfox-Widdicombe, C. E., Archer, S. D., Burkill, P. H., and Stefels, J. 2004. Microzooplankton grazing in *Phaeocystis* and diatom-dominated waters in the southern North Sea in spring. Journal of Sea Research, 51: 37–51.

Stern, R., Kraberg, A., Bresnan, E., Kooistra, W. H. C. F., Lovejoy, C., Montresor, M., Morán, X. A. G., *et al.* 2018. Molecular analyses of protists in long-term observation programmes—current status and future perspectives. Journal of Plankton Research, 40: 519–536.

Stoecker, D., Gifford, D., and Putt, M. 1994. Preservation of marine planktonic ciliates: losses and cell shrinkage during fixation. Marine Ecology Progress Series, 110: 293–299.

Strom, S. L., Miller, C. B., and Frost, B. W. 2000. What sets lower limits to phytoplankton stocks in high-nitrate, low-chlorophyll regions of the open ocean? Marine Ecology Progress Series, 193: 19–31.

Sukhotin, A., and Berger, V. 2013. Long-term monitoring studies as a powerful tool in marine ecosystem research. Hydrobiologia, 706: 1–9.

Sündermann, J., and Pohlmann, T. 2011. A brief analysis of North Sea physics. Oceanologia, 53: 663–689.

Suter, E. A., Pachiadaki, M., Taylor, G. T., and Edgcomb, V. P. 2022. Eukaryotic Parasites Are Integral to a Productive Microbial Food Web in Oxygen-Depleted Waters. Frontiers in Microbiology, 12: 1–16.

Suzuki, M. T., and Giovannoni, S. J. 1996. Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. Applied and Environmental Microbiology, 62: 625–630.

Taylor, J. D., and Cunliffe, M. 2014. High-throughput sequencing reveals neustonic and planktonic microbial eukaryote diversity in coastal waters. Journal of Phycology, 50: 960–965.

Teeling, H., Fuchs, B. M., Bennke, C. M., Krüger, K., Chafee, M., Kappelmann, L., Reintjes, G., *et al.* 2016. Recurring patterns in bacterioplankton dynamics during coastal spring algae blooms. eLife, 5: 1–31.

Thines, M., Nam, B., Nigrelli, L., Beakes, G., and Kraberg, A. 2015. The diatom parasite *Lagenisma coscinodisci* (Lagenismatales, Oomycota) is an early diverging lineage of the Saprolegniomycetes. Mycological Progress, 14: 75.

Tillmann, U., Hesse, K.-J., and Tillmann, A. 1999. Large-scale parasitic infection of diatoms in the Northfrisian Wadden Sea. Journal of Sea Research, 42: 255–261.

Tillmann, U., and Rick, H. J. 2003. North Sea phytoplankton: A review. Senckenbergiana Maritima, 33: 1–69.

Turner, J. T., Tester, P. A., and Hettler, W. F. 1985. Zooplankton feeding ecology - A laboratory study of predation on fish eggs and larvae by the copepods *Anomalocera ornata* and *Centropages typicus*. Marine Biology, 90: 1–8.

Utermöhl, H. 1958. Zur Vervollkommnung der quantitativen Phytoplankton-Methodik. Internationale Vereinigung für Theoretische und Angewandte Limnologie (Mitteilung No. 9), 9: 1–38.

Vaulot, D., Eikrem, W., Viprey, M., and Moreau, H. 2008. The diversity of small eukaryotic phytoplankton (≤3 µm) in marine ecosystems. FEMS Microbiology Reviews, 32: 795–820.

Vincent, F., and Bowler, C. 2020. Diatoms Are Selective Segregators in Global Ocean Planktonic Communities. mSystems, 5.

Walther, G.-R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T. J. C., Fromentin, J.-M., *et al.* 2002. Ecological responses to recent climate change. Nature, 416: 389–395.

Wangensteen, O. S., Palacín, C., Guardiola, M., and Turon, X. 2018. DNA metabarcoding of littoral hardbottom communities: High diversity and database gaps revealed by two molecular markers. PeerJ, 2018: 1–30.

Wasmund, N., Tuimala, J., Suikkanen, S., Vandepitte, L., and Kraberg, A. 2011. Long-term trends in phytoplankton composition in the western and central Baltic Sea. Journal of Marine Systems, 87: 145–159.

Widdicombe, C. E., Eloire, D., Harbour, D., Harris, R. P., and Somerfield, P. J. 2010. Long-term phytoplankton community dynamics in the Western English Channel. Journal of Plankton Research, 32: 643–655.

Williams, O. J., Beckett, R. E., and Maxwell, D. L. 2016. Marine phytoplankton preservation with Lugol's: a comparison of solutions. Journal of Applied Phycology, 28: 1705–1712.

Wiltshire, K. H., and Dürselen, C. D. 2004. Revision and quality analyses of the Helgoland Reede long-term phytoplankton data archive. Helgoland Marine Research, 58: 252–268.

Wiltshire, K. H., and Manly, B. F. J. 2004. The warming trend at Helgoland Roads, North Sea: Phytoplankton response. Helgoland Marine Research, 58: 269–273.

Wiltshire, K. H., Malzahn, A. M., Wirtz, K., Greve, W., Janisch, S., Mangelsdorf, P., Manly, B. F. J., *et al.* 2008. Resilience of North Sea phytoplankton spring bloom dynamics: An analysis of long-term data at Helgoland Roads. Limnology and Oceanography, 53: 1294–1302.

Wiltshire, K. H., Kraberg, A., Bartsch, I., Boersma, M., Franke, H. D., Freund, J., Gebühr, C., *et al.* 2010. Helgoland roads, North Sea: 45 years of change. Estuaries and Coasts, 33: 295–310.

Wiltshire, K. H., Boersma, M., Carstens, K., Kraberg, A. C., Peters, S., and Scharfe, M. 2015. Control of phytoplankton in a shelf sea: Determination of the main drivers based on the Helgoland Roads Time Series. Journal of Sea Research, 105: 42–52.

Winder, M., Berger, S. A., Lewandowska, A., Aberle, N., Lengfellner, K., Sommer, U., and Diehl, S. 2012. Spring phenological responses of marine and freshwater plankton to changing temperature and light conditions. Marine Biology, 159: 2491–2501.

Wolfe, D. A., Champ, M. A., Flemer, D. A., and Mearns, A. J. 1987. Long-term biological data sets: Their role in research, monitoring, and management of estuarine and coastal marine systems. Estuaries, 10: 181–193.

Wollschläger, J., Nicolaus, A., Wiltshire, K. H., and Metfies, K. 2014. Assessment of North Sea phytoplankton via molecular sensing: A method evaluation. Journal of Plankton Research, 36: 695–708.

Wollschläger, J., Wiltshire, K. H., Petersen, W., and Metfies, K. 2015. Analysis of phytoplankton distribution and community structure in the German Bight with respect to the different size classes. Journal of Sea Research, 99: 83–96.

Xiao, X., Sogge, H., Lagesen, K., Tooming-Klunderud, A., Jakobsen, K. S., and Rohrlack, T. 2014. Use of High Throughput Sequencing and Light Microscopy Show Contrasting Results in a Study of Phytoplankton Occurrence in a Freshwater Environment. PLoS ONE, 9: e106510.

Yang, H., Hu, Z., Shang, L., Deng, Y., and Tang, Y. Z. 2020. A strain of the toxic dinoflagellate *Karlodinium veneficum* isolated from the East China Sea is an omnivorous phagotroph. Harmful Algae, 93: 101775.

Yang, J. 2014. The role of ciliates in structuring the plankton communities of the German Bight based on ecological , experimental and taxonomic investigations School of Engineering and Science. Jacobs University, Bremen.

Yang, J., Löder, M. G. J., and Wiltshire, K. H. 2014. A survey of ciliates at the long-term sampling station ‘Helgoland Roads’, North Sea. Helgoland Marine Research, 68: 313–327.

Yang, J., Löder, M. G. J., Gerdts, G., and Wiltshire, K. H. 2015. Structural composition and temporal variation of the ciliate community in relation to environmental factors at Helgoland Roads, North Sea. Journal of Sea Research, 101: 19–30.

Yang, J., Zhang, X., Xie, Y., Song, C., Zhang, Y., Yu, H., and Burton, G. A. 2017. Zooplankton community profiling in a eutrophic freshwater ecosystem-lake tai basin by DNA metabarcoding. Scientific Reports, 7: 1–11.

Yang, J., Löder, M. G. J., Wiltshire, K. H., and Montagnes, D. J. S. 2021. Comparing the Trophic Impact of Microzooplankton during the Spring and Autumn Blooms in Temperate Waters. Estuaries and Coasts, 44: 189–198.

Yang, J. W., Wu, W., Chung, C.-C., Chiang, K.-P., Gong, G.-C., and Hsieh, C. 2018. Predator and prey biodiversity relationship and its consequences on marine ecosystem

functioning—interplay between nanoflagellates and bacterioplankton. The ISME Journal, 12: 1532–1542.

Yarimizu, K., Sildever, S., Hamamoto, Y., Tazawa, S., Oikawa, H., Yamaguchi, H., Basti, L., *et al.* 2021. Development of an absolute quantification method for ribosomal RNA gene copy numbers per eukaryotic single cell by digital PCR. Harmful Algae, 103: 102008.

Yeh, H. D., Questel, J. M., Maas, K. R., and Bucklin, A. 2020. Metabarcoding analysis of regional variation in gut contents of the copepod *Calanus finmarchicus* in the North Atlantic Ocean. Deep Sea Research Part II: Topical Studies in Oceanography, 180: 104738.

Yeh, Y. C., McNichol, J., Needham, D. M., Fichot, E. B., Berdjeb, L., and Fuhrman, J. A. 2021. Comprehensive single-PCR 16S and 18S rRNA community analysis validated with mock communities, and estimation of sequencing bias against 18S. Environmental Microbiology, 23: 3240–3250.

Yoch, D. C. 2002. Dimethylsulfoniopropionate: Its sources, role in the marine food web, and biological degradation to dimethylsulfide. Applied and Environmental Microbiology, 68: 5804–5815.

Zafeiropoulos, H., Viet, H. Q., Vasileiadou, K., Potirakis, A., Arvanitidis, C., Topalis, P., Pavloudi, C., *et al.* 2020. PEMA: a flexible Pipeline for Environmental DNA Metabarcoding Analysis of the 16S/18S ribosomal RNA, ITS, and COI marker genes. GigaScience, 9: 1–12.

Zamora-Terol, S., Novotny, A., and Winder, M. 2020. Reconstructing marine plankton food web interactions using DNA metabarcoding. Molecular Ecology, 29: 3380–3395.

Zamora-Terol, S., Novotny, A., and Winder, M. 2021. Molecular evidence of host-parasite interactions between zooplankton and Syndiniales. Aquatic Ecology, 55: 125–134.

Zarauz, L., and Irigoien, X. 2008. Effects of Lugol's fixation on the size structure of natural nano-microplankton samples, analyzed by means of an automatic counting method. Journal of Plankton Research, 30: 1297–1303.

Zhang, J., Kobert, K., Flouri, T., and Stamatakis, A. 2014. PEAR: A fast and accurate Illumina Paired-End reAd mergeR. Bioinformatics, 30: 614–620.

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 Microbiology Ecology, 52: 79–92.

Zimmermann, J., Glöckner, G., Jahn, R., Enke, N., and Gemeinholzer, B. 2015. Metabarcoding vs. morphological identification to assess diatom diversity in environmental studies. Molecular Ecology Resources, 15: 526–542.

Zingone, A., Harrison, P. J., Kraberg, A., Lehtinen, S., McQuatters-Gollop, A., O'Brien, T., Sun, J., *et al.* 2015. Increasing the quality, comparability and accessibility of phytoplankton species composition time-series data. Estuarine, Coastal and Shelf Science, 162: 151–160.

Zingone, A., D'Alelio, D., Mazzocchi, M. G., Montresor, M., Sarno, D., and Team, L.-M. 2019. Time series and beyond: multifaceted plankton research at a marine Mediterranean LTER site. Nature Conservation, 34: 273–310.

A Appendix

Supplementary Material to Publication I

Supplementary Materials and Methods

Bioinformatics processing

Sequence processing, operational taxonomic unit (OTU) clustering, and annotation was done with an internally developed pipeline at the Alfred Wegener Institute as described below, wrapping common bioinformatics tools and "*GNU parallel*" (Tange, 2011) for fast and massive parallel workflow execution. Raw sequences were processed by the tool *Trimmomatic*, version 0.38 (Bolger *et al*., 2014), which scanned each sequence from 5' to 3' and trimmed the 3'-end if in a sliding window of 4 bp the average quality dropped below a Phred Q-score of 10. Sequences passing this filter were only retained if their paired-end partner also passed. *PEAR*, version 0.9.10 (Zhang *et al.*, 2014) with default settings was used to merge the paired-end reads. Sequences which could not be merged were discarded.

To guarantee the same orientation, the sequences were filtered. Hereby, the sequence of the PCR forward primer had to occur before the sequence of the reverse complement of the reverse PCR primer. If sequences did not match this pattern, their reverse complement was also scanned. *Cutadapt,* version 1.17 (Martin, 2011) with the following settings was used for this task: for the forward primer sequence (19 bp long) a minimum sequence overlap of 16 bp and a maximum number of mismatches of 4 bp and for the reverse primer sequence (15 bp long) a minimum sequence overlap of 12 bp and a maximum number of mismatches of 3 bp were required to keep a sequence. Primer matching segments and additionally possible remaining artificial subsequences were finally truncated from the amplicon sequences. The remaining sequences were feature-filtered by *VSEARCH,* version 2.3.0 (Rognes *et al*., 2016): Sequences were discarded, i) if they were outside a 50 bp radius above or below the median length of the targeted amplicon (376 bp), i. e. below 326 bp or above 426 bp, ii) if they carried any ambiguity or iii) if the expected base error (sum of all base error probabilities) of a sequence was above 0.5.

Each sample was de-replicated individually (abundances of each amplicon kept in the sequence headers) and chimera were sample-wise predicted *de novo* by the tool *VSEARCH* (version 2.3.0) with default settings and removed from the sample files. Only samples with at least 10000 sequences after filtering were considered for further analyses (49 out of 50 samples). Cleaned sample files were pooled and now de-replicated in total while total amplicon abundances were kept in the sequence headers. Finally, about 4.3

million sequences (if one re-replicates the pooled amplicons) survived all filtering steps and were used as input of the OTU-clustering. OTU-clustering was done by the tool *Swarm*, version 2.1.8 (Mahé *et al.*, 2014, 2015) with default settings. Scripts needed to create the OTU-table from the *Swarm*-clustering output and from the abundance information in the headers of the chimera-cleaned sequence files were taken from *https://github.com/torognes/swarm/wiki/Working-with-several-samples*, adjusted and executed.

