U.PORTO

MESTRADO EM CIÊNCIAS DO MAR - RECURSOS MARINHOS

ESPECIALIZAÇÃO EM BIOLOGIA E ECOLOGIA MARINHAS

Prokaryotic Dynamics Across the North Pacific Subtropical Front

Eva Silva Lopes

M

2020

U.PORTO

Eva Silva Lopes. Prokaryotic Dynamics Across the North Pacific Subtropical Front

Prokaryotic Dynamics Across the North Pacific Subtropical Front

Eva Silva Lopes

M.ICBAS **2020**

INSTITUTO DE CIÊNCIAS BIOMÉDICAS ABEL SALAZAR

Eva Silva Lopes

Prokaryotic Dynamics Across the North Pacific Subtropical Front

Dissertação de candidatura ao grau de Mestre em Ciências do Mar – Recursos Marinhos, especialização em Ecologia e Biologia Marinhas submetida ao Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto.

Orientador - Prof. Doutora Catarina Magalhães

Categoria – Investigadora Auxiliar; Professora Auxiliar

Afiliações – Faculdade de Ciências da Universidade do Porto; Centro de Investigação Interdisciplinar Marinha e Ambiental

Coorientador – Doutor Miguel Semedo

Categoria – Postdoc

Afiliações – Faculdade de Ciências da Universidade do Porto; Centro de Investigação Interdisciplinar Marinha e Ambiental

Acknowledgments

Firstly, I would like to thank my supervisor, Dra Catarina Magalhães, for the opportunity given and the confidence deposited in me for this research. Even though we were an ocean apart, Dra Catarina always had time for supporting and guiding me. Definitely, Dra Catarina is a worthy and the best supervisor that someone can have in the corner when doing things out of the comfort zone.

I also would like to thank to my co-advisor, Dr Miguel Semedo, for always having time for me and for his eternal patience for my endless questions and doubts. Dr Miguel always made sure that I fully understood what he was explaining, never diminished my doubts and he made me see how funny and interesting RStudio could be. By also being on my side in every step, Dr Miguel is one of the best people to have in your corner when doing things outside of the comfort zone.

I cannot forget to thank Professor Susana Moreira, for introducing me to my advisor, sharing her vast knowledge with me and just for listening to me when I needed it.

To Dra Maria Paola Tomasino and Dra Roberta Piredda, for sharing their wisdom with me and always answering my many questions with a smile on their face.

My colleagues and friends from CIIMAR were also fundamental pieces to this process, especially to Inês Carvalho and Marta Sousa. Thank you for always calming me down on my drama queen moments.

I also would like to acknowledge my friends, especially Andreia Silva and Rita Fernandes. Thank you for listening to me, even when I did not make any sense, advise me and just for being present when I needed the most.

Last but not least, I would like to thank my family for the incredible support, for giving me the opportunity of a better education and for being present in the little moments.

Declaração de Honra

Declaro que, o presente relatório final de mestrado, é da minha autoria e não foi utilizado previamente noutro curso ou unidade curricular desta, ou de outra instituição. As referências a outros autores (afirmações, ideias, pensamentos) respeitam escrupulosamente as regras de atribuição e, encontram-se devidamente indicadas no texto e nas referências bibliográficas, de acordo com as normas de referenciação. Tenho consciência que a prática de plágio e auto plágio constitui um ilícito académico.

Abstract

Microbial communities are crucial for ecosystem function, such as primary production and nutrient cycling. However, few studies have addressed the distribution of microplankton communities in the deep oligotrophic region of Pacific Ocean. Moreover, the biogeochemical and physical drivers of microbial community structure are not fully understood in these areas. This research aims to investigate the patterns of prokaryotic communities distribution in the North Pacific Subtropical Front (NPSF), an area with very limited biological surveys.

During the “Exploring fronts with multiple robots” campaign, that took place at 1000 nautical miles off the coast of southern California in June 2018, seawater samples were collected at five different depths (Surface [5m], Above DCM [80m], DCM (Deep Chlorophyll Maximum) [108-130m], Below DCM [175-200m], Mesopelagic [500m]). Samples were also collected across the ocean front, targeting the different water masses (Polar, Front, Subtropical), together with biogeochemical data. Prokaryotic community composition and diversity were studied by using16S rRNA metabarcoding approach and next-generation sequencing (NGS).

Results, regarding the analysis of depth effect, demonstrated a clear vertical stratification of the water column. The phylum Proteobacteria was present across all the water columns but some phyla demonstrated depth preference. The phylum Cyanobacteria was abundant at surface environments and, on the other hand, the phylum Thaurmachaeota started to appear at higher depths. In addition, a biodiversity hotspot was found around the 200 m depth. Concerning the water mass effect, results did not demonstrate a big influence of the water mass type on the distribution of the prokaryotic communities. A similar microbiome was observed in the three water masses. At both depths (Surface and DCM), the number of shared amplicon sequence variants (ASVs) between the water masses was relatively high, 411 at Surface, and 937 at DCM from a total of 6445 ASVs retrieved from all samples. Subsequently, the horizontal distribution of the prokaryotic communities did not demonstrate a visible habitat preference.

Ultimately, this study demonstrated a strong influence of depth on the microbial community structure and a weak influence of the horizontal gradients in the North Pacific Subtropical Front.

Keywords: Ocean Fronts; North Pacific Subtropical Front; Metabarcoding; Prokaryotes; Depth; Water Mass, Temperature; Salinity; DCM

Resumo

As comunidades microbianas são fundamentais para o funcionamento dos ecossistemas, como por exemplo na produção primária e ciclo dos nutrientes. Contudo, apenas alguns estudos analisaram a distribuição das comunidades microbianas na zona profunda oligotrófica do Oceano Pacifico. Os factores biogeoquímicos e físicos que influenciam a estrutura das comunidades microbianas também não estão devidamente analisados nesta área. Esta tese tem como objetivo analisar os padrões de distribuição das comunidades procariotas na Frente Subtropical do Pacifico Norte (FSPN), uma área com pesquisas biológicas bastante limitadas.

Durante a campanha " Exploring fronts with multiple robots", que ocorreu a 1000 milhas náuticas da costa do sul da Califórnia em junho de 2018, amostras de água do mar foram recolhidas em cinco profundidades diferentes (Superfície [5m], Acima do DCM [80m], DCM (Máximo de clorofila profunda) [108-130m], Abaixo do DCM [175-200m], Mesopelágico [500m]). Amostras também foram recolhidas ao longo da frente oceânica, diferenciando as diferentes massas de água (Polar, Frontal, Subtropical), juntamente com dados biogeoquímicos. A composição e diversidade da comunidade procariótica foram estudadas por uma abordagem de metabarcoding do gene 16S rRNA usando sequenciamento de próxima geração (NGS).

Os resultados, relativos à análise do efeito da profundidade, demonstraram uma estratificação vertical evidente da coluna de água. O filo Proteobacteria estava presente em toda a coluna de água, mas alguns filos demonstraram preferência por profundidade. O filo Cyanobacteria revelou ser abundante em ambientes superficiais e, por outro lado, o filo Thaurmachaeota começou a aparecer em maiores profundidades. Além disso, um hotspot de biodiversidade foi encontrado em torno dos 200 m de profundidade. Em relação ao efeito da massa de água, os resultados não demonstraram uma grande influência do tipo de massa de água na distribuição das comunidades procarióticas. Microbiomas semelhantes foram observados nas três massas de água. Em ambas as profundidades (Superfície e DCM), o número de ASVs partilhados entre as massas de água foi relativamente alto, 411 na Superfície e 937 no DCM de um total de 6445 ASVs recuperados de todas as amostras. Consequentemente, a distribuição horizontal das comunidades procarióticas não demonstrou uma preferência de habitat visível.