The most abundant amplicon of each OTU cluster was used as representative for the respective OTU. These sequences were annotated with the default classifier implemented in *mothur*, version 1.38.1 (Schloss *et al.*, 2009) with the *Protist Ribosomal Reference database* (*PR2*), version 4.10 (Guillou *et al.*, 2013) as reference set and a confidence cut-off of 90. The reference set was primer filtered and truncated the same way as the queries. The annotation of a representative sequence was used as annotation of the full OTU cluster and the annotation was added to the corresponding line of the OTU table.

Supplementary Figures

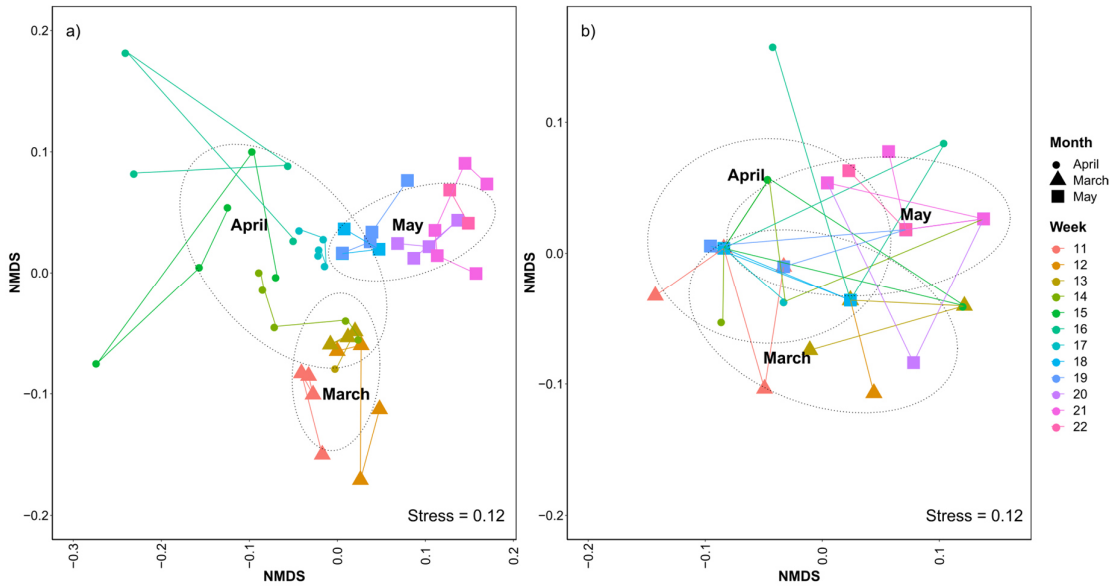


Fig. SI. Two-dimensional NMDS plots of community dissimilarities on a) genus level and b) phylum level based on presence-absence data of the respective taxon. Dissimilarity matrices are based on Bray-Curtis dissimilarities. Samples belonging to the same month are depicted by the same shape of points, colour coding shows weekly variations, lines connect the sampling days in weekly chronology. Ellipses were drawn at 80% confidence level.

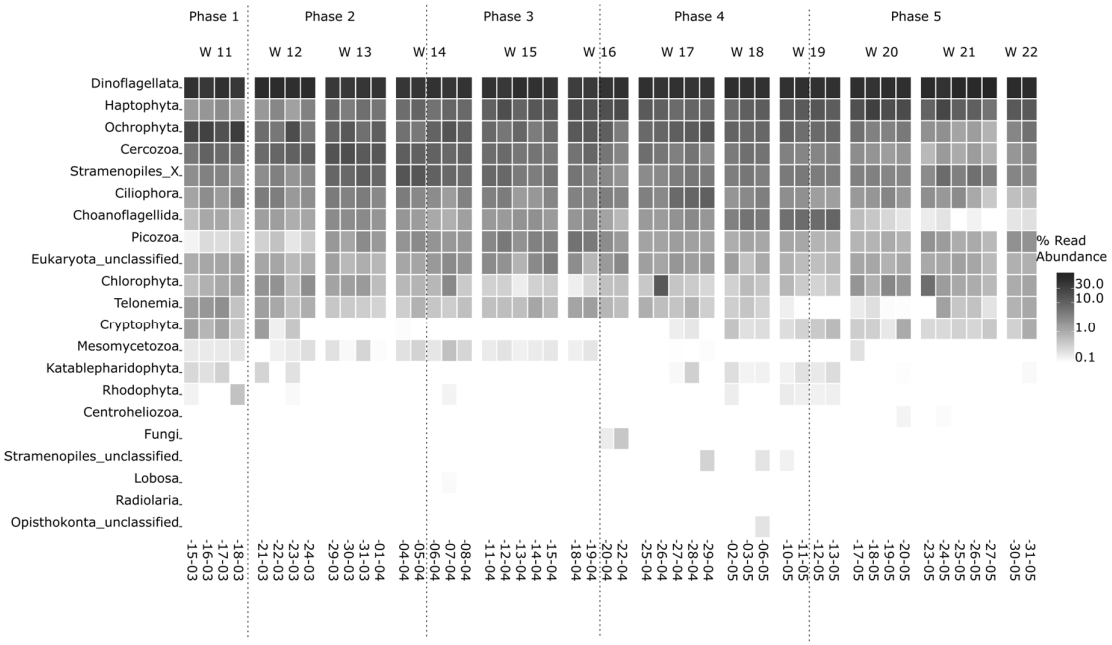


Fig. SII. Community composition on phylum level from March 15 to May 31, 2016, phyla are displayed according to the sum of their abundances in all samples.

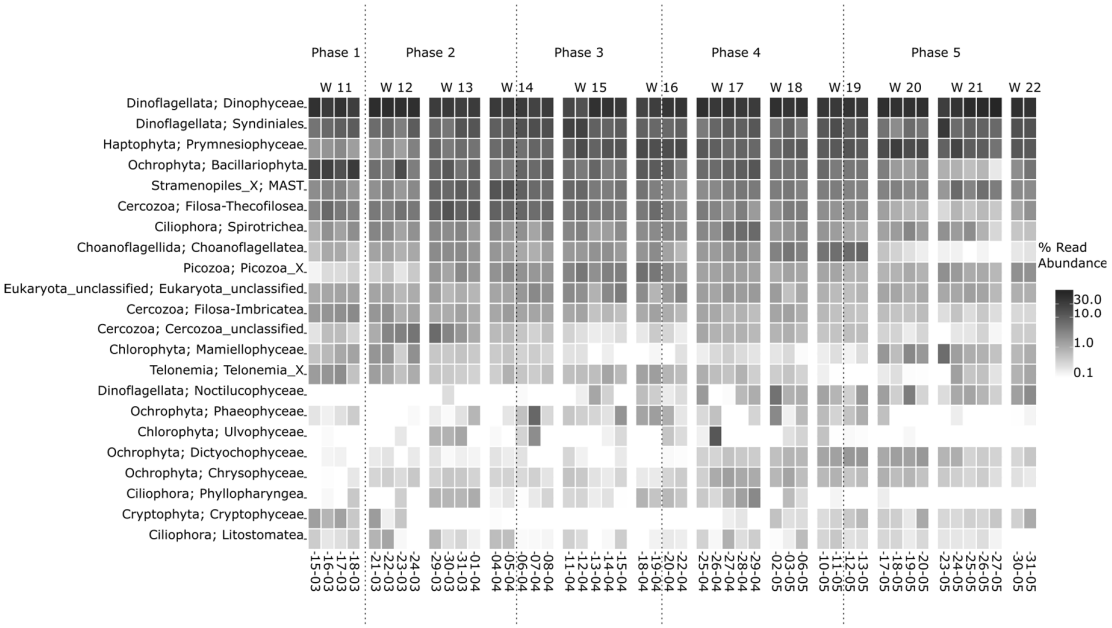


Fig. SIII. Community composition of the 22 most abundant classes from March 15 to May 31, 2016. First taxa is given on phylum level, second taxa on class level. Classes are displayed according to the sum of their abundances in all samples.

Supplementary Material to Publication II

Supporting Information: Additional information on bioinformatic pipeline and analysis

Material and Methods

The chosen pipeline, which is focusing on a relative strict parameter set and a high confidence cut-off of annotation (here named as “default”) aiming for a high reliability, was compared with other settings of the same tools. Additionally, for clustering into operational taxonomic units (OTUs) by the tool swarm (version 2.2.2) [1,2], we compared the results of the default settings with changed up d values (Table 2). Prior to comparison, the threshold of 0.001% of all reads was applied in each dataset.

OTU tables of these additional parameter settings can be found on zenodo.org (DOI: 10.5281/zenodo.4319940).

Table 1. Different settings and steps of the bioinformatic pipeline. The term “default” refers to the settings as they have been chosen for this study.

tool and action	tool version and reference	"relaxed" settings	"default" settings	"strict" settings	"very strict" settings
trimming with trimmomatic (length sliding window - average quality score within)	version 0.38 [3]	3-5	3-8	1-15	1-15
merge with vsearch (length of minimum overlap of paired end reads - max number of mismatches allowed)	version 2.3.0 [4]	25-5	50-5	50-0	50-0
primer removal with cutadapt (max [x*100]% mismatches in at least [y*100]% primer-to-sequence overlap)	version 1.19 [5]	0.2-0.75	0.1-0.9	0.1-0.9	0-1
eemax-filtration with vsearch (expected error of x allowed in sequence)	version 2.3.0	1	0.25	0.1	0.1
length filtration with vsearch (min len x - max len y)	version 2.3.0	300-550	300-550	300-550	300-550
classification with mothur (cutoff value of taxonomic level)	version 1.38.1 [6]	0.6	0.8	0.9	0.9

Table 2. Combinations of different settings for comparison of results. and steps of the bioinformatic pipeline. The term “default” refers to the settings as they have been chosen for this study.

	chosen d value in swarm	parameter settings (relaxed, default, strict, very strict)
option 1 (this study)	1	default
option 2	2	default
option 3	3	default
option 4	5	default
option 5	10	default
option 6	1	relaxed
option 7	2	relaxed
option 8	3	relaxed
option 9	1	strict
option 10	2	strict
option 11	3	strict
option 12	1	very strict

Results

Higher d values resulted in less OTUs in the datasets, since potential OTUs that are found with lower d values are merged with other close OTUs when distances are too big. Except for the very strict settings, which filters all sequences that do not fit 100% in primer-to-sequence overlap, parasitoid sequences of all 10 phyla were found (Table 3). Identification of the sequences was not possible for the Perkinsea sequence, when (very) strict settings with different d values were used (option 10-12). Additionally, in option 9-12 (marked in red) some samples were removed by the pipeline due to poor quality. For all options, new sequences that were identified as parasitoids were found in the dataset after the 0.001% threshold. In general, the distribution of parasitoids in the different phyla remained similar. For example, parasitoid Dinoflagellata distributed the most OTUs, followed by Cercozoa parasitoids. Additionally, in all options the three most abundant OTUs (including non-parasitoids) belonged to the genera *Paracalanus*, *Temora* and *Gyrodinium*.

Table 3. Comparison of results of different pipeline settings. “Old” sequences refer to sequences that are present in the dataset of the 2790 OTUs of the study (option 1).

	number of OTUs (at 0.001% threshold)	number of samples	total Parasitoids (0.001%)	number of "old" / "new" parasitoid sequences found	number of phyla (sequences / identified)	comment
option 1 (this study)	2790	280	461		10/10	
option 2	2274	280	414	405/9	10/10	
option 3	2125	280	381	370/11	10/10	
option 4	1921	280	340	324/16	10/10	
option 5	1539	280	274	258/16	10/10	
option 6	2904	280	491	403/88	10/10	
option 7	2313	280	431	353/78	10/10	
option 8	2165	280	405	324/81	10/10	
option 9	2522	267	466	423/43	10/10	
option 10	2208	267	410	375/35	10/9	Perkinsea sequence found but not identified by PR2
option 11	2068	267	379	345/34	10/9	Perkinsea sequence found but not identified by PR2
option 12	2200	266	419	390/29	09/08	parasitic Metazoa sequences not found/identified, Perkinsea sequence found but not identified by PR2

References

1. Mahé F, Rognes T, Quince C, de Vargas C, Dunthorn M. Swarm v2: highly-scalable and high-resolution amplicon clustering. PeerJ. 2015;3: e1420. doi:10.7717/peerj.1420

2. Mahé F, Rognes T, Quince C, de Vargas C, Dunthorn M. Swarm: robust and fast clustering method for amplicon-based studies. PeerJ. 2014;2: e593. doi:10.7717/peerj.593

3. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30: 2114–2120. doi:10.1093/bioinformatics/btu170

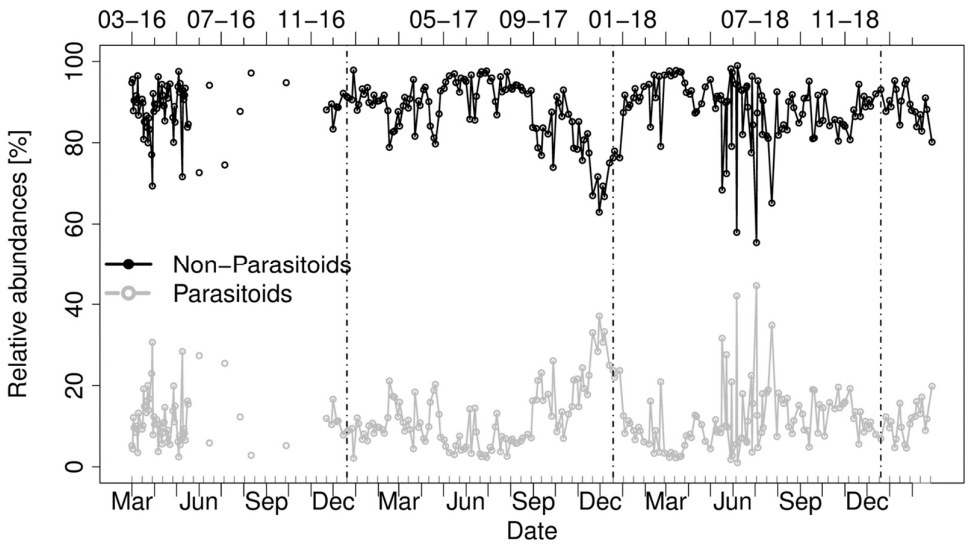
4. Rognes T, Flouri T, Nichols B, Quince C, Mahé F. VSEARCH: a versatile open source tool for metagenomics. PeerJ. 2016;4: e2584. doi:10.7717/peerj.2584

5. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal. 2011;17: 10–12. doi:https://doi.org/10.14806/ej.17.1.200

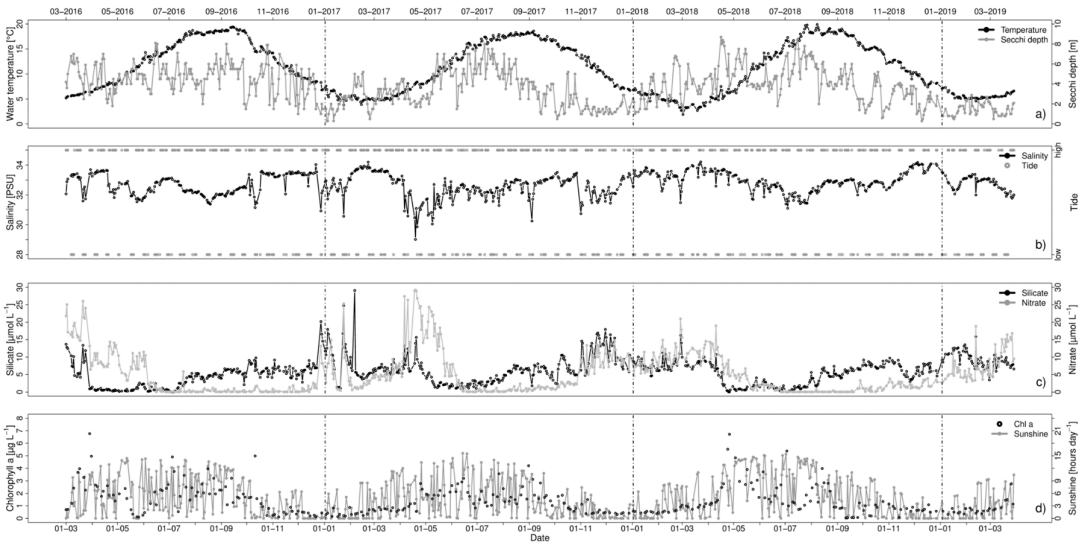
6. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, *et al.* Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for

Describing and Comparing Microbial Communities. Appl Environ Microbiol. 2009;75: 7537–7541. doi:10.1128/AEM.01541-09

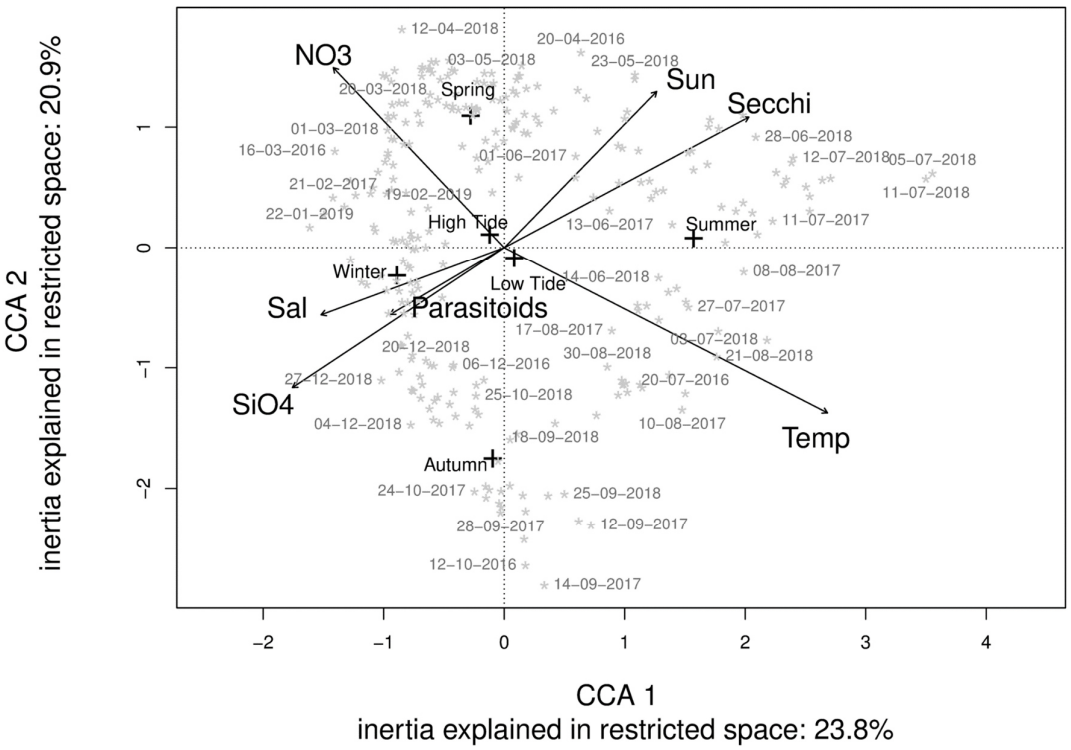
Supporting Information: Figures



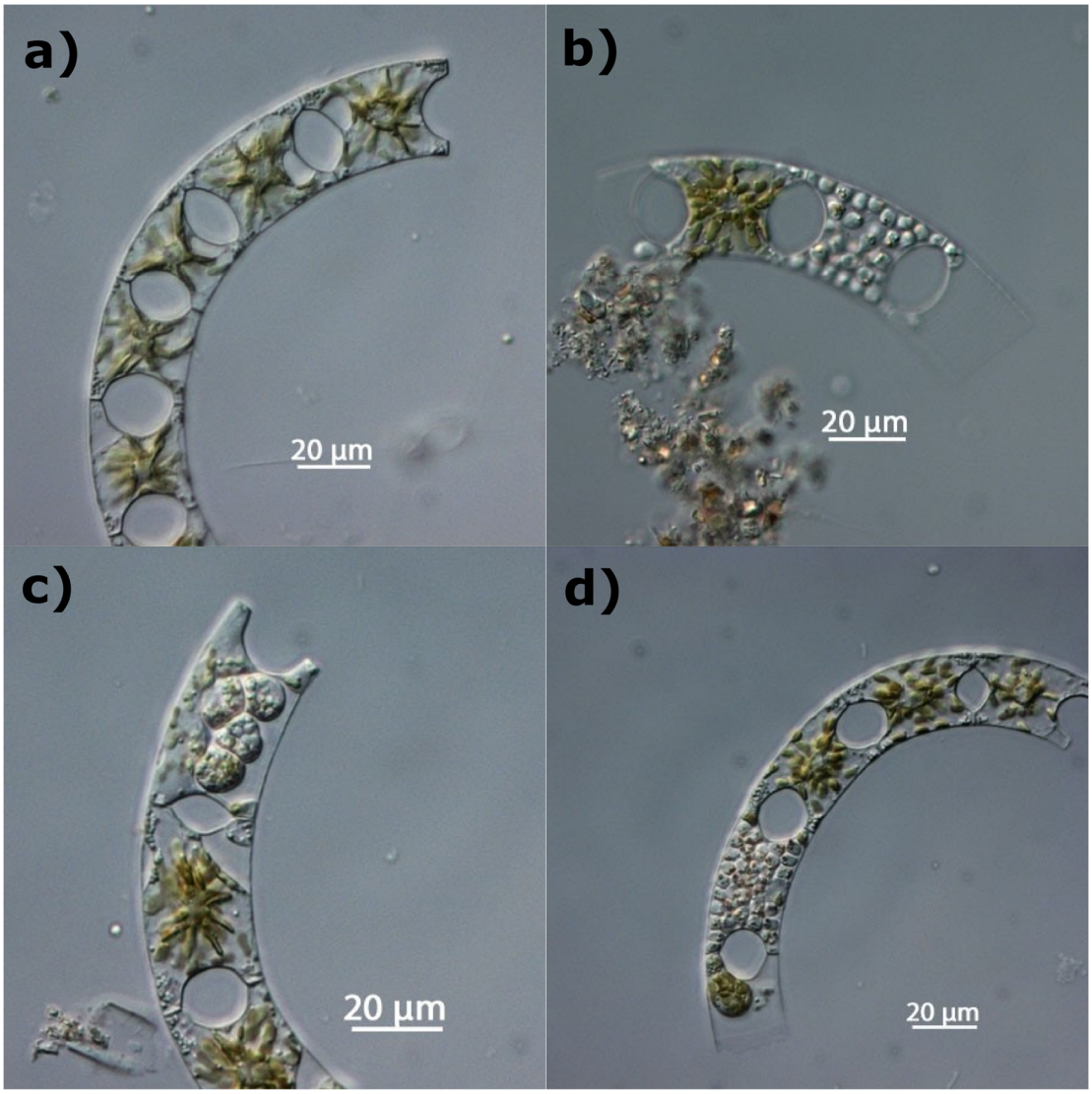
S1 Fig. Relative abundances [%] of parasitoids and non-parasitoid OTUs. Non-parasitoid OTUs include all remaining OTUs, that were not identified as Parasitoids; Vertical lines indicate turn of the years.



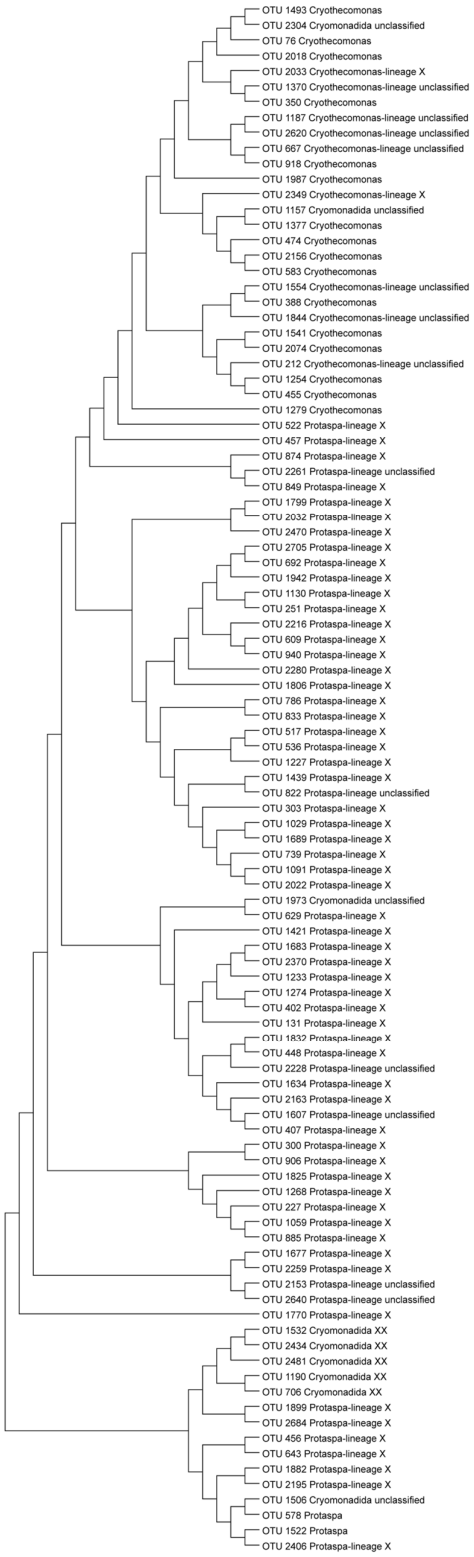
S2 Fig. Overview of environmental conditions, a) water temperature, Secchi depth, b) Salinity, Tide, c) Silicate, Nitrate, d) Chlorophyll a, Sunshine duration from March 2016 to March 2019. Vertical lines indicate turn of the years. Note the different scaling of the axes.



S3 Fig. Canonical Correspondence Analysis (CCA) of the samples (grey asterisks with sampling date) including significant parameters in black: Temperature (temp), salinity (sal), silicate (SiO4), nitrate (NO3), sunshine duration (sun), total parasitoid occurrence (parasitoids), seasons (spring, summer, autumn, winter) and tide (low tide, high tide). 12.2% of total inertia could be explained by all variables in full space, in restricted space CCA1 explained 23.8% of the variance and CCA2 explained 20.9%.



S4 Fig. Live cells of the centric diatom Eucampia zodiacus collected at Helgoland Roads, a) without parasitic infection (3rd August 2017), b)-d) with parasitic infection (b) 27th July 2017, c-d) 29th August 2017). Figures retrieved from planktonnet.awi.de.



S5 Fig. Maximum likelihood tree of Cryomonadida OTUs. The evolutionary history was inferred by using the Maximum Likelihood method and Tamura-Nei model [101]. The tree with the highest log likelihood (-3807.40) is shown. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Tamura-Nei model, and then selecting the topology with superior log likelihood value. This analysis involved 101 nucleotide sequences. There were a total of 397 positions in the final dataset. Evolutionary analyses were conducted in MEGA X [100].

Supplementary Material to Publication III

Supplementary Results

Successional patterns and diversity of kingdoms

Stramenopiles, with 551 OTUs, reached relative sequence abundances up to 56.8 percent (Suppl. Figure 1b). Further identified taxa were subdivided in the phylum Ochrophyta and in Stramenopiles_X (Table 2). Relative sequence abundances of up to 60.7 percent were found for Archaeplastida (94 OTUs). These included the phyla Chlorophyta and Streptophyta (Table 2). 427 OTUs were identified as Rhizaria, while relative sequence abundances reached up to 15.9 percent (Suppl. Figure 1c). Taxa were divided into two phyla: Cercozoa and Radiolaria (Table 2). 140 OTUs were assigned to the Hacrobia and up to 37.5 percent relative abundance was found. The kingdom included six Phyla: Centroheliozoa, Cryptophyta, Haptophyta, Katablepharidophyta, Picozoa and Telonemia (Table 2). Amoebozoa and Apusozoa with 27 and 16 OTUs each only reached maximum relative sequence abundances of up to 3.3 and 0.5 percent (Suppl. Figure 1d). Additionally, these two kingdoms were the only ones that could not be found in every sample (247 and 233 samples, respectively). While the Amoebozoa included only one phylum namely Lobosa, the Apusozoa included Apusomonadidae, Hilomonadea and Mantamonadidea (Table 2). 142 OTUs could not be classified for any kingdom (Eukaryota unclassified), relative sequence abundances of up to 11.7 percent are therefore of unknown taxonomy on any lower level (Suppl. Figure 1e).

Successional patterns and diversity of potential consumers of phytoplankton

Meso- and macrozooplankton

Arthropoda were the most abundant Metazoa. They included only 1 OTU of Chelicerata and 339 OTUs of Crustacea, of which only 11 OTUs belonged to Branchiopoda, Malacostraca and Ostracoda. All other OTUs were identified as Maxillopoda, of which most OTUs belonged to **copepod** genera.