Esta pesquisa demonstrou uma forte influência da profundidade na estrutura da comunidade microbiana dentro da profundidade de 500m e uma fraca influência dos gradientes horizontais na NPSF.

Palavras-chave: Frentes oceânicas; Frente Subtropical do Pacífico Norte; Metabarcoding; Procariontes; Profundidade; Massa de água, Temperatura; Salinidade; DCM

Table of Contents

Acknowledgments i
Abstract iii
Resumo iv
Table of Contents vi
List of Abbreviations viii
List of Figures x
List of Tables xii
1. Introduction 1
1.1. Ocean Fronts 1
1.2. Importance in terms of Ocean Productivity and Fisheries 2
1.3. Biogeochemical and Microplankton Dynamics in Ocean Fronts 3
1.4. North Pacific Subtropical Front 4
1.5. Description of the Schmidt Ocean Institute Campaign 5
1.6. Genomic methodologies to study microplankton communities 6
2. Objectives 7
3. Methodologies 8
3.1. Study Area – Sampling Sites 8
3.2. In Situ Measurements (CTD – Casts) 9
3.3. Water Sampling for Prokaryotic Communities Analysis 13
3.4. 16S rRNA Amplicon Analysis 15
4. Results 17
Part A – Depth Effect 17
4.1. Biogeochemical Gradients across Depth 17
4.1.1. Environmental Parameters Across Depth 17
4.2. Prokaryotic Communities Distribution Across Depth 21
4.2.2. Venn Diagram Analysis 25
4.2.3. Prokaryotic Taxonomic Analysis 27
4.3. Environmental vs Biological Relationships 32
Part B – Water Mass Effect 35
4.4. Biological and Biogeochemical Gradients Across the Water Masses 35
4.4.1. Environmental Gradients Across the Water Masses 35
4.5. Prokaryotic Distribution Across the Water Masses 39
4.5.1. Alpha and Beta Diversity 41
4.5.2. Venn Diagram Analysis 46

4.5.3. Prokaryotic Taxonomic Analysis 48
4.6. Environmental vs Biological Relationships 59
4.6.1. Shared ASVs 59
4.6.2. Exclusive ASVs 61
5. Discussion 63
5.1. Prokaryotic communities distribution within water column depth 63
5.1.1. Environmental Controls on Depth Prokaryotic Communities Distribution 64
5.1.2. Below DCM as a Hotspot of Biodiversity 66
5.1.3. Connectivity between the different water column depths 67
5.2. Prokaryotic communities distribution across the different water masses at Surface and DCM 68
5.2.1. Prokaryotic diversity across the different water masses 70
5. Conclusions and Future Perspectives 72
6. Scientific Communications (based on the work of this thesis) 73
7. References 74
8. Supplementary Information 98

List of Abbreviations

α diversity – Alpha Diversity

°C – degrees Celsius

16S rDNA – 16S ribosomal deoxyribonucleic acid

ASV – Amplicon Sequence Variant

C – Carbon

CCA – Canonical Correspondence Analysis

CO_2 – Carbon Dioxide

CTD – Conductivity, Temperature, Depth

DCM – Deep Chlorophyll Maxim

DOM – Dissolved Organic Matter

eDNA – environmental DNA

Fig. – Figure

Log – Logarithm

m – metre

mg/m^3 – milligrams per meter cubed

ml/L – millilitre per litre

NA – Not Available

NPSF – North Pacific Subtropical Front

NTU – Nephelometric Turbidity Unit

OTU – Operational Taxonomic Unit

PCA – Principal Component Analysis

PCR – Polymerase Chain Reaction

PSU – Practical Salinity Unit

SOI – Schmidt Ocean Institute

Temp – Temperature

TZCF – Transition Zone Chlorophyll Front

List of Figures

Figure 1: The major 10 types of ocean fronts (based on Belkin & Cornillon, 2007).....2
Figure 2: Map of the area of research and location of the ocean stations (adapted from Google Earth). 9
Figure 3: Picture of Rosette, one of the gears used to collect on board water samples. ...13
Figure 4: Sterivex filters used on the sample filtration procedure. 14
Figure 5: Oxygen (a) and fluorescence (b) water column profiles in transect A samples. .18
Figure 6: Principal Component Analysis of the samples to study the depth effect. 21
Figure 7: Alpha Diversity of the samples 1A-20A. 23
Figure 8: Hierarchical cluster of log transformation of samples 1A-20A.....24
Figure 9: Multi-dimension scaling of the samples of samples 1A-20A. 24
Figure 10: Venn diagram of the samples 1A-20A. 26
Figure 11: Exclusive ASVs (%) for each depth classification. 27
Figure 12: Phyla identification across the water column. 27
Figure 13: Class identification across the water column. 28
Figure 14: Top 5 phyla identification along the water column. 29
Figure 15: Proteobacteria genera across the water column. 30
Figure 16: Cyanobacteria genera across the water column. 30
Figure 17: Thaumarchaeota genera across the water column. 31
Figure 18: Marinimicrobia (SAR406 clade) genera across the water column. 32
Figure 19: Euryarchaeota genera across the water column. 32
Figure 20: Canonical Correspondence Analysis of the samples 1A-20A. 33
Figure 21: Spearman correlation analysis. 35
Figure 22: Water masses classification and Temperature-Salinity relationship (Surface-a; DCM-b). 36
Figure 23: Rarefaction curve of samples collected at Surface (a) and DCM (b). 41
Figure 24: Alpha Diversity of the samples Water Mass-Surface. 42
Figure 25: Alpha Diversity of the samples Water Mass-DCM. 43
Figure 26: Hierarchical clusters of log transformation with (a) and without (b) the sample 53A of Water Mass-Surface. 44
Figure 27: Hierarchical clusters of log transformation with (a) and without (b) the sample 56A of Water Mass-DCM. 44
Figure 28: Venn diagram for Surface water masses. 46
Figure 29: Venn diagram for DCM Water Mass. 47
Figure 30: Exclusive ASVs distribution according to salinity (a) and temperature (b) per sample at Surface. 48

Figure 31: Exclusive ASVs distribution according to salinity (a) and temperature (b) per sample at DCM. 48
Figure 32: Phyla identification across the samples at Surface. 49
Figure 33: Phyla identification across the samples at DCM. 50
Figure 34: Class identification across the samples at Surface. 51
Figure 35: Class identification across the samples at DCM. 52
Figure 36: Shared genera between the water masses at Surface samples. 54
Figure 37: Shared genera between the water masses at DCM samples. 55
Figure 38: Exclusive genera between the water masses at Surface samples. 57
Figure 39: Exclusive genera between the water masses at DCM samples. 58
Figure 40: CCA analysis of the shared ASVs at Surface samples. 59
Figure 41: CCA analysis of the shared ASVs at DCM samples. 60
Figure 42: CCA analysis of the exclusive ASVs at Surface samples. 61
Figure 43: CCA analysis of the exclusive ASVs at DCM samples. 62