Most OTUs belonged to the genera *Pseudocalanus* (166 OTUs) and *Calanus* (85 OTUs), however these OTUs had low relative Crustacea sequence abundance, 3.83 and 1.87 %, respectively (Suppl. Table 5). Highest relative Crustacea sequence abundances were found for *Paracalanus* (4 OTUs with 34.83%), *Temora* (9 OTUs with 31.34%) and *Oithona* (4 OTUs with 11.39%). 99.9% of the *Paracalanus* reads thereby belonged to OTU 1, which had the highest total read number of the whole dataset.

Paracalanus (4 OTUs) was always abundant, with the lowest relative sequence abundances occurring especially in May, and the highest relative sequence abundances in February 2018 (Figure 2). *Temora* (9 OTUs) was also always highly abundant, except for autumn months. *Oithona* (4 OTUs) was not found in high relative sequence abundances in 2016 and 2019 and had its highest relative sequence abundances in summer 2017 and spring 2018. *Calanus* (85 OTUs) had sporadic peaks in spring 2016, winter 2018/2019 and was also present in other winters, as well as in spring and autumn 2018 (Figure 2). *Pseudocalanus* (166 OTUs) was peaking in spring 2016, 2017 and 2018 and had an additional small peak in June 2016. *Centropages* (12 OTUs) relative sequence abundances were highest in winter 2016/2017 and winter 2018/2019, with lower peaks occurring in summer and autumn 2017 and 2018 and in spring 2018 (Figure 2). High relative sequence abundances for *Tachidius* (1 OTU) were found in autumn 2017 and winter 2018/2019.

Several genera of other high abundant Metazoa also reached relative sequence abundances of above 10% (Suppl. Table 3). Two genera each of Cnidaria and Craniata had one single high peak in abundance during the whole timeframe: *Cyanea* (1 OTU) in April 2018, *Metridium* (1 OTU) in June 2018, *Kareius* (1 OTU) in February 2017 and *Gadus* (1 OTU) in March 2018. Four taxa belonged to Annelida, which showed a distinct succession throughout the years. *Polycirrus* (3 OTUs) was mainly present during Winter (December 2016, November to December 2017 & 2018). It was followed by *Nephtys* (3 OTUs), which was present in January to March in 2017, 2018 & 2019 as well as in March 2016. Nereis (2 OTUs) was also present in March of all years and in April 2016 and 2017. *Lanice* (1 OTU) had its highest peaks in April and May in 2016 and 2017 but was also present in January 2019, February 2017 and in March 2016, 2017 and 2019. *Cephalotrix* (4 OTUs, Nemertea) did not occur in high relative sequence abundances except for two individual peaks, in August and September 2017, respectively. In contrast, *Microstomum* (1 OTU, Platyhelminthes) had several high peaks in abundance in spring and summer 2017. *Hiatella* (2 OTUs, Mollusca) peaked individually in spring 2016 and all following winters. *Phoronis* (3 OTUs, Brachiopoda) had several high peaks in abundance in all three years from end of spring till autumn, except for 2016, where only small peaks were detected in spring.

Microzooplankton

Out of the five taxa summarized on genus level, which occurred in every sample (Suppl. Table 2), three taxa could be identified as a distinct genus. All three distinct genera belonged to the Class of Dinophyceae (*Gymnodinium*, *Heterocapsa* and *Gyrodinium*) and were either mixotrophic or heterotrophic. *Gyrodinium* and *Heterocapsa* also belonged to the genera that reached relative sequence abundances of above 20 % in at least one sample (Suppl. Table

4). However, all genera reached these high relative sequence abundances only occasionally. During most times, relative sequence abundances were even below 5%.

The most abundant **Dinoflagellata** belonged to the order of Gymnodiniales (Dinophyceae) and reached over 30 percent of relative Dinoflagellata sequence abundance (Suppl. Table 7). A high number of Dinophyceae OTUs could not be identified further, leaving up to 25% of relative Dinoflagellata sequence abundance without any information on family and lower levels. Several Syndiniales Groups (e.g. Dino-Group-II, Dino-Group-I) were present with a high number of OTUs and a high relative sequence abundance.

Regarding mixotrophic Dinophyceae the most abundant genera were *Lepidodinium* (1 OTU), *Heterocapsa* (4 OTUs), *Alexandrium* (4 OTUs), *Akashiwo* (1 OTU), *Tripos* (5 OTUs), *Gymnodinium* (10 OTUs) and *Prorocentrum* (9 OTUs).

Peaks of *Heterocapsa* with relative sequence abundances above 10% were found in spring 2016, 2017 and 2018 and in summer 2018, lower peaks were also found during other seasons such as summer 2016, winter 2016/2017 and autumn 2017 and 2018. *Alexandrium* was most abundant in summer 2018, and present with lower relative sequence abundances in autumn 2017 and 2018. *Akashiwo* was being mostly present in autumn 2016 and 2017. *Gymnodinium* only reached such high relative sequence abundances during spring 2016, with smaller peaks occurring throughout the years (Figure 2). *Tripos* was present during summer 2016 and autumn 2018 with relative sequence abundances above 10%, in summer 2017 and 2018 as well as winter 2017/2018 and winter 2018/2019 smaller peaks occurred. For *Prorocentrum* one high peak was found in July 2016, additional lower peaks in abundance occurred especially in winter 2016/2017 and autumn 2018 (Figure 2). While the mixotrophic genus *Lepidodinium* (OTU 30) was also found in summer of 2016 and 2017, it reached its highest relative sequence abundances with peaks above 10% during summer 2018, more specifically in June and August.

One heterotrophic Dinophyceae genus, *Gyrodinium* (16 OTUs), reached relative sequence abundances above 10% and was abundant throughout all years (Figure 2). Except for summer 2018 it also always exceeded the mixotrophic *Lepidodinium* in abundance. *Noctiluca* (1 OTU) was abundant in spring and summer in 2016, 2017 and 2018. Peaks above 10% were found in summer 2018 in June, July and August and in July 2017.

Most **Ciliophora** belonged to the Class of Spirotrichea (100 OTUs, comprising Choreotrichida, Euplotia, Hypotrichia, Strombidiida and Tintinnida) with over 75 % of relative Ciliophora sequence abundance (Suppl. Table 8). Highest OTU diversity was found for *Askenasia* sp. (CONThreeP) with 25 OTUs. Compared to the high relative sequence abundances of Crustacea or Dinoflagellata genera, Ciliophora genera were generally rare.

On genus level, the highest relative sequence abundances (above 5%) were found in the class Spirotrichea. For the spirotrich order Tintinnida the highest relative sequence abundances were caused by the genus *Tintinnidium* with 5.99% (4 OTUs). Only one genus (*Leegaardiella*, 5 OTUs, Order Choreotrichida) was found to be abundant over 10% in at least one sample. In 2016, *Leegaardiella* was present from March to April, in 2017 from January till June, in 2018 from January till May, with an additional peak in July and in 2019 in February and March.

Additionally, **heterotroph flagellates** such as two genera of Chrysophyceae were highly abundant, *Spumella* (1 OTU) in spring and autumn 2018 and *Paraphysomonas* (7 OTUs) in winter 2017.

Successional patterns and diversity of autotroph and small mixotroph prey

The Stramenopiles comprised 209 OTUs of Bacillariophyta (Figure 2). Out of the taxa identifiable to genus level, the diatoms *Chaetoceros* and *Thalassiosira* showed the highest diversity in OTUs, with 29 and 21 OTUs, respectively. Maximum relative sequence abundances were found in early spring and summer 2016, spring and summer 2017 and spring 2018. The peak in early spring 2016 was mainly caused by two OTUs namely *Ditylum* (OTU 84), which reached a relative sequence abundance of above 10% during that time, and *Pseudo-nitzschia* (OTU 46). For summer 2016 *Rhizosolenia* (OTU 39, 179) was found to be the main contributor with peaks above 10%, other contributing taxa included *Actinocyclus* (OTU 256), *Thalassiosira* (OTU 264, 60), *Skeletonema* (OTU 308), *Chaetoceros* (OTU 120, 278) and *Ceratulina* (OTU 92). The peak in spring 2017 was mainly caused by *Thalassiosira* (OTU 107, 132, 487) with over 10% of relative sequence abundance and unclassified genera (OTU 68, 196). The peaks in summer 2017 were more diverse and comprised different diatoms, such as *Leptocylindrus* (OTU 114), *Rhizosolenia* (OTU 39, 93), *Ceratulina* (OTU 92), *Eucampia* (OTU 338), *Tenuicylindrus* (OTU 139, formerly known as *Leptocylindrus*) and unclassified genera (OTU 61, 755). In spring 2018 the peaks were mainly caused by *Coscinodiscus* (OTU 113) with over 10% of relative sequence abundances, by *Chaetoceros* (OTU 507, 510, 717), *Thalassiosira* (OTU 60) and additional unclassified genera (OTU 246, 87).

Other autotroph genera, which reached a relative sequence abundance of 10% in at least one sample, included two green macroalgal taxa belonging to the class of Ulvophyceae, two to the Prymnesiophyceae and one to Trebouxiophyceae.

While *Ulva* (5 OTUs) was only present occasionally and in low relative sequence abundances, it had one peak above 10% in April 2016. The second green algal taxon of the genus *Dilabifilum* (5 OTUs) was found more often. It was present in low relative sequence

abundances occasionally during all years but reached relative sequence abundances above 6% and higher a few times in spring and summer 2018.

Highly abundant Prymnesiophyceae were *Emiliania* (1 OTU) and *Phaeocystis* (5 OTUs), with both being present most of the time. Peaks above 10 percent were reached during spring months, lower peaks were also found during summer or autumn. During spring 2016 and 2018 *Phaeocystis* reached its peak before *Emiliania*. In spring 2017 both were not found in high relative sequence abundances, however *Emiliania* peaked (above 6%) during summer of that year.

For *Picochlorum* (4 OTUs, class Trebouxiophyceae) peaks of 5% or below were found in 2017, however it reached massive peaks during 2018, peaks above 10% were found in spring and summer of that year, with a maximum of over 56% in May 2018.

Mixotrophic genera included *Chrysochromulina* (25 OTUs) and Chrysophyceae Clade-C (2 OTUs). *Chrysochromulina* had one high peak in June 2017 and Chrysophyceae Clade-C had one high peak in February 2019.

Beta diversity

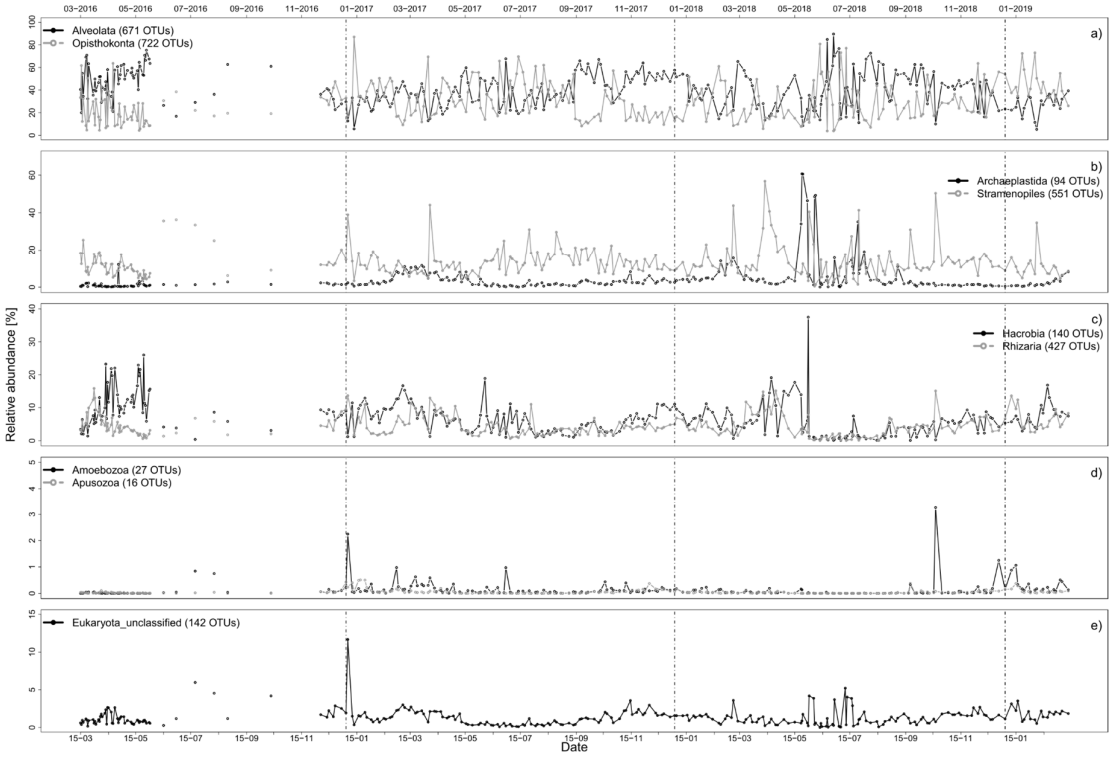
The beta diversity matrix was divided into five branches at h = 0.8. Cutting at h = 0.5 revealed 35 clusters (Suppl. Figure 5 and 6). Branch 1 (Suppl. Figure 5) included mostly spring samples from 2016 (49), 2017 (19) and 2018 (18) as well as 2 samples from Summer 2017 (June) and 4 samples from winter 2017/2018 (February). At h = 0.5 branch 1 was divided into 7 clusters. The first cluster included samples from spring 2018 and winter 2017/2018, which indicate an earlier begin of spring communities in 2018. While cluster 2 included samples from spring 2017 and 2018, cluster 4 only included 3 samples from spring 2018. The samples from summer 2017 were clustered with samples from spring 2017 (cluster 5), which indicates a delayed spring community period in 2017. The higher resolved spring 2016 was divided into 3 different clusters (cluster 3,6 and 7).

Branch 2 (Suppl. Figure 5) included samples from all seasons, except for samples from spring 2016 and spring 2018. It was divided into 12 clusters (8-19) and comprised the two biggest clusters (11,9). Cluster 11 comprised samples from spring 2017 (March), winter 2016/2017 and winter 2018/2019 and cluster 9 consisted of samples from autumn 2017, autumn 2018, winter 2016/2017 and winter 2018/2019. Other clusters were also in accordance to similar timeframes. For example, cluster 19 included samples from autumn of all three years, as well as from summer 2016 and summer 2017.

Branch 3, 4 and 5 (Suppl. Figure 6) only consisted of small clusters (16 clusters, maximum 6 samples), with 7 clusters comprising only 1 sample. In general, these small clusters

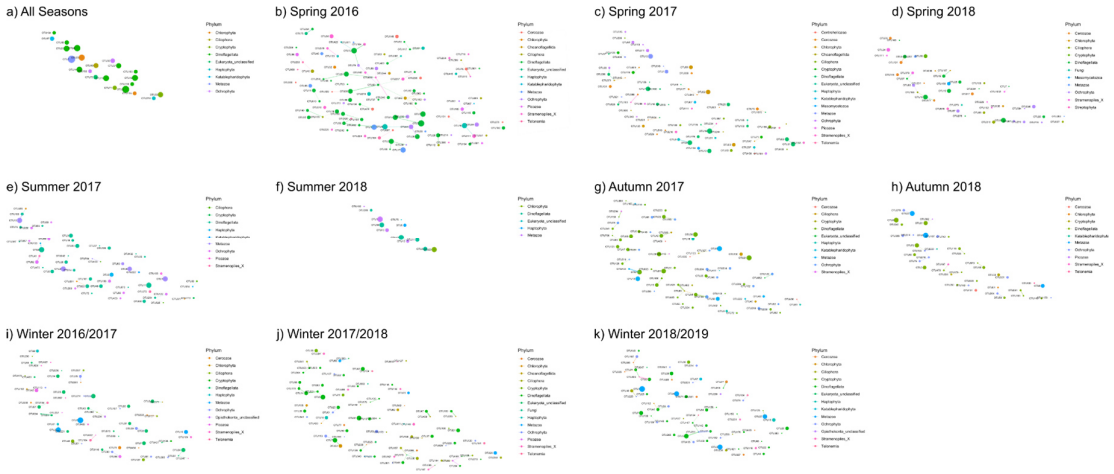
included samples from 2018, indicating rapid community changes as well as different community compositions in spring and summer of that year. Additionally four clusters included samples from one season each (winter 2017/2018, autumn 2018, summer 2016 and summer 2017).

Supplementary Figures

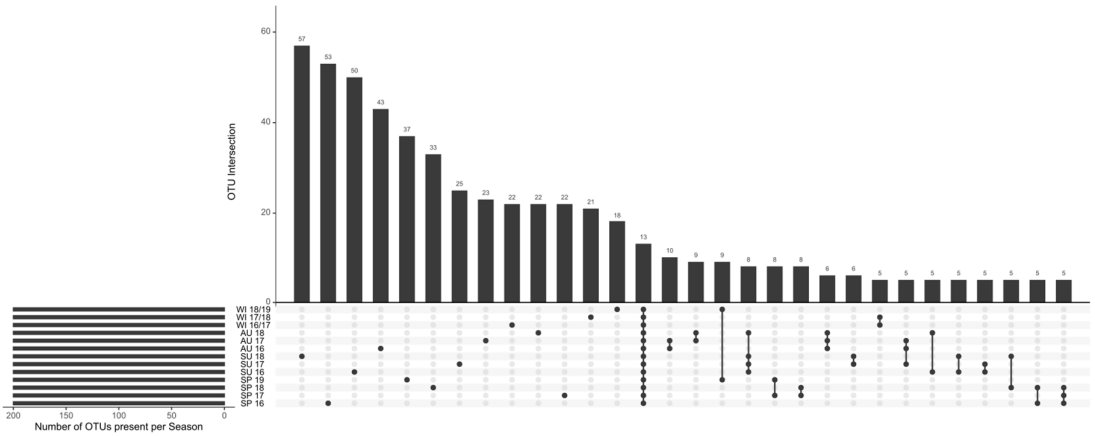


Suppl. Figure 1: Relative abundance of each kingdom from March 2016 to March 2019, a) Alveolata and Opisthokonta, b) Archaeplastida and Stramenopiles, c) Hacrobia and Rhizaria, d) Amoebozoa and Apusozoa, e) Eukaryota_unclassified; kingdoms sorted by relative abundance, perpendicular dotted lines mark the transition into a new year.

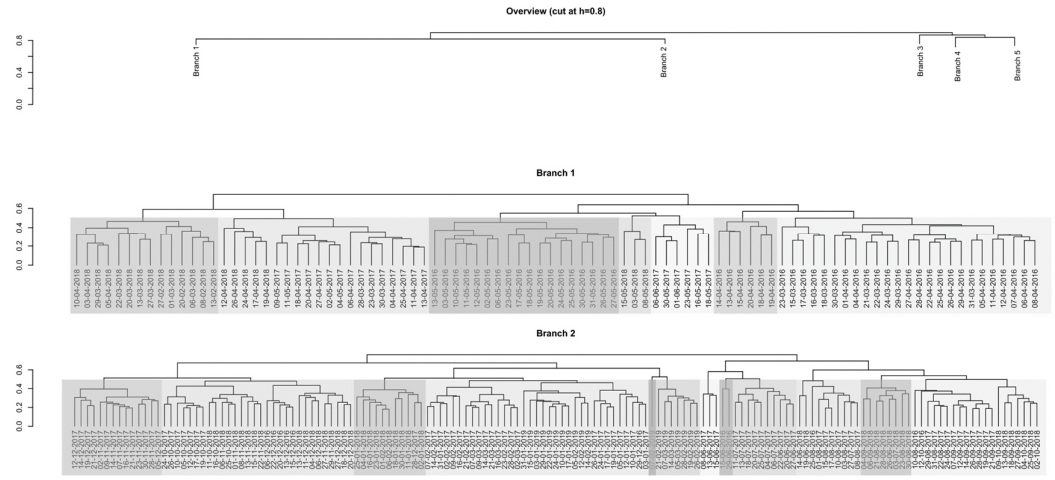
Suppl. Figure 2: Relative abundance of a) Ciliophora and Dinoflagellata without Syndiniales, b) *Copepoda,* c) *Acartia* spp. (3 OTUs), and d) *Acartia* spp. (OTU 429 & 732) and *Acartia* sp. (OTU 564) from March 2016 to March 2019; perpendicular dotted lines mark the transition into a new year.



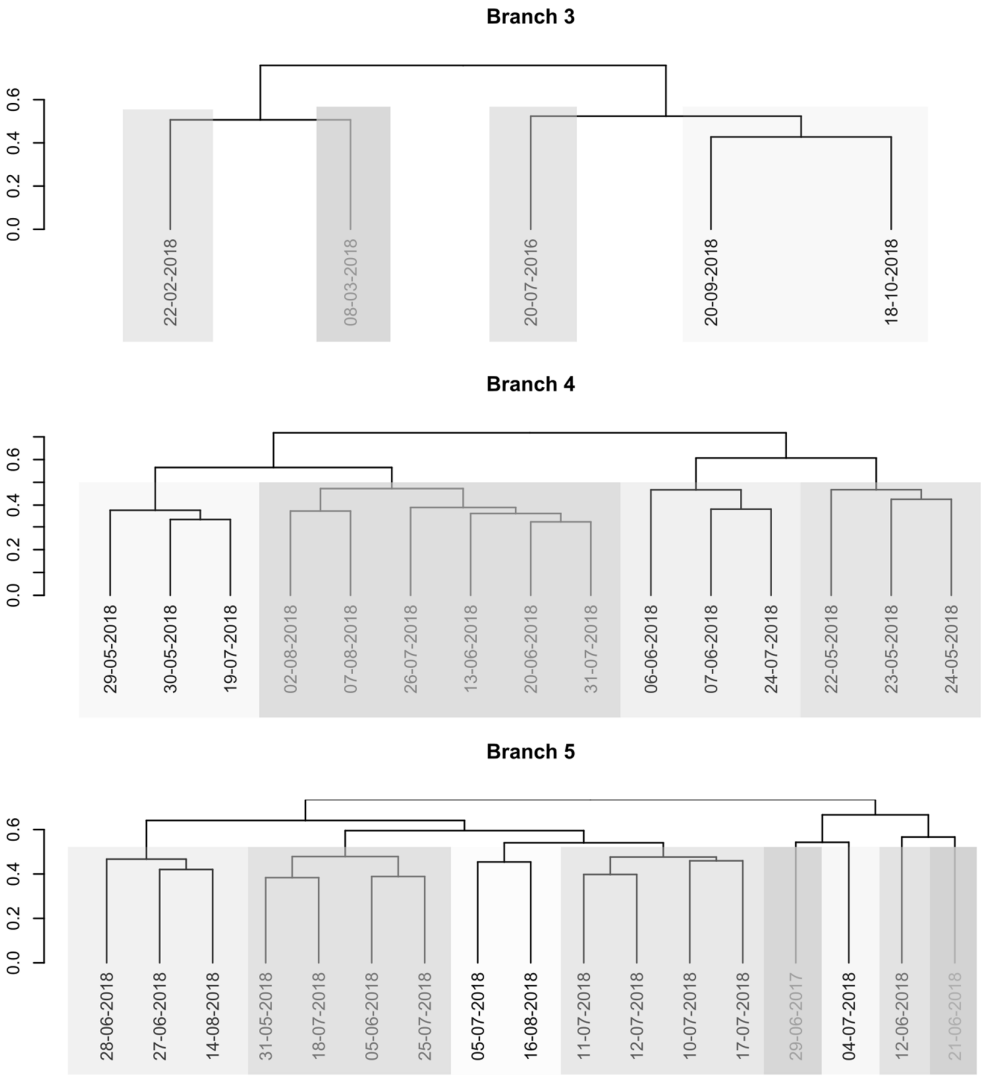
Suppl. Figure 3: Co-occurrence networks of all samples and per season; displayed are only connections with negative edges, colour code shows the associated phylum for each OTU. The width of the edges display edge weights. Node sizes were set proportional to the relative sequence abundances of the respective OTUs.



Suppl. Figure 4: Comparison of shared OTUs of the 200 most abundant OTUs per season, OTU Intersection displays the number of shared OTUs; SP = Spring, SU = Summer, AU = Autumn, WI = Winter, display of intersections was limited to at least 5 intersections.



Suppl. Figure 5: Visualized beta diversity matrix: Overview of the different branches, Branch 1 (cluster 1 to 7) and 2 (cluster 8 to 19). It was calculated using the betadiver function (vegan package) and Whittaker index; visualization of the matrix was done with the hclust function.



Suppl. Figure 6: Visualized beta diversity matrix: Branch 3 (cluster 20 to 23), 4 (cluster 24 to 27) and 5 (cluster 28 to 35). It was calculated using the betadiver function (vegan package) and Whittaker index; visualization of the matrix was done with the hclust function.

Supplementary Material to Publication IV

Supplementary Text

Supplementary Results

Alpha diversity at genus level and general overview

Bacillariophyta

For Bacillariophyta 27 genera were found by all methods (Figure 1, Suppl.Table 1). Two additional genera, that were present in the counts, were found in the dataset of the lower threshold. In total, 19 taxa were only found by metabarcoding, whereas 11 taxa were only found by counts. Three genera that are usually counted were only found in the sequences (details on all genera in Suppl. Table 1). Six genera, that were counted by microscopy, could only be revealed by BLAST alignment, as PR2 could not provide a sufficient identification. These included inter alia the newly counted or recorded *Plagiolemma* and *Mediopyxis*. Three genera, that were present in the counts, only had 1 OTU each with under 50 reads, and therefore were not detected in the MB datasets, whereas two genera were not found by metabarcoding at all. One genus has been reported since 2009, but was not part of the regular counts, while three genera are not regularly counted and have only been reported before 2009. Twelve genera have never been reported before at Helgoland.

Dinoflagellata

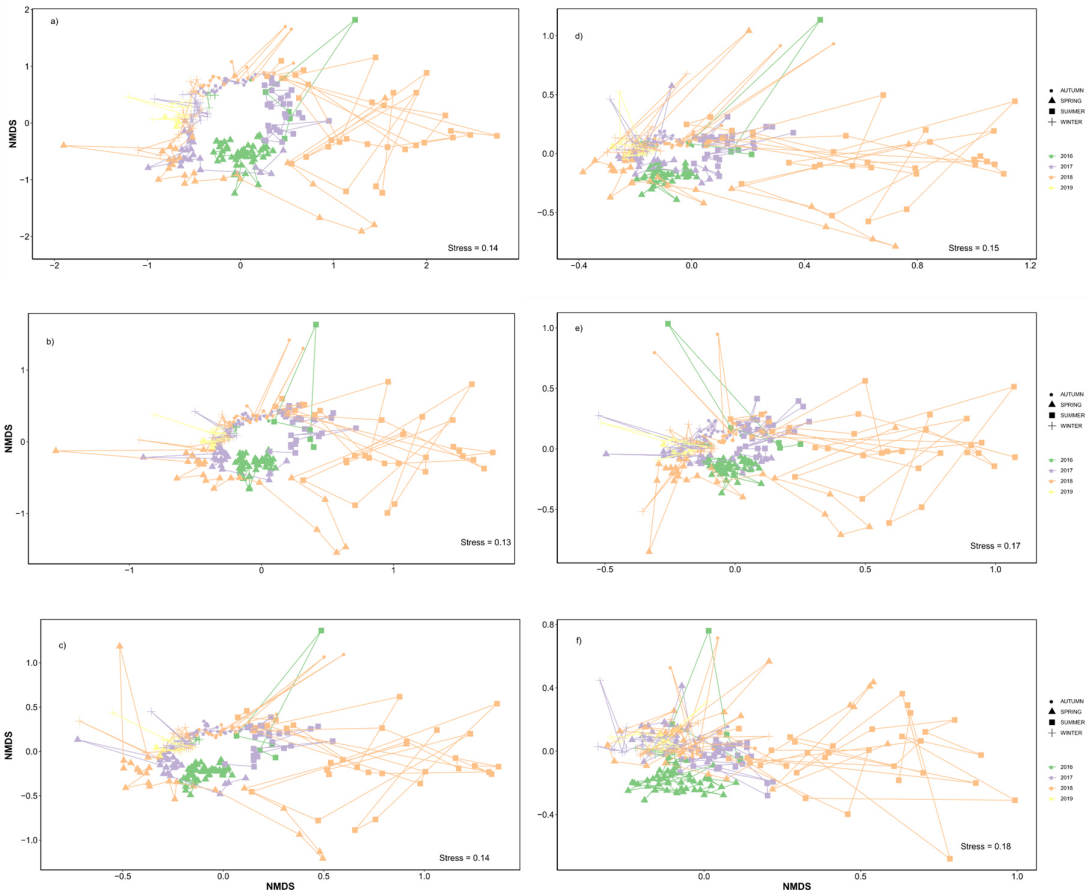
For Dinoflagellata 13 genera were found by all methods (Figure 1, Table 1). Of the genera that were found all both methods, two genera were based on synonyms (*Tripos* and *Ceratium*) or spelled differently (*Phalacroma*/*Phalachroma*) in the different datasets. One additional genus, that was present in the counts, was found in the dataset of the lower threshold. In total, 30 taxa were only found by metabarcoding, whereas 4 taxa were only found by counts. Out of five genera found by metabarcoding, one genus was not detected in counts and all genera have been observed occasionally since 2009 or more recently (details on all genera in Suppl. Table 2). Five other genera were only recorded until 2004. Two genera that are part of the counts were not found by metabarcoding. Two other genera could not be identified by PR2 but BLAST alignment identified potential OTUs in the MB datasets. 20 genera were not reported at the LTER before.