List of Tables

Table 1: Metadata of the samples analysed for the depth effect.....10
Table 2: Metadata of the samples analysed for the water mass effect at Surface. 11
Table 3: Metadata of the samples analysed for the water mass effect at DCM. 12
Table 4: PCA correlations of the environmental data and samples analysed for depth effect.....19
Table 5: Environmental data with chlorophyll depth classifications.....20
Table 6: Cleaning Process of the 16S rRNA sequences for each sample (*: the ASVs classified as Eukaryota/Chloroplast/Mitochondria were also removed). 22
Table 7: Output of permanova analysis of the samples 1A-20A. 25
Table 8: Envfit test results for the depth effect samples. 33
Table 9: Environmental data of the samples collected at Surface. 37
Table 10: Environmental data of the samples collected at DCM.....38
Table 11: Cleaning Process of the ASVs present at Surface (*: the ASVs with Eukaryota/Chloroplast/Mitochondria were also removed from the ASV table). 39
Table 12: Cleaning Process of the ASVs present at DCM (*: the ASVs with Eukaryota/Chloroplast/Mitochondria were also removed from the ASV table). 40
Table 13: Tukey test results of alpha diversity and the water masses at Surface. 42
Table 14: Tukey test results of alpha diversity and the water masses at DCM. 43
Table 15: Output of permanova analysis of Water mass-Surface.....45
Table 16: Output of permanova analysis of Water mass-DCM. 45
Table 17: Envfit test results of the shared ASVs at Surface. 60
Table 18: Envfit test results of the shared ASVs at DCM.....60
Table 19: Envfit test results of the exclusive ASVs at Surface.....61
Table 20: Envfit test results of the exclusive ASVs at DCM. 62

1. Introduction

1.1. Ocean Fronts

An ocean front can be defined as a strait area of intensified horizontal gradients of water physicochemical properties, such as temperature, salinity, and nutrients, which divides extensive areas according to their water masses or according to the water column stratification, i.e. their vertical structure (Belkin, 2002). The length and duration of an ocean front are variable. Fronts can go from a few meters to many hundreds of kilometers, in terms of width. Additionally, the vertical extension can also be very big, as there are records for major fronts to be extended as two kilometers deep (Belkin & Cornillon, 2007; Belkin et al., 2009). Regarding the duration, some fronts can only occur for a few days, being designated as short-lived fronts, and some fronts are quasi-stationary and seasonally persistent. These types of fronts appear and disappear in the same area during the same season. However, the most salient fronts are present year-round (Belkin & Cornillon, 2007; Belkin et al., 2009). The physicochemical characteristics of ocean fronts are not constant either, especially salinity and temperature. Typically, salinity and temperature can vary 1 ppt (parts per thousand) and 5°C between water masses. Yet, the differences between salinity and temperature in strong fronts are characterized by higher amplitudes, going from 2 to 3 ppt and 10°C to 15°C (Belkin & Cornillon, 2007). Ocean fronts are also very dynamic zones, continuously changing through space and time. Changes of adjacent water masses within a front are made by frontal mixing, vertical transport, and air-sea interaction (Hopkins et al., 2010).

Ocean fronts can be divided and classified into ten types (fig. 1) (Belkin & Cornillon, 2007). In this study, samples were collected in a subtropical front. These oceanic fronts are created due to the Eckman wind convergence, which connects waters with different temperatures and maintains the emerging front (Belkin & Cornillon, 2007). This type of fronts can be found in the North Atlantic, especially in the Sargasso Sea, and in the North Pacific (Belkin & Cornillon, 2007). Normally, different types of fronts interact with each other, for instance, temperature fronts are commonly associated with salinity fronts. Thermohaline fronts are frequently accompanied by density fronts, as seawater density is based on temperature, salinity, and pressure (depth) (Belkin et al., 2009).

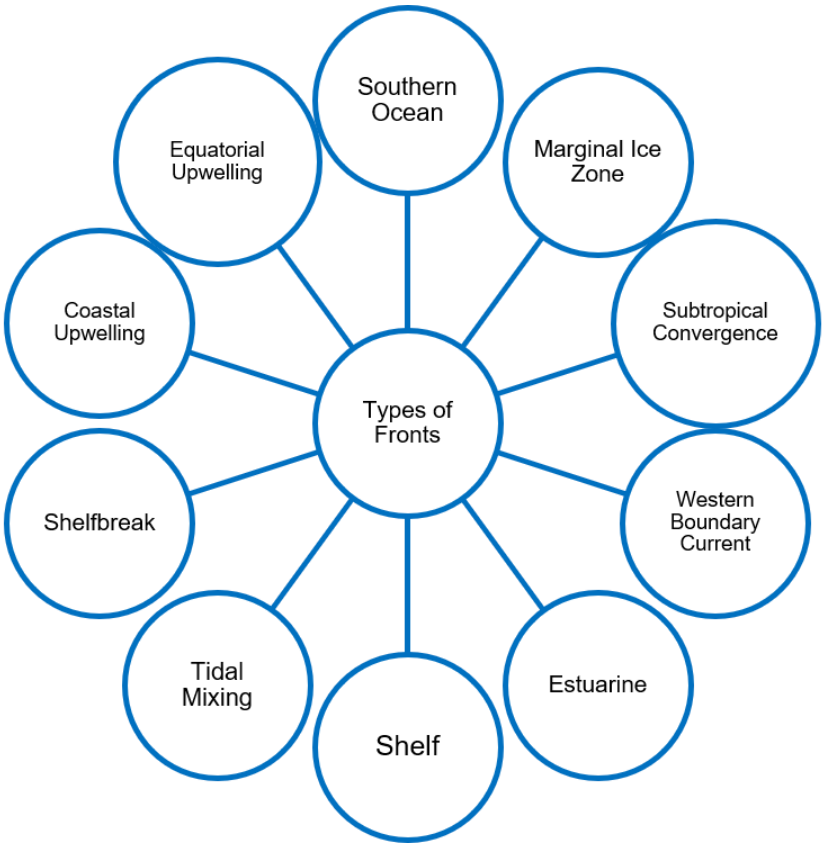


Figure 1: The major 10 types of ocean fronts (based on Belkin & Cornillon, 2007).

Ocean fronts can have crucial importance for society, such as in marine transportation (optimization of ships routes by taking advantage of front currents), fishing industry (several important fishery grounds are located in fronts), monitorization and prediction of climate change (based on the location of ocean currents/fronts), marine mining (patterns of ocean currents affects the offshore platform's activities), pollution control (by knowing fronts currents patterns and its variability) and coastal zone management by the fact that natural hazards are dependent on the large scale ocean circulation/frontal systems. In addition, knowledge about ocean fronts characteristics can also facilitate submarine navigation (Belkin & Cornillon, 2007).

1.2. Importance in terms of Ocean Productivity and Fisheries

Ocean fronts are fundamental for the biological productivity of marine ecosystems and can influence the distribution of all oceanic live forms, from microbes to top predators (Kahru et al., 2012; Mann & Lazier, 2013; Belkin et al., 2014). Fronts are highly associated with sites of high plankton productivity, biodiversity, and are crucial for several marine organisms that spawn, feed, and/or utilise these areas as nursery grounds (Smith & Lane, 1999; Ohman et

al., 2012). This elevated productivity is explained by the presence of a strong horizontal pressure gradient, which is formed in the fronts, that is balanced by the Coriolis force, resulting in strong along-fronts currents (Acha et al., 2015). When these forces are not synchronized, vertical circulations are formed. With these new circulations, the conditions for upwelling occurring are present and nutrient-rich water can reach new areas of the ocean. Consequently, phytoplankton growth will be stimulated (Sournia, 1994). Furthermore, frontal instabilities can result in vertical restratification, decreasing the turbulent flux of phytoplankton out of the euphotic zone, leading to a bigger light exposition (Taylor & Ferrari, 2011). As it follows, fronts have ideal growth conditions to phytoplankton in comparison to surrounding waters (Traganza et al., 1987; Taylor & Ferrari, 2011). As a result, the phytoplankton presence attracts nekton organisms, such as squid and fish, transferring the energy to higher trophic levels (Acha et al., 2015). The presence of several marine organisms at oceanic fronts is very well registered. Such marine organism are turtles, pelagic fishes (e.g.: tuna and swordfish), whales, and marine birds (Podestá et al., 1993: Polovina et al., 2000; Zainuddin et al., 2008; Boost et al., 2009; Okkonen et al., 2011). The presence of these animals at oceanic fronts can also be explained by the feeding behaviour and migration routes (Olson et al., 1994). Sagarminaga & Arrizabalaga (2014) suggested that the association between tunas and fronts is due to confinement to a physiologically optimum temperature range, the use of frontal gradients to occur thermoregulation, as well as the presence of aliment. These three factors are also frequently associated with cold upwelled waters.