Other taxa

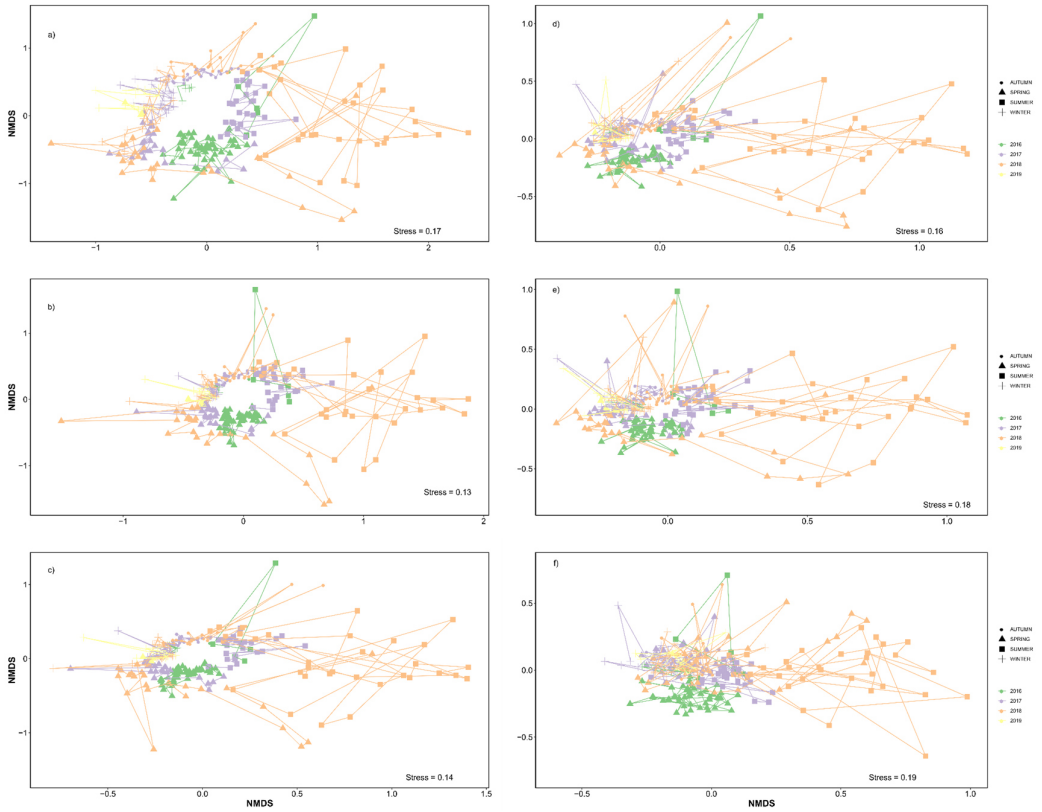
9 further genera are counted at the phytoplankton LTER. They include genera respectively species of Chlorophyta, Ochrophyta, Cercozoa, Haptophyta and Ciliophora. Additionally,

different size groups of indetermined taxa such as coccolithophorids and flagellates are counted. Most genera counted at the LTER were found in the metabarcoding dataset at a threshold of 50 reads. For example the genus *Scenedesmus* was not present in the dataset with a higher threshold. As several *Chattonella* sp. are now accepted as *Fibrocapsa* or *Pseudo-chattonella* sp., PR2 only identified these genera. *Mesodinium* and *Myrionecta* could not be found by PR2. BLAST alignment also did not suggest any Ciliophora OTU fitting to these genera, however PR2 and BLAST identification were weak for Ciliophora in general. Out of 368 Ciliophora OTUs at the 50 reads threshold, PR2 could not identify at least 188 OTUs on genus level. After BLAST alignment of the unclassified OTUs, over 42% of these OTUs could still not be identified clearly.

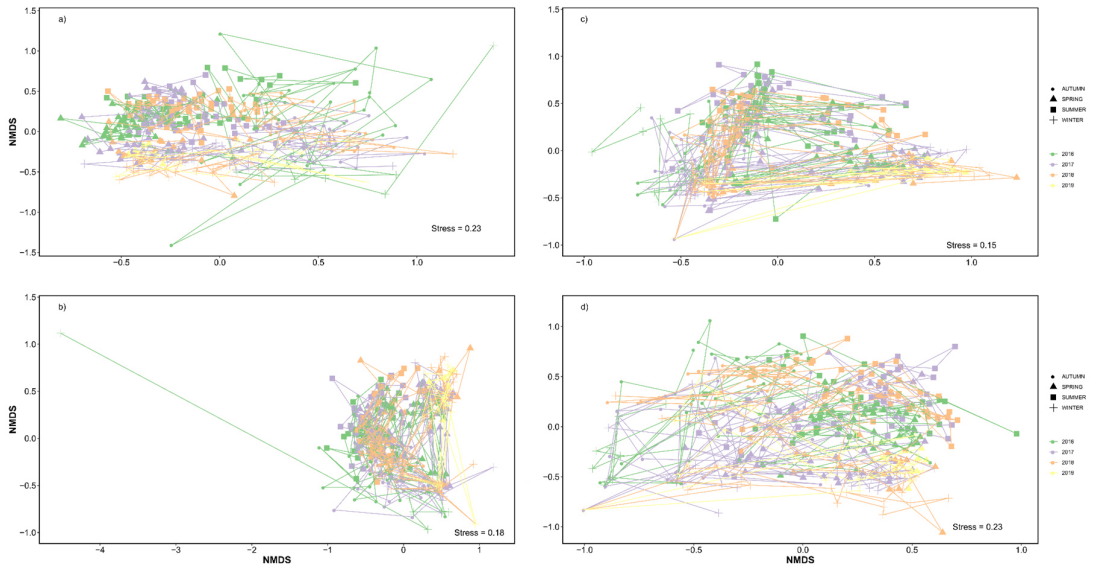
Supplementary Figures



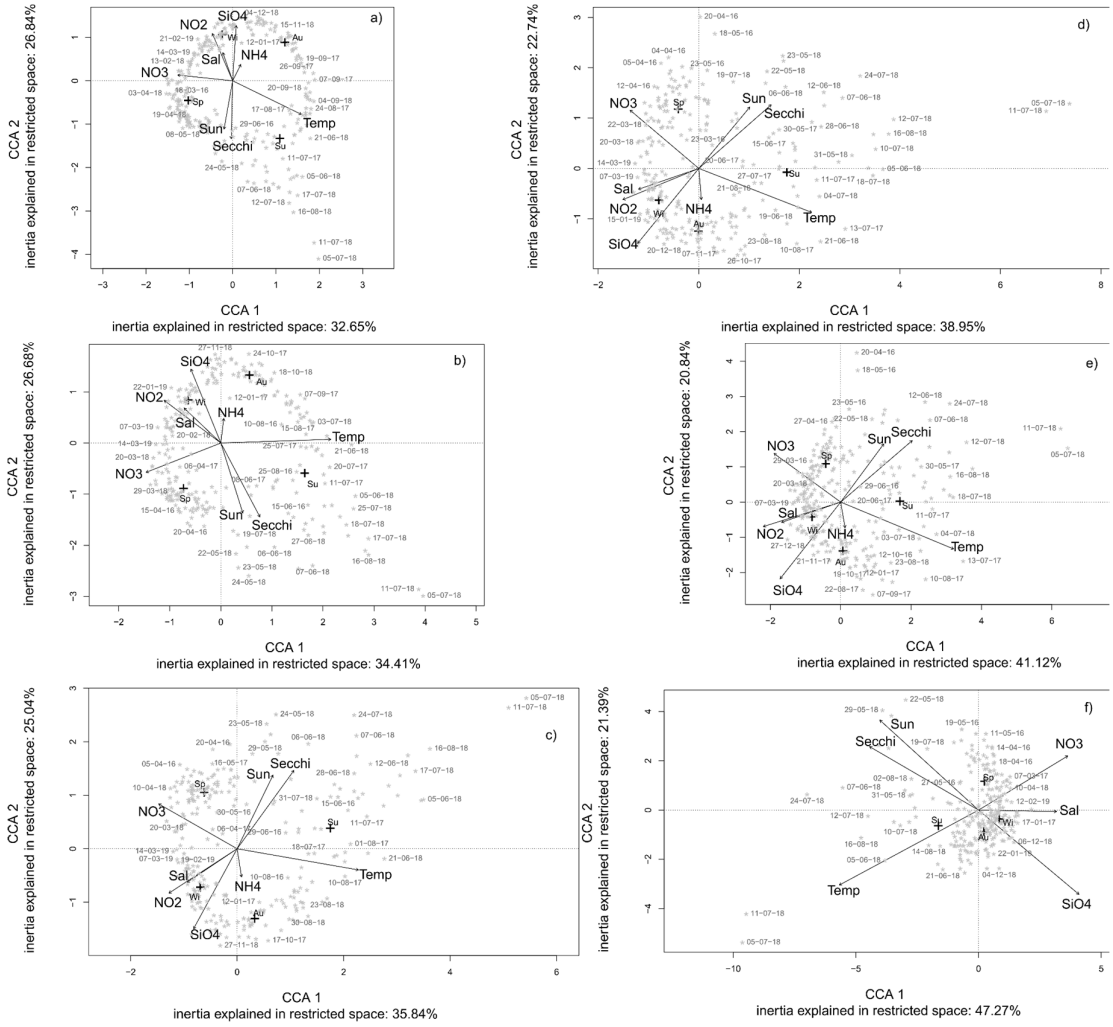
Suppl. Figure 1: NMDS plots of the metabarcoding dataset (0.001%; 205 reads), a) OTU level, b) genus level, c) family level, d) order level, e) class level and f) phylum level, seasons are displayed by different shapes, years are color coded.



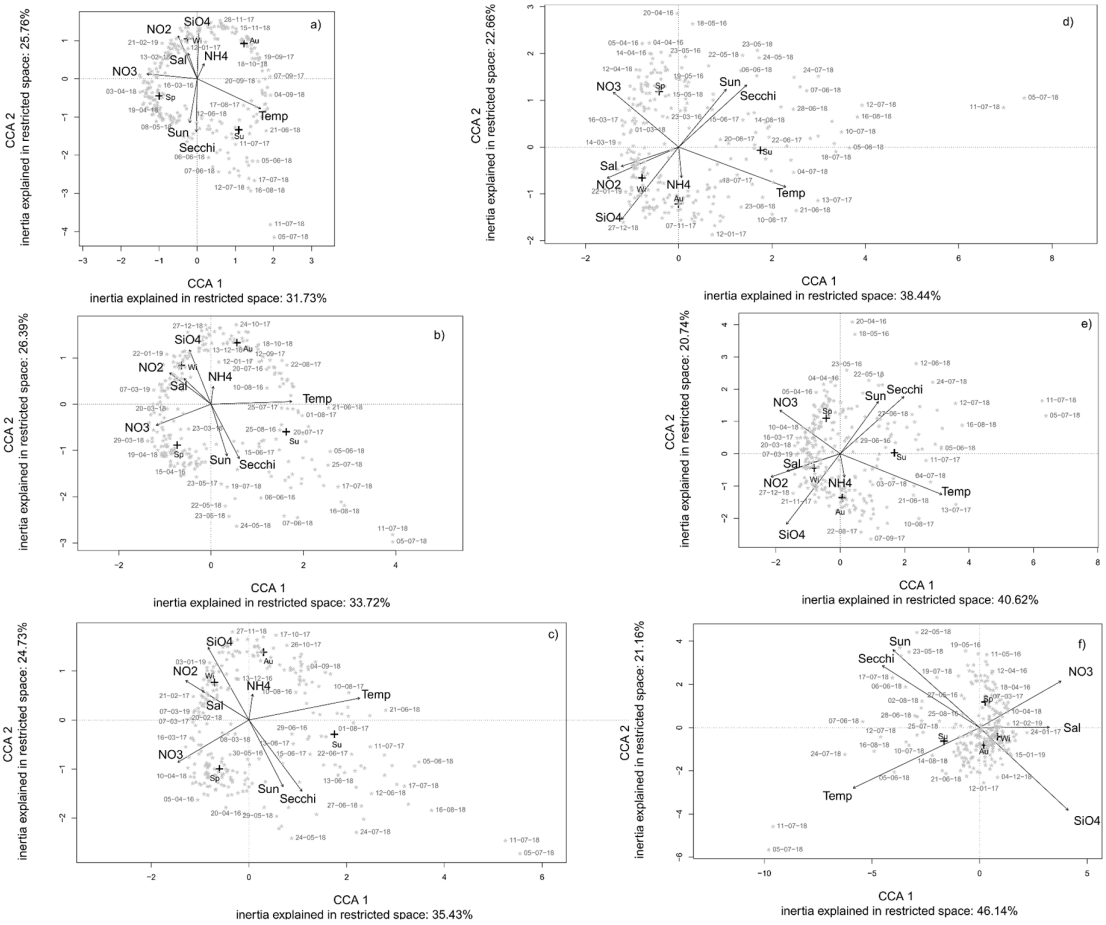
Suppl. Figure 2: NMDS plots of the metabarcoding dataset (50 reads), a) OTU level, b) genus level, c) family level, d) order level, e) class level and f) phylum level, seasons are displayed by different shapes, years are color coded.



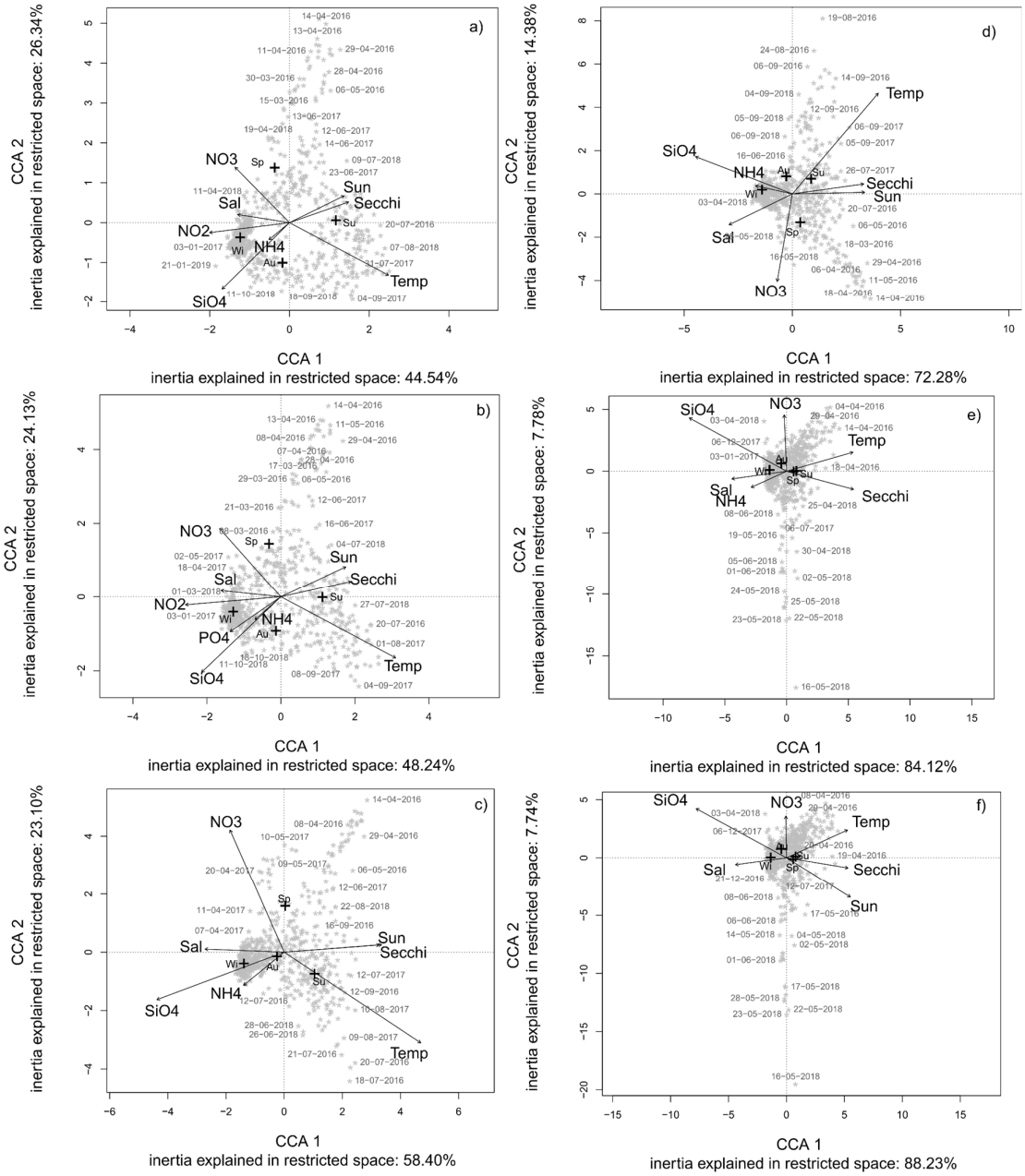
Suppl. Figure 3: NMDS plots of the zooplankton LTER dataset, a) last assigned taxa, b) family level, c) order level, d) class level, seasons are displayed by different shapes, years are color coded.



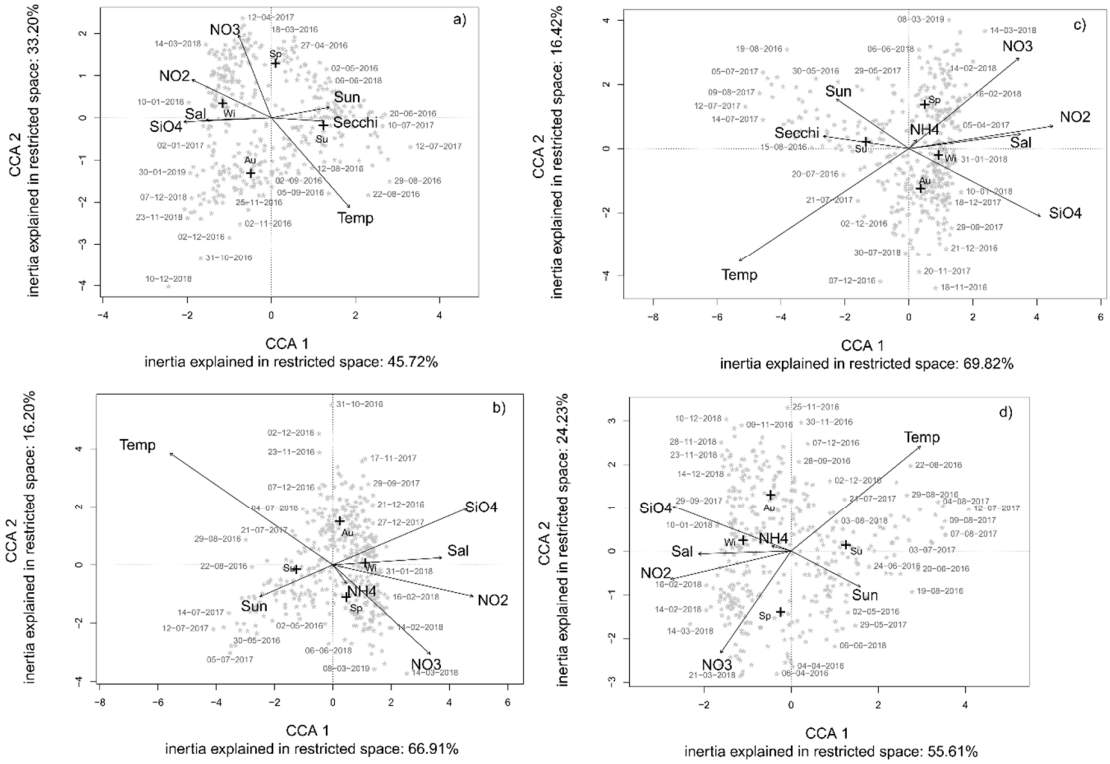
Suppl. Figure 4: CCA plots of the metabarcoding dataset (0.001%; 205 reads), a) OTU level, b) genus level, c) family level, d) order level, e) class level and f) phylum level, samples are plotted as grey asterisks with partially displayed sampling dates, including significant parameters in black (if applicable): seasons (Sp = spring, Su = summer, Au = autumn, Wi = winter), Temperature (Temp), salinity (Sal), Secchi depth (Secchi), sunshine duration (Sun), silicate (SiO4), nitrate (NO3), nitrite (NO2), ammonium (NH4), phosphate (PO4); additional information on total inertia (full space) and inertia in restricted space can be found in Suppl. Table 1.



Suppl. Figure 5: CCA plots of the metabarcoding dataset (50 reads), a) OTU level, b) genus level, c) family level, d) order level, e) class level and f) phylum level, samples are plotted as grey asterisks with partially displayed sampling dates, including significant parameters in black (if applicable): seasons (Sp = spring, Su = summer, Au = autumn, Wi = winter), Temperature (Temp), salinity (Sal), Secchi depth (Secchi), sunshine duration (Sun), silicate (SiO4), nitrate (NO3), nitrite (NO2), ammonium (NH4), phosphate (PO4); additional information on total inertia (full space) and inertia in restricted space can be found in Suppl. Table 1.



Suppl. Figure 6: CCA plots of the phytoplankton LTER dataset, a) last assigned taxa, b) genus level, c) family level, d) order level, e) class level and f) phylum level, samples are plotted as grey asterisks with partially displayed sampling dates, including significant parameters in black (if applicable): seasons (Sp = spring, Su = summer, Au = autumn, Wi = winter), Temperature (Temp), salinity (Sal), Secchi depth (Secchi), sunshine duration (Sun), silicate (SiO4), nitrate (NO3), nitrite (NO2), ammonium (NH4), phosphate (PO4); additional information on total inertia (full space) and inertia in restricted space can be found in Suppl. Table 1.



Suppl. Figure 7: CCA plots of the zooplankton LTER dataset a) last assigned taxa, b) family level, c) order level, d) class level, samples are plotted as grey asterisks with partially displayed sampling dates, including significant parameters in black (if applicable): seasons (Sp = spring, Su = summer, Au = autumn, Wi = winter), Temperature (Temp), salinity (Sal), Secchi depth (Secchi), sunshine duration (Sun), silicate (SiO4), nitrate (NO3), nitrite (NO2), ammonium (NH4), phosphate (PO4); additional information on total inertia (full space) and inertia in restricted space can be found in Suppl. Table 1.

Supplementary Tables

Suppl. Table 1: CCA statistic for each CCA model, variables are listed in order, in which they were added to the model, MB = Metabarcoding, P-LTER = "Phytoplankton LTER", Z-LTER = "Zooplankton LTER" #S/#T = Number of Samples/ Taxa, Last a.t. = Last assigned taxa, Temp = temperature, Sal = salinity, Sun = sunshine duration, Sec = Secchi depth, SiO4 = silicate, NO3 = nitrate, NO2 = nitrite, NH4 = ammonium, PO4 = phosphate, *not significant in additional ANOVA for added terms

Dataset	Taxonomic Level (#S/#T)	significant variables in model	Inertia explained	
			in full space [%] (CCA1/CCA2)	in restricted space [%] (CCA1/CCA2)
MB (50 reads)	OTU (273/6108)	Season, Temp, SiO4, Sal, NO3, Sec, NO2, Sun, NH4	17.25 (5.47/4.44)	57.49 (31.73/25.76)
	Genus (273/985)	Season, Temp, Sal, SiO4, NO3, Sun, Sec, NO2, NH4	18.89 (6.37/4.99)	60.11 (33.72/26.39)
	Family (273/518)	Season, Temp, SiO4, Sal, NO3, Sun, NO2, Sec, NH4	21.97 (7.46/5.21)	60.16 (35.43/24.73)
	Order (273/259)	Season, SiO4, Sal, Sun, Temp, NO3, Sec, NO2, NH4	20.69 (7.95/4.69)	61.10 (38.44/22.66)
	Class (273/138)	Season, SiO4, Sal, Sun, Temp, NO2, NO3, Sec, NH4	20.26 (8.23/4.20)	61.35 (40.62/20.74)
	Phylum (273/36)	Season, SiO4, Sun, Sal, NO3, Sec, Temp	21.54 (9.94/4.56)	67.30 (46.14/21.16)
MB (0.001%; 205 reads)	OTU (273/2790)	Season, Temp, Sal, SiO4, NO3, NO2, Sun, Sec, NH4	19.76 (6.45/5.30)	59.50 (32.65/26.84)
	Genus (273/734)	Season, Temp, Sal. SiO4, NO3, Sec, Sun, NO2, NH4	20.13 (6.93/5.37)	61.09 (34.41/26.68)
	Family (273/420)	Season, Temp, Sal, SiO4, NO3, Sun, NO2, Sec, NH4	21.87 (7.84/5.48)	60.88 (35.84/25.04)
	Order (273/212)	Season, SiO4, Sal, Sun, Temp, NO3, NO2, Sec, NH4	21.31 (8.30/4.85)	61.69 (38.95/22.74)
	Class (273/115)	Season, SiO4, Sal, Sun, Temp, NO2, NO3, Sec, NH4	20.44 (8.40/4.26)	61.95 (41.12/20.84)
	Phylum (273/30)	Season, SiO4, Sun, Sal, NO3, Sec, Temp	21.52 (10.17/4.60)	68.66 (47.27/21.39)

Suppl. Table 1: CCA statistic for each CCA model (continued).

Dataset	Taxonomic Level (#S/#T)	significant variables in model	Inertia explained	
			in full space [%] (CCA1/CCA2)	in restricted space [%] (CCA1/CCA2)
P-LTER	Last a.t. (713/173)	Season, Temp, SiO4, NO3, Sal, Sec, NH4, NO2, Sun	23.53 (10.48/6.20)	70.89 (44.54/26.34)
	Genus (713/68)	Season, Temp, SiO4, NO3, Sal, Sec, NO2, NH4, Sun, PO4	30.87 (14.89/7.45)	72.37 (48.24/24.13)
	Family (713/25)	Season, Temp, NO3, SiO4, Sal, Sec, NH4, Sun	33.10 (19.34/7.65)	81.50 (58.40/23.10)
	Order (713/16)	Season, Temp, SiO4, NO3, Sal, Sun, NH4, Sec	33.69 (24.35/4.84)	86.66 (72.28/14.38)
	Class (713/11)	Season, Temp, SiO4, NO3, Sal, Sec, NH4	32.98 (27.75/2.57)	91.90 (84.12/7.78)
	Phylum (713/7)	Season, Temp, SiO4, NO3, Sal, Sec, Sun	36.58 (32.27/2.83)	95.97 (88.23/7.74)
Z-LTER	Last a.t. (407/43)	Season, Temp, NO2, Sal, NO3, SiO4, Sec, Sun	24.40 (11.16/8.10)	78.92 (45.72/33.20)
	Family (407/13)	Season, Temp, SiO4, Sal, NH4, Sun, NO3, NO2	25.76 (17.24/4.17)	83.11 (66.91/16.20)
	Order (407/8)	Season, Temp, SiO4, Sun, NH4, Sal, NO3, Sec, NO2	32.64 (22.78/5.36)	86.24 (69.82/16.42)
	Class (407/17)	Season, Temp, SiO4, Sal, NH4, NO2, Sun, NO3	29.91 (16.64/7.25)	79.85 (55.61/24.23)
	Phylum (407/2)	NOT APPLICABLE		

Suppl. Table 2: List of Bacillariophyta genera identified or known at Helgoland with different methods (MB50 = metabarcoding with 50 reads threshold, MB205 = metabarcoding with 205 reads threshold, PC = counted in "Phytoplankton LTER"), if a genus was not counted, but has been reported before, the respective source was added in the comments (Kraberg *et al*. 2019 for reports from 2009 onwards, Hoppenrath 2004 only for reports until 2004).