1.3. Biogeochemical and Microplankton Dynamics in Ocean Fronts

Biogeochemistry in marine environments is influenced by the metabolic activities of the macroplanktonic communities. In the marine environments, the availability of nitrogen controls plankton productivity as well as the exportation of carbon (Dugdale & Goering, 1967; Eppley & Peterson, 1979). Plankton productivity in oligotrophic areas is determined by several factors, such as the fixation of nitrogen (Shiozaki et al., 2009). Nitrogen fixation serves as a valuable supply of nitrogen to the oligotrophic areas and consequently, balances the community production and carbon export (Capone et al., 2005). In the North Pacific, the availability of nitrogen through nitrogen fixation is comparable to the concentration of nitrogen by upwelling and/or advection (Dore et al., 2002; Johnson et al., 2010)

As stated above, ocean fronts are distinguished by the elevated primary production rate and by the presence of strong plankton communities (Fiala et al., 1994; Samo et al., 2012). As a result, the quantity of nutrients in these transition areas suffers a boost, which leads to higher concentrations of chlorophyll *a* (Riemann et al., 2011), since phytoplankton find conditions for optimal growth. As a consequence, the neighbouring waters also suffer an increase in biomass and organic matter (Samo et al., 2012). The result of primary production is turned into dissolved organic matter (DOM) by several mechanisms regulated by marine prokaryotes (Hagström et al., 1984; Azam & Malfatti, 2007). How the prokaryotes use the DOM, determines the relative portioning of carbon into labile/semi-labelled, and refractory organic matter classes. Ultimately, this process has an influence on the global ocean-atmosphere carbon dioxide (CO_2) exchange and carbon export (del Giorgio & Cole, 1998). Ultimately, these processes affect the exportation of carbon via the biological pump running through the microbial loop (Martin et al., 1994; Karl et al., 1998). At fronts, nutrients injection generally supports elevated bacterial production because of the increased phytoplankton biomass (Videau et al., 1994; Morán et al., 2001). Additionally, the increased transference of nutrients and carbon leads to the increment of the biogeochemical fluxes to deeper areas (Woodsona et al., 2015). As a consequence, these interactions between hydrodynamics and biological responses are capable of influencing the composition and diversity of the prokaryotic communities of these regions (Sempére et al., 2003).

1.4. North Pacific Subtropical Front

In the north Pacific Ocean, two major domains can be distinguished: the North Pacific Subarctic Front (NPSAF) and the North Pacific Subtropical Front (NPSF) (Roden, 1972; Chen et al., 2019). This thesis focused on the North Pacific Subtropical Front, and this front is located between 28ºN and 32ºN (Nakamura et al., 1997; Nakamura & Kazmin, 2003; Zhang et al., 2019).

The formation of NPSF is associated with net heat flux and turbulent entrainment as well as surface wind stress advection (Qiu et al., 2014; Wang et al., 2016). As a result of the surface wind stress, the Ekman transport creates a surface layer convergence at the southern edge of the front. A proxy is present to indicate regions of high surface water convergence. This proxy can be defined as the Transition Zone Chlorophyll Front (TZCF), where chlorophyll reaches values of 0.2-0.3 mg/m^3, and normally corresponds to the area of the highest chlorophyll values. As a result, TZCF reveals the zone of surface convergence, where cold vertically mixed waters chlorophyll-rich on the north side sinks beneath warm stratified waters chlorophyll poor on the south side (Polovina et al., 2001;

Pichel et al., 2007). The sea surface temperature gradient in this front is generally low, variating from 0.7˚C/100km to 0.9˚C/100km (Qiu et al., 2014). Additionally, the North Pacific Subtropical Front demonstrates a strong seasonal variation, and its locations change latitudinally around 10˚ over the year (Ayers & Lozier, 2010). NPSF reaches its maximum strength in the Winter and Spring and is practically absent in the Summer and Autumn. The weakening of the front occurs mainly from May to August (Kobashi & Xie, 2011; Qiu & Kawamura, 2012). Additionally, the North Pacific Subtropical Front has a considerable influence on the atmospheric boundary layer, surface wind speed, and deep convection in the troposphere (Wang et al., 2016).

The North Pacific Subtropical Front is a symbol of intensified trophic transferences (Olson et al., 1994). The dense cold phytoplankton rich waters descend below warmer water, which leads to a phytoplankton convergence (Roden, 1980). The presence of phytoplankton attracts other organisms, like zooplankton, capable of maintaining their position in the downwelling. As a consequence, the presence of zooplankton attracts the presence of higher trophic level predators. Ultimately, a complete food web is put together in the NPSF (Olson et al., 1994). Besides being a biodiversity hotspot, NPSF also plays an important role in the migration for several species, such as loggerhead turtles (*Caretta caretta*), albacore tuna (*Thunnus alalunga*), laysan albatross (*Phoebastria immutabilis*), black-footed albatross (*Phoebastria nigripes*), as well as, elephant seals (*Mirounga angustirostris*) (Le Boeuf et al., 2000; Polovina et al., 2000; Polovina et al., 2001; Shaffer et al., 2005). Swordfish (*Xiphias gladius*) can also be found in these waters. The presence of these species makes this area a known local for fisheries and increases the economical relevance of the NPSF (Seki et al., 2002; Howell et al., 2008).

1.5. Description of the Schmidt Ocean Institute Campaign

This thesis is part of the project “Exploring fronts with multiple robots” promoted by the Schmidt Ocean Institute. This project took place at 1000 nautical miles off the coast of southern California, on board of the research vessel *Falkor* and it included an international team of researchers and partners (LSTS and CIIMAR - Porto University, Portugal; Technical University of Cartagena - Spain; Norwegian University of Science and Technology - Norway; Harvard, University of Rhode Island, and Lamond-Doherty Observatory - United States).

A common difficulty in working with ocean robotics are the delays and disruptions when messages are being sent over long distances underwater. Consequently, a goal of this

project was to develop more efficient means of communication in these dynamic environments. By extending *Falkor*'s communications systems with underwater and radio communications. Falkor acted as a "nerve centre" which allowed the mixed-initiative planning and execution control by researchers present in the planning and execution of the control loops. Resulting in real-time updates of coordination and communication software to be used for data assimilation and modelling (https://schmidtocean.org/cruise/exploring_fronts_with_multiple_aerial-surface-underwater-vehicles/). Another goal of this campaign was to use autonomous vehicles to find and sample the physical, chemical, and biological features of the ocean. With this, the campaign became a more dynamic campaign, by collecting samples at the moment. This is, with the autonomous vehicles characterization of the physical, chemical, and biological proprieties properties, the researchers knew on-time where to collect the better samples.

In this study the biological samples collected during this campaign were used to study the distribution of the prokaryotic communities across depth and horizontal gradients of the NPSF.