Genus	MB50	MB205	PC	COMMENT
Achnanthes	YES		YES	
Actinocyclus	YES	YES	YES	
Actinoptychus	YES	YES	YES	
Amphora	YES	YES		not reported
Asterionellopsis	YES	YES	YES	
Asteroplanus	YES	YES		not detected in counts, Kraberg *et al*. 2019
Attheya			YES	PR2 identified 1 OTU below 50 reads
Bacillaria	YES	YES	YES	
Bacteriastrum			YES	not identified by PR2 in MB50/MB205, BLAST alignment identified 1 OTU above the MB50/MB205 threshold (OTU 777)
Berkeleya	YES			not reported
Biddulphia	YES	YES	YES	
Cerataulina	YES	YES	YES	
Chaetoceros	YES	YES	YES	
Corethron	YES	YES	YES	
Coscinodiscus	YES	YES	YES	
Cyclotella	YES	YES		not reported
Cylindrotheca	YES	YES	YES	
Cymbella	YES	YES		not reported
Dactyliosolen			YES	PR2 identified 1 OTU below 50 reads
Delphineis	YES	YES		Kraberg *et al*. 2019
Detonula			YES	not identified by PR2 in MB50/MB205, BLAST alignment identified 1 OTU above the MB50/MB205 threshold (OTU 97)
Ditylum	YES	YES	YES	
Eucampia	YES	YES	YES	
Eunotogramma	YES	YES		only reported in Hoppenrath 2004
Grammonema	YES			not reported
Guinardia	YES	YES	YES	
Lauderia	YES	YES	YES	
Leptocylindrus	YES	YES	YES	
Licmophora	YES			not reported
Lithodesmium	YES	YES	YES	
Lyrella	YES			not reported

Suppl. Table 2: List of Bacillariophyta genera identified or known at Helgoland (continued).

Genus	MB50	MB205	PC	COMMENT
Mediopyxis			YES	not identified by PR2 in MB50/MB205, BLAST alignment identified 1 OTU above the MB50 threshold (OTU 4449)
Melosira	YES		YES	
Meuniera	YES	YES	YES	
Minidiscus	YES	YES		only reported in Hoppenrath 2004
Minutocellus	YES	YES		not reported
Navicula	YES	YES		not detected in counts, Kraberg *et al*. 2019
Nitzschia	YES	YES		only reported in Hoppenrath 2004
Odontella	YES	YES	YES	
Paralia	YES	YES	YES	
Phaeodactylum	YES	YES		not reported
Plagiogrammopsis			YES	not identified by PR2 in MB50/MB205, BLAST alignment identified 1 OTU above the MB50/MB205 threshold (OTU 1864)
Plagiolemma			YES	not identified by PR2 in MB50/MB205, BLAST alignment identified 3 OTUs above the MB50/MB205 threshold (OTU 285, OTU 631, OTU 897)
Pleurosigma	YES	YES	YES	
Podosira			YES	PR2 identified 1 OTU below 50 reads
Porosira	YES	YES	YES	
Proboscia			YES	not identified by PR2 in MB50/MB205, BLAST alignment identified 1 OTU above the MB50 threshold (OTU 4901)
Psammodictyon	YES			not reported
Pseudo-nitzschia	YES	YES	YES	
Rhizosolenia	YES	YES	YES	
Roperia			YES	not found by metabarcoding (PR2/BLAST)
Skeletonema	YES	YES	YES	
Stellarima	YES			not detected in counts, Kraberg *et al*. 2019
Stephanopyxis	YES	YES	YES	
Tabularia	YES	YES		not reported
Tenuicylindrus	YES	YES		not reported
Thalassionema	YES	YES	YES	
Thalassiosira	YES	YES	YES	
Triceratium			YES	not found by metabarcoding (PR2/BLAST)

Suppl. Table 3: List of Dinoflagellata genera identified or known at Helgoland with different methods (MB50 = metabarcoding with 50 reads threshold, MB205 = metabarcoding with 205 reads threshold, PC = counted in "Phytoplankton LTER"), if a genus was not counted, but has been reported before, the respective source was added in the comments (Kraberg *et al*. 2019 for reports from 2009 onwards, Hoppenrath 2004 only for reports until 2004).

Genus	MB50	MB205	PC	COMMENT
Akashiwo	YES	YES	YES	
Alexandrium	YES	YES		Kraberg *et al*. 2019
Amphidinium			YES	not identified by PR2 in MB50/MB205, BLAST alignment identified OTUs above the MB50/MB205 threshold (OTU 581, 3476) & PR2 identified OTUs below 50 reads
Amphidoma	YES	YES		not reported
Amylax	YES	YES		Kraberg *et al*. 2019
Ansanella	YES	YES		not reported
Archaeperidinium	YES	YES		not reported
Azadinium	YES	YES		not reported
Balechina	YES			not reported
Crypthecodinium	YES	YES		not reported
Dinophysis	YES	YES	YES	
Diplopsalis	YES		YES	
Euduboscquella	YES	YES		not reported
Fragilidium	YES	YES		only reported in Hoppenrath 2004
Gonyaulax	YES	YES		not detected in counts, Kraberg *et al*. 2019
Gymnodinium	YES	YES	YES	
Gyrodinium	YES	YES	YES	
Hematodinium	YES	YES		not reported
Herdmania	YES			not reported
Heterocapsa	YES	YES	YES	
Islandinium	YES	YES		not reported
Karenia	YES	YES		only reported in Hoppenrath 2004
Karlodinium	YES	YES		not reported
Katodinium			YES	not identified by PR2 in MB50/MB205, BLAST alignment identified potential OTUs above the MB50/MB205 threshold (OTU 1215, 2792, 5329, 5641, 5693)
Kofoidinium	YES	YES		not reported
Lepidodinium	YES	YES		Kraberg *et al.* 2019
Levanderina	YES			not reported
Margalefidinium	YES			not reported
Noctiluca	YES	YES	YES	
Paragymnodinium	YES			not reported

Suppl. Table 3: List of Dinoflagellata genera identified or known at Helgoland (continued).

Genus	MB50	MB205	PC	COMMENT
Pelagodinium	YES	YES		not reported
Phalachroma / Phalacroma	YES	YES	YES	*Phalachroma* in PR2, *Phalacroma* in LTER
Polykrikos	YES	YES	YES	
Prorocentrum	YES	YES	YES	
Proterythropsis	YES	YES		only reported in Hoppenrath 2004
Protoceratium	YES	YES		Kraberg *et al*. 2019
Protoperidinium	YES	YES	YES	
Pseudo-phalacroma	YES			not reported
Pyrocystis			YES	not found by metabarcoding (PR2/BLAST)
Pyrophacus			YES	not found by metabarcoding (PR2/BLAST)
Scrippsiella	YES	YES	YES	
Spatulodinium	YES	YES		only reported in Hoppenrath 2004
Symbiodinium	YES			not reported
Syndinium	YES	YES		not reported
Torodinium	YES	YES	YES	
Tripos/Ceratium	YES	YES	YES	Tripos in PR2, Ceratium in LTER
Warnowia	YES	YES		only reported in Hoppenrath 2004
Woloszynskia	YES	YES		not reported

Suppl. Table 4: List of Crustacea genera identified or known at Helgoland with different methods (MB50 = metabarcoding with 50 reads threshold, MB205 = metabarcoding with 205 reads threshold, ZC = counted in "Zooplankton LTER"), if a genus is not counted but has been reported before the respective source was added in the comments.

Genus	MB50	MB205	ZC	COMMENT
Acartia	YES	YES	YES	
Ameira	YES			not reported
Anthessius	YES	YES		not reported
Balanus	YES	YES		Harms 1993
Calanus	YES	YES	YES	
Centropages	YES	YES	YES	
Crangon	YES	YES		Harms 1993, Greve *et al*. 2004
Cyclopina	YES			Greve *et al*. 2004
Diarthrodes	YES			Harms 1993
Euterpina	YES	YES		Greve *et al*. 2004
Galathea	YES			Harms 1993, Greve *et al*. 2004
Harpacticus	YES	YES		Harms 1993
Isias	YES	YES		not reported
Itunella	YES	YES		not reported
Lichomolgus	YES	YES		not reported
Oithona	YES	YES	YES	
Paracalanus	YES	YES	YES *	*counted as a genus-complex together with *Pseudocalanus*
Penilia	YES	YES		Greve *et al*. 2004
Pseudanthessius	YES	YES		not reported
Pseudocalanus	YES	YES	YES *	*counted as a genus-complex together with *Paracalanus*
Tachidius	YES	YES		not reported
Temora	YES	YES	YES	
Thoralus	YES			Harms 1993
Tisbe	YES	YES		Harms 1993, Greve *et al*. 2004
Verruca	YES	YES		Harms 1993
Zaus	YES	YES		Harms 1993

Suppl. Table 5: Number of genera found per phyla (excluding Bacillariophyta, Dinoflagellata and Crustacea) from March 2016 to March 2019, both metabarcoding datasets include all genera cut to 50 or 205 reads per OTU respectively, genera listed here include the ones clearly identified by PR2, strain IDs or other taxonomic groups do not count as clear identification and further BLAST identification is only mentioned in the text.

Phylum	Number of genera (50 reads)	Number of genera (0.001%,205 reads)
Apicomplexa	6	4
Apusomonadidae	1	1
Cercozoa	27	17
Chlorophyta	38	25
Choanoflagellida	5	3
Ciliophora	51	40
Conosa	2	0
Cryptophyta	6	6
Fungi	19	8
Haptophyta	9	9
Hilomonadea	2	2
Katablepharidophyta	2	2
Lobosa	9	5
Mantamonadidea	1	1
Mesomycetozoa	3	3
Metazoa	174	109
Ochrophyta	17	17
Perkinsea	2	1
Radiolaria	2	1
Rhodophyta	1	0
Stramenopiles_X	20	15
Streptophyta	15	9
Telonemia	1	0

Abstracts for Publications V and VI

Publication V: Spatial dynamics of eukaryotic microbial communities in the German Bight

Abstract

Monitoring changes in eukaryotic microbial communities is critical for understanding ecosystem dynamics, trophic interactions and the impacts of climate change. Long-term time series are an important tool for monitoring changes in ecological communities, but time series from a single location may not be representative of regional dynamics. In the German Bight, the Helgoland Roads time series is such a long-term series. Here, we consider the spatial dynamics of the eukaryotic microbes as an indicator of the representativeness of the Helgoland Roads site for the coastal German Bight, which is located in the North Sea. The eukaryotic microbial community in the German Bight was analysed at Helgoland Roads and two coastal stations (Cuxhaven and Wilhelmshaven) between March and October 2016 using metabarcoding. In addition, an oceanographical model was used to check for potential hydrological connectivity between the stations during the sampling period. Our results showed that the communities were different at the three stations. Helgoland was dominated by dinoflagellates, whereas the coastal stations had more diverse communities. Furthermore, differences were observed in the dinoflagellate and diatom communities between the three stations. Lagrangian particle tracking applied to the model results, showed limited connectivity between Helgoland and the coastal stations in 2016. The differences between Helgoland and the coastal stations were correlated with the different hydrological regimes and associated nutrient contents. Our observations suggest the presence of different eukaryotic microbial communities separated by complex hydrological conditions in the coastal German Bight.

Publication VI: Phytoplankton Responses to Marine Climate Change – An Introduction

Abstract

Phytoplankton are one of the key players in the ocean and contribute approximately 50% to global primary production. They serve as the basis for marine food webs, drive chemical composition of the global atmosphere and thereby climate. Seasonal environmental changes and nutrient availability naturally influence phytoplankton species composition. Since the industrial era, anthropogenic climatic influences have increased noticeably – also within the ocean. Our changing climate, however, affects the composition of phytoplankton species composition on a long-term basis and requires the organisms to adapt to this changing environment, influencing micronutrient bioavailability and other biogeochemical parameters. At the same time, phytoplankton themselves can influence the climate with their responses to environmental changes. Due to its key role, phytoplankton has been of interest in marine sciences for quite some time and there are several methodical approaches implemented in oceanographic sciences. There are ongoing attempts to improve predictions and to close gaps in the understanding of this sensitive ecological system and its responses.

B Acknowledgements

First of all, I would like to thank everyone who has accompanied and supported me, special thanks to my supervisors for their trust and support on my journey.

I cordially thank Prof. Dr. Karen Wiltshire for giving me the opportunity to work on this exciting topic. Thank you, Karen, for your support, patience and all the fruitful discussions throughout the years.

I thank Prof. Dr. Ruth Schmitz-Streit, PD Dr. Bernhard Fuchs, PD Dr. Cédric Meunier and Prof. Dr. Reid Brennan for their willingness to be part of my examining committee.

Furthermore, I would like to thank all members of my PhD committee at AWI: Prof. Dr. Karen Wiltshire, Prof. Dr. Maarten Boersma, Dr. Katja Metfies, Dr. Alexandra Kraberg and PD Dr. Bernhard Fuchs for your time and the interesting and constructive discussions not only during my PhD committee meetings. I want to thank all of you for your great supervision and endless support.

Special thanks to Kristine Carstens, Silvia Peters, Dr. Alexandra Kraberg, Christine Grauel, Dr. Cédric Meunier, Dr. Antje Wichels and Dr. Eva-Maria Brodte as well as the late Dr. Mirco Scharfe for taking me in and helping me with anything during my time on Helgoland. Special thanks also to everyone involved in my work during the time I spent in Bremerhaven, namely Dr. Katja Metfies, Swantje Rogge, Kerstin Oetjen, Stefan Neuhaus, Pim Sprong, Nora-Charlotte Pauli and Lora Strack van Schijndel.

I would like to thank all other PhD colleagues I met at AWI and elsewhere, especially everyone at Helgoland and my fellow AWI DokTeam members. Thank you to my fellow office mates Subrata Sarker, Prayoga Wahyu Triputra, Constantin Büttner, Lea Blum, Linda-Maria Wolf und Johanna Strauß for all the time spent together. Many thanks to my flatmates and friends, thank you for a wonderful time on Helgoland: Melanie, Elanor, Sebastian, Annika and Michaela.

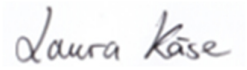
Last but not least, I want to express my deepest gratitude to my family and friends. This thesis would not have been possible without your unrestricted support and understanding.

B Acknowledgements

C Eidesstattliche Erklärung

Hiermit erkläre ich, Laura Käse, durch meine Unterschrift, dass ich die vorliegende Abhandlung, abgesehen von der Beratung meiner Betreuer/in, nach Inhalt und Form selbstständig angefertigt und alle Stellen, die ich wörtlich dem Sinne nach aus Veröffentlichungen entnommen habe, als solche kenntlich gemacht habe, mich auch keiner anderen als der angegebenen Literatur oder sonstiger Hilfsmittel bedient habe. Die Arbeit enthält drei Publikationen, die bereits veröffentlich worden sind. Mein Eigenanteil dieser einzelnen Publikationen ist in einer separaten Auflistung angegeben.

Weiterhin erkläre ich, dass die Arbeit zu keiner Zeit an einer anderen Stelle im Rahmen eines Prüfungsverfahrens vorgelegen hat und, dass es sich hierbei um das erste und bisher einzige Promotionsverfahren handelt. Die Arbeit wurde unter Einhaltung der Regeln guter wissenschaftlicher Praxis der Deutschen Forschungsgemeinschaft durchgeführt und mir wurde kein akademischer Grad entzogen.

Witzenhausen, 14.03.2022 

Ort, Datum Unterschrift