1.6. Genomic methodologies to study microplankton communities

A common approach to study the prokaryotic communities diversity in natural complex samples is the 16S rRNA metabarcoding approach. This approach appeared as a response to technology in order to fill the scientific needs. DNA metabarcoding analysis can be defined as the identification of high-throughput multispecies/higher-level taxon by using all the DNA, even the degraded part, from an environmental sample (Taberlet et al., 2012). Metagenomics is another possible methodology to analyse microplankton diversity. Metagenomics can be defined as the study and analysis of the complete genomes of a community, avoiding the need of an amplification step (Handelsman et al., 1998). Metagenomic analyses in the prokaryotic communities have been used in several different types of marine environments, such as open oceans, coastal zones, and estuaries at different latitudes, longitudes, and depths (see DeLong, 2005). This type of approach is capable of analysing the complete genetic information of prokaryotic communities. As a consequence, the relative abundance of all genes can be determined and used to characterize natural communities (Debroas et al., 2009; Oh et al., 2011). Metagenomic analyses are generally used to perform the genome analysis of uncharacterized taxa, identify differences in the metabolic profiles across environments, and identify specific genomic adaptations to life in constating habitats (Beja et al., 2000; Dinsdale et al., 2008; Eiler et al., 2014).

Metabarcoding and metagenomics can be used according to the main objectives of the study. While metagenomics is applied to the analysis of all genomes within a particular ecosystem, metabarcoding is applied to the analysis of a subset of a gene(s) (Pavan-Kumar et al., 2015). This approach is essentially used to describe species diversity in the ecosystems, being reported as a better analysis to identify the rare taxa within an ecosystem (Pavan-Kumar et al., 2015). Additionally, metabarcoding results are considered trustworthy, valid, taxonomically comprehensive and is a practical biodiversity measurement (Yang et al., 2014). According to Ji et al. (2013), Metabarcoding uses the universal polymerase chain reaction (PCR) in order to mass-amplify taxonomically informative genes, retrieved from environmental DNA (eDNA) (Ji et al. 2013). The result of the PCR is an amplicon, which is resolved by next generation sequencing (NGS), generating an enormous amount of sequences. Generally, this output has sequencing errors that are eliminated through its processing in specific NGS bioinformatic pipelines. These bioinformatic analyses are going to cluster the sequences into operational taxonomic units (OTUs) that are assigned to its respective taxonomy, using taxonomy databases (Ji et al., 2013).

2. Objectives

The main goal of this study is to investigate the patterns of horizontal and vertical distribution of prokaryotic communities across the North Pacific Subtropical Front (NPSF), to understand how the biogeochemical and physical parameters that characterized this remote region of the Pacific Ocean influence the distribution of the planktonic prokaryotic communities. The main goal of this thesis will be achieved by accomplish three specific objectives:

- Understand the physical and biogeochemical gradients present in the North Pacific Subtropical Front.
- Evaluate if the prokaryotic communities are influenced by water column depth in NPSF.
- Verify if prokaryotic communities distribution is influenced by the horizontal water masses stratification that characterizes NPSF.

3. Methodologies

3.1. Study Area – Sampling Sites

Under the project “Exploring fronts with multiple robots”, seawater samples were collected along the North Pacific Subtropical front, 1000 nautical miles off the coast of southern California (fig. 2), on board of the Schmidt Ocean Institute (SOI) research vessel *Falkor* (https://schmidtocean.org/cruise/exploring_fronts_with_multiple_aerial-surface-underwater-vehicles/).

The sample collection started on the 1st of June 2018 and it ended on the 14th of June 2018, within the region 29.69-30.76°N, 132.11-132.08°W. A total of thirty-three samples were analysed in this thesis. The collected samples were divided into two different groups to study two different effects that may influence the distribution of the prokaryotic community: Depth (Part A) and Water Mass (Part B).

The samples (n= 20) analysed for the depth effect were classified according to their depth and their chlorophyll values, obtained by the fluorescence sensor. As a result, five different depth classifications were used: Surface (5m) (n=5), Above DCM (80m) (n=1), DCM (108-130m) (n=5), Below DCM (175-200m) (n=4) and Mesopelagic (500m) (n=5). The deep chlorophyll maximum (DCM) samples were the ones with the highest values of fluorescence. The DCM depths registered were between 108 to 130 m, similar to the DCM depths previously registered by Letelier et al. (2004) and Sauzèd et al. (2020).

The samples (n= 23) analysed for the water mass effect were classified according to their salinity values. As a result, three different types of water masses were identified: Polar (n= 11), Front (n= 8) and Subtropical (n= 4). The salinity values of the Polar water samples varied from 34.07 to 34.29, the Front samples varied from 34.39 to 34.66, and the Subtropical water samples varied from 34.78 and 34.96. Additionally, these samples were subdivided into two different groups regarding their depth: Surface and DCM. The salinity values and water mass classification matched with Aksenov et al. (2010), Cao et al. (2020), and Woo & Pattiaratchi (2008). While at Surface the three water masses (Polar, Front, and Subtropical) were found, at DCM, only the water masses Polar and Front were present. This was due to the salinity values of DCM samples not being high enough to classify the samples as Subtropical.

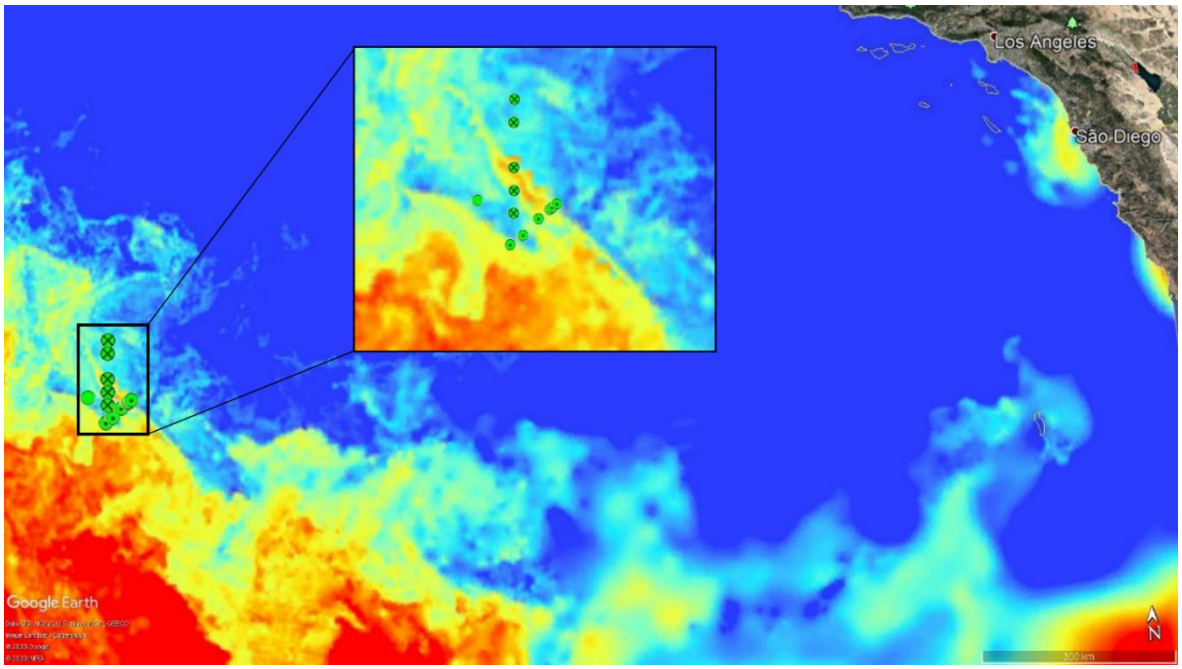


Figure 2: Map of the area of research and location of the ocean stations (adapted from Google Earth).

3.2. In Situ Measurements (CTD – Casts)

The collected samples were distributed at two different transects (A – 93.39 km – and B – 119.31 km). In the transect A, samples were analysed in terms of depth and horizontal water mass effect. Samples from transect B were exclusively analysed for the water mass effect (table 2 and 3). In all stations in situ measurements were performed using a SEABIRD SBE 9 PLUS CTD (conductivity-temperature-depth) profiler. In total, twelve Casts-CTD were deployed, five for the depth effect (table 1) and twelve for the water mass effect (table 2 and 3). In addition to conductivity (mS/cm), temperature (°C), and depth (m), the Casts-CTD sensors measured other environmental parameters, such as salinity (PSU), oxygen (ml/L), turbidity (NTU), and fluorescence (mg/m3).

Table 1: Metadata of the samples analysed for the depth effect.

Chlorophyll Depth	Sample ID	Cast	CTD	Transect	Date	Time	Latitude	Longitude	Gear
Surface (5m)	4A	1	1	A	1/Jun	06:17:18	30.586	-132.08606	Rosette
	8A	2	2		1/Jun	14:00	29.9179	-132.0818	
	12A	3	3		2/Jun	05:20	30.084	-132.08	
	16A	4	6		2/Jun	09:25	30.255	-132.083	
	20A	5	8		2/Jun	15:08	30.7575	-132.081	
Above DCM (80m)	19A	5	8	A	2/Jun	15:08	30.7575	-132.081	Rosette
DCM (108-130m)	3A	1	1	A	1/Jun	06:17:18	30.586	-132.08606	Rosette
	7A	2	2		1/Jun	14:00	29.9179	-132.0818	
	11A	3	3		2/Jun	05:20	30.084	-132.08	
	15A	4	6		2/Jun	09:25	30.255	-132.083	
	18A	5	8		2/Jun	15:08	30.7575	-132.081	
Below DCM (175-200m)	2A	1	1	A	1/Jun	06:17:18	30.586	-132.08606	Rosette
	6A	2	2		1/Jun	14:00	29.9179	-132.0818	
	10A	3	3		2/Jun	05:20	30.084	-132.08	
	14A	4	6		2/Jun	09:25	30.255	-132.083	
Mesopelagic (500m)	1A	1	1	A	1/Jun	06:17:18	30.586	-132.08606	Rosette
	5A	2	2		1/Jun	14:00	29.9179	-132.0818	
	9A	3	3		2/Jun	05:20	30.084	-132.08	
	13A	4	6		2/Jun	09:25	30.255	-132.083	
	17A	5	8		2/Jun	15:08	30.7575	-132.081	

Table 2: Metadata of the samples analysed for the water mass effect at Surface.

Water Mass	*Sample ID*	*Cast*	*CTD*	*Transect*	*Date*	*Time*	*Latitude*	*Longitude*	*Gear*
Polar	4A	1	1	A	1/Jun	06:17:18	30.586	-132.08606	Rosette
	20A	5	8	A	2/Jun	15:08	30.7575	-132.081	Rosette
	24A	9	NA	NA	6/Jun	09:17	30.0104	-132.4	Surf from wet lab
	57A	35	TowYo3	B	12/Jun	00:28	29.96118	-131.74621	Surf from wet lab
	59A	36	TowYo1	B	12/Jun	01:50	29.98813	-131.70628	Surf from wet lab
Front	8A	2	2	A	1/Jun	14:00	29.9179	-132.0818	Rosette
	55A	34	TowYo4	B	11/Jun	23:29	29.94787	-131.76732	Surf from wet lab
	73A	43	35	B	14/Jun	04:11	29.75496	-131.97031	Surf from wet lab
Subtropical	12A	3	3	A	2/Jun	05:20	30.084	-132.08	Rosette
	16A	4	6	A	2/Jun	09:25	30.255	-132.083	Rosette
	53A	33	TowYo11	B	11/Jun	22:06	29.87848	-131.8647	Surf from wet lab
	71A	42	34	B	14/Jun	02:45	29.68509	-132.11068	Surf from wet lab

NA – Not Available

Table 3: Metadata of the samples analysed for the water mass effect at DCM.

Water Mass	*Sample ID*	*Cast*	*CTD*	*Transect*	*Date*	*Time*	*Latitude*	*Longitude*	*Gear*
Polar	7A	2	2	A	1/Jun	14:00	29.9179	-132.082	Rosette
	15A	4	6	A	2/Jun	09:25	30.255	-132.083	Rosette
	18A	5	8	A	2/Jun	15:08	30.7575	-132.081	Rosette
	54A	34	TowYo4	B	11/Jun	23:29	29.94787	-131.767	Small winch
	56A	35	TowYo3	B	12/Jun	00:28	29.96118	-131.746	Small winch
	58A	36	TowYo1	B	12/Jun	01:50	29.98813	-131.706	Small winch
Front	3A	1	1	A	1/Jun	06:17:18	30.586	-132.086	Rosette
	11A	3	3	A	2/Jun	05:20	30.084	-132.08	Rosette
	52A	33	TowYo11	B	11/Jun	22:06	29.87848	-131.865	Small winch
	70A	42	34	B	14/Jun	02:45	29.68509	-132.111	Small winch
	72A	43	35	B	14/Jun	04:11	29.75496	-131.97	Small winch

NA – Not Available

In this project, different types of gears and techniques were used to collect the samples. The samples from the transect A were collected by only one type of gear, a Rosette system with 24*12L niskin bottles (fig. 3). This gear contains a carousel holding several bottles, which can be activated from the ship by projecting an electric pulse through the conducting CTD support cable (Thomson & Emery, 2014). In order to collect the samples on the transect B, two different gears and techniques were used: a small winch to collect water at DCM depth and surface water was directly collected from the wet lab itself. Winches used on this type of project are electrically or hydraulically powered devices, that their torque is suitable for the howling and traction speed, using a multistage gearbox. The winches' drums ensure that the cables with measuring sensors can be managed (Carral et al., 2017). The technique of "Surface from the wet lab" was direct water input to the research vessel laboratory, through a tap that connected the water at the surface to a wet lab.

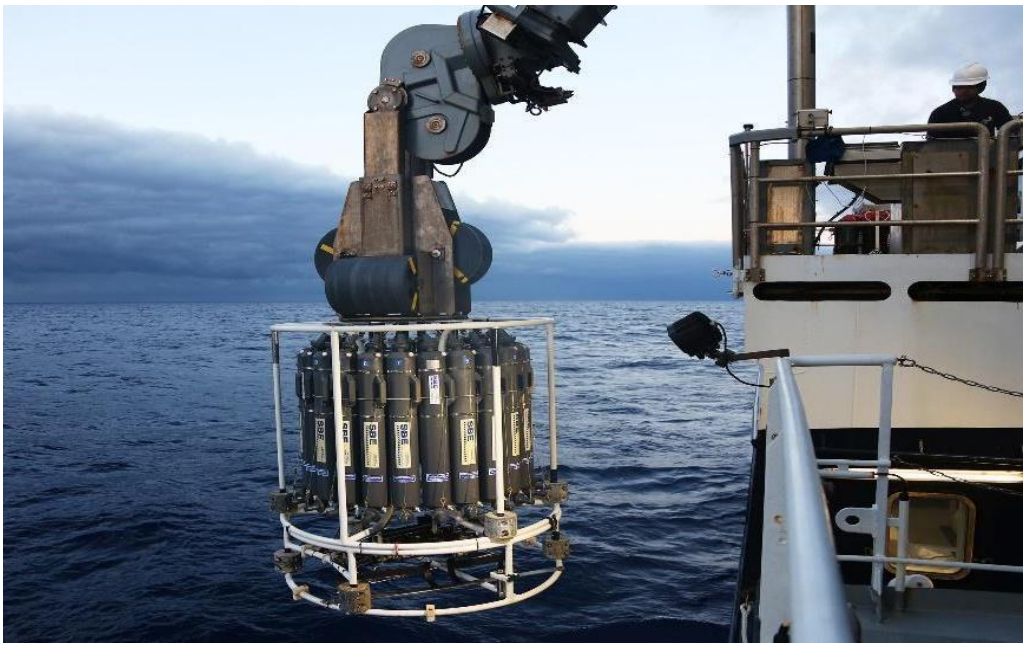


Figure 3: Picture of Rosette, one of the gears used to collect on board water samples.

3.3. Water Sampling for Prokaryotic Communities Analysis

The sampling methodology used was based on the standardized approach developed during the international Ocean Sampling Day campaign, with the goal of performing a global sampling campaign of ocean microbiomes (Kopf et al., 2015). Onboard of the Falkor research vessel, CIIMAR researchers (Dra Catarina Magalhães and Dra Maria Paola Tomasino) collected plankton samples by concentrating 3.75L of water column and then filtered it through a Sterivex® Filter with a 0.22µm pore size (fig. 4). The collected filters

were stored on board at -80°C until being processed at CIIMAR for eDNA isolation and sequencing analysis.

Total DNA was isolated directly from the Sterivex filters using the DNeasy PowerWater Sterivex Kit (QIAGEN), following the manufacturer's instruction.

The 16S rRNA gene was amplified by PCR with the primer pair 515YF (5′ - GTGYCAGCMGCCGCGGTAA - 3′) and Y906R-jed (5′ - CCGYCAATTYMTTTRAGTTT- 3′), which was designed by Caporaso et al. (2011, 2012) and latter modified by Apprill et al. (2015) and Parada et al. (2016). The initial PCR reaction included 12.5ng of DNA template in a total volume of 25µL. Firstly, the PCR protocol involved a 3min denaturation step, being followed by 25 cycles of 98˚C for 20s, 60˚C for 30s, 72˚C for 30s, and, finally, an extension at 72˚C for 5min. Negative controls were included in all PCR amplification procedures. Lastly, PCR products were one-step purified and normalized using SequalPrep 96-well plate kit (ThermoFisher Scientific, Waltham, USA) (Comeau et al., 2017), pooled and paired-end sequenced in the Illumina MiSeq® sequencer with the V3 chemistry, according to manufacturer instructions (Illumina, San Diego, CA, USA) at Genoinseq (Cantanhede, Portugal). This primer set targets the V4-V5 hypervariable regions of the 16S rRNA gene. The amplified 16S rRNA gene was used for sequencing through the Illumina MiSeq platform using 2*300 bp, V3 chemistry (Illumina). These steps were performed by BIOCANT laboratories (Cantanhede, Portugal). A detailed description of these methods can be found in Ribeiro et al. (2018) and Sousa et al. (2019).

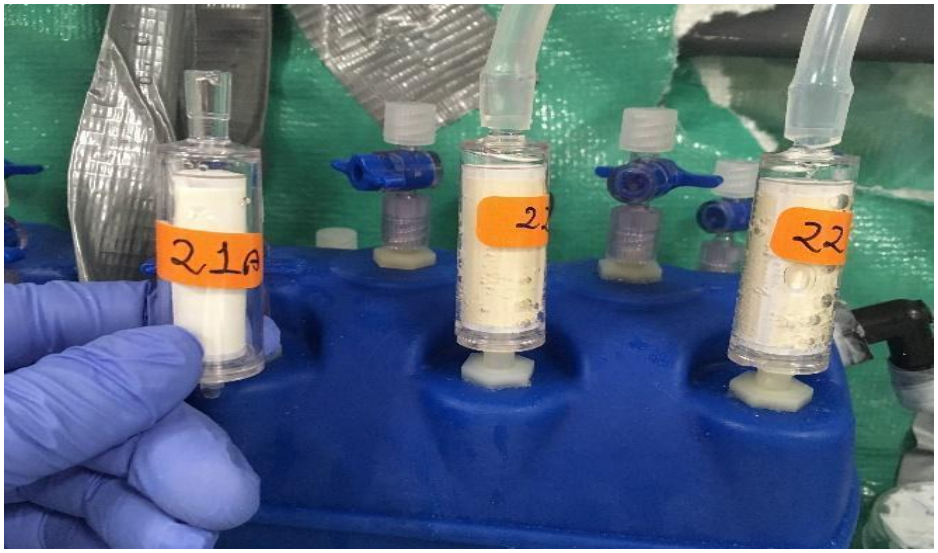


Figure 4: Sterivex filters used on the sample filtration procedure.

3.4. 16S rRNA Amplicon Analysis

Initially, the sequences were quality filtered, dereplicated and denoised. Afterward, the primer cut forward and reverse reads of the dataset were merged from raw Illumina fastq files (DADA2 v. 0.99.8) (Callahan et al., 2016). As a result, the first ASV table was created. Furthermore, chimeras were determined and removed with UCHIME (Edgar et al., 2011). In addition, the exclusive reads were assigned against the SILVA database (v. 1.2.8) (Quast et al., 2013) and the taxonomy was assigned. Lastly, undesirable lineages, such as “Chloroplast”, “Eukaryota” and “Mitochondria” were eliminated from the dataset. Subsequently, the final ASV table was generated.

Several approaches to analyse the distribution and diversity of the prokaryotic community across the Pacific Ocean samples were performed. The alpha diversity was analysed by using the number of observed ASVs and the Shannon and Chao-1 indexes. The richness (Observed ASVs) is the number of ASVs present in each sample. The Shannon index accounts for both abundance and evenness of the species (Shannon, 1948). The Chao-1 index estimated the diversity based on abundance, it uses the number of sequences belonging to a certain ASV in each sample (Chao, 1984).

The alpha diversity analysis was obtained with phyloseq R package (v. 1.30.0; McMurdie & Holmes, 2013). With the aim to analyse the dissimilarity between the samples (beta diversity), and the correlation between the samples and the environmental gradients, a principal component analysis (PCA) and a canonical correspondence analysis (CCA), were performed. The beta diversity results were obtained using the vegan R package (v. 2.5-6; Oksanen et al., 2007) and the multidimensional plot with the phyloseq R package. A rarefaction curve was also produced to visualize if sequencing depths were enough to represent the sample prokaryotic diversity. A permanova analysis was also performed to verify if the different effects (water column depth and front water masses) influence or not the dissimilarity between samples. These results were also obtained with the vegan R package. In addition, a Venn diagram was generated to calculate the shared and exclusive ASVs between the different water column depths and water masses. These analyses were performed in the vegan R package and with the Venn R package (v. 1.9.; Dusa, 2020). Taxonomy bar plots were also generated. Finally, heatmaps and Spearman correlations were also included in the downstream analysis, in order to study the interactions between the environmental gradients, the alpha diversity indexes and the distribution of the different prokaryotic phyla. These results were obtained with the corrplot R package (v. 0.84; Wei & Simko, 2017) and with Hmisc (v. 4.3-1; Harrel, 2020). Additionally, the visualization plots were performed with the ggplot2 R package (v. 3.2.1; Wickham, 2016).

In the majority of the analysis performed (alpha and beta diversity, phyloseq, permanova, and Venn diagram), in order to have all samples with the same sequencing depth, i.e. the total number of sequences, the ASVs suffered a process of subsampling and were rarefied. As a result, 4834 and 3805 (Surface – 1994; DCM – 1811) reads for the depth and water mass effect, respectively. The taxonomic classification of the prokaryotic community of this analysis was ranked at seven levels: kingdom, phylum, class, order, family, genus, and species.

4. Results

Part A – Depth Effect

4.1. Biogeochemical Gradients across Depth

4.1.1. Environmental Parameters Across Depth

In order to study how the environmental data varies with depth, CTDs profiles from the different stations were analysed. In figure 5 it is possible to observe the oxygen (fig. 5a) and fluorescence (fig. 5b) profiles from Transect A samples. The oxygen profiles gave us information about the distribution of photosynthetic (high O_2 concentration depths), and chemolithotrophic organisms (low O_2 concentration depths). The fluorescence profiles gave indication of the distribution of the primary producers in the water column, since where fluorescence values were higher, chlorophyll values are also at its maximum. As it is visible in fig. 5a, the concentration of oxygen decreased with depth. Oxygen values reached their maximum between 0-100m and were close to zero at 500 m (having a mean value of 0.93 ml/L). Fluorescence values (fig. 5b) did not suffer high variation, being almost zero at all depths. However, between the 100-200m, there was a peak of these values, reaching almost 0.8mg/m^3. These values of fluorescence and depth corresponded to the deep chlorophyll maximum (DCM).

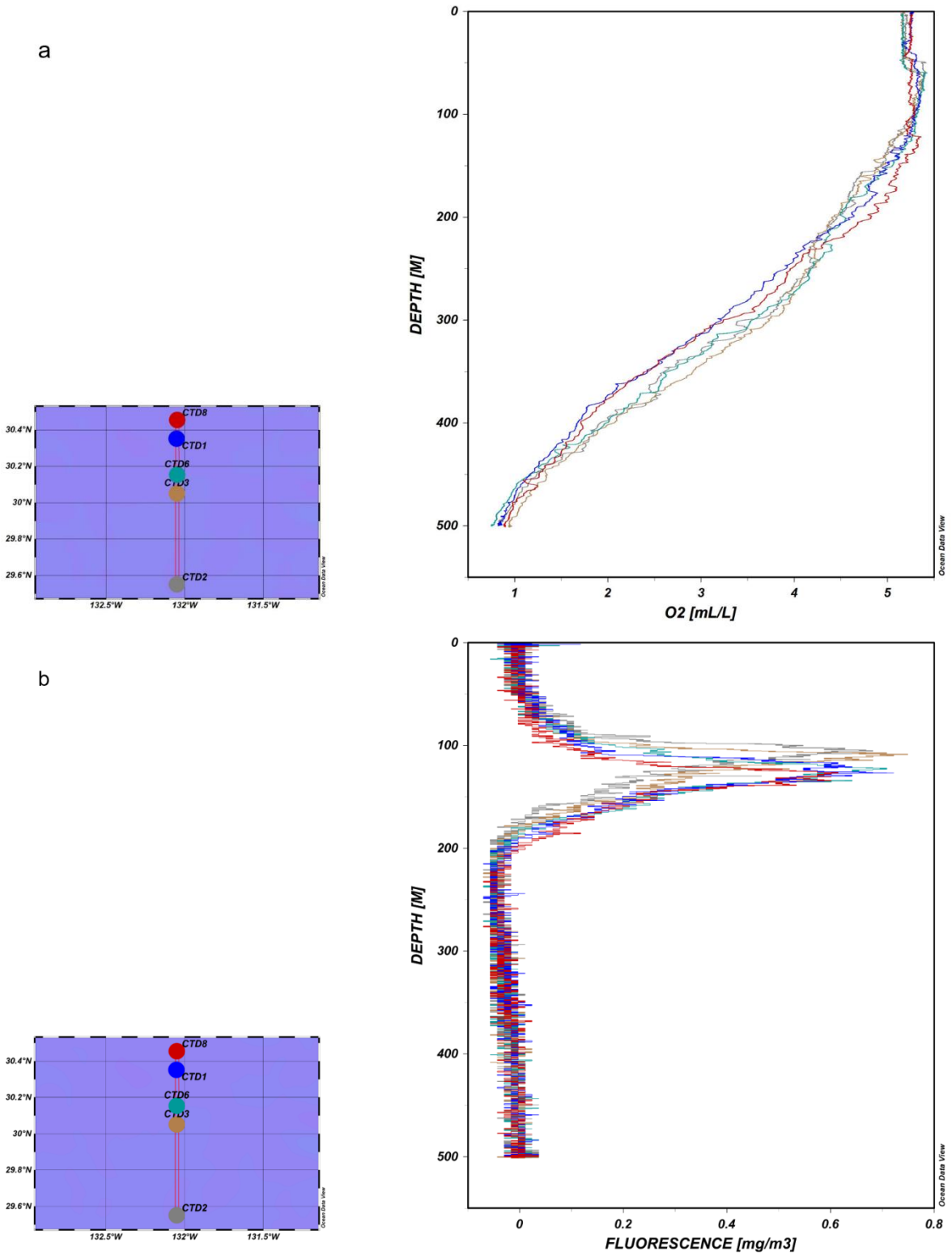


Figure 5: Oxygen (a) and fluorescence (b) water column profiles in transect A samples.

Principal component analysis (PCA) was performed using the variables presented in Table 5 as input. Through the water column, PC1 (Comp. 1) explains 68%, of the variation between the sample's distribution, and PC2 (Comp. 2) explains 15%. These two axes combined (fig. 6) explain 83% (68%+15%=83%) of the variance. PC1 is positively correlated with all environmental variables, except for depth. PC2, however, is only positively correlated with temperature and salinity (table 4). Figure 6 allows observing four main clusters, being Above DCM associated with Surface, providing a distribution pattern in the PCA analysis by the samples. These clusters represent distinct environments, with specific conditions in terms of the environmental variables analysed. Consequently, given that the five different depths are split along the x-axis (PC1) in the plot, it is possible to observe that the samples classified as DCM presented higher values of fluorescence. It is also visible that some samples classified as Surface (8A, 12A, 16A) and Above DCM had the highest values of salinity. In addition, an inversely proportional relation between depth and temperature and salinity was recognized. Samples at lower depths presented higher values of temperature and salinity, and vice-versa.

In table 4, the following environmental parameters are presented: temperature, salinity, depth, oxygen, turbidity, and fluorescence. With depth, it is possible to observe a decreasement of temperature, oxygen, and turbidity. Salinity values were relatively stable along the water column. Fluorescence values, and consequently chlorophyll values, reached a maximum at DCM (108-130m) and started to decrease Below DCM (175-200m). Additionally, according to table 4 and oxygen and fluorescence profiles (fig. 5), these parameters reached their maximum values at the same depth (DCM). Consequently, at 500m (Mesopelagic) all parameters had the lowest values, except for salinity values.

Table 4: PCA correlations of the environmental data and samples analysed for depth effect.

Environmental Data	**Comp. 1**	**Comp. 2**
Temperature	0.488	0.165
Salinity	0.291	0.481
Depth	-0.477	-0.116
O2	0.464	-0.109
Turbidity	0.435	-0.109
Fluorescence	0.211	-0.839

Table 5: Environmental data with chlorophyll depth classifications.

Chlorophyll Depth	Sample ID	Temperature (°C)	Salinity (PSU)	Depth (m)	O_2 (ml/L)	Turbidity (NTU)	Fluorescence (mg/m^3)
Surface (5m)	4A	18.06	34.2	5	5.26	0.06	-0.02
	8A	19.1	34.7	5	5.21	0.06	-0.02
	12A	19.5	34.8	5	5.17	0.06	-0.02
	16A	19.5	34.8	5	5.16	0.06	-0.003
	20A	18.7	34.2	5	5.25	0.06	0.01
Above DCM (80m)	19A	18.7	34.7	80	5.26	0.06	0.01
DCM (108-130m)	3A	16.6	34.6	120	5.22	0.05	0.63
	7A	15.6	34.2	110	5.21	0.06	0.56
	11A	16.5	34.4	108	5.26	0.06	0.67
	15A	16.2	34.3	122	5.26	0.06	0.67
	18A	15.8	34.1	130	5.33	0.07	0.56
Below DCM (175-200m)	2A	11.4	33.8	195	4.63	0.04	-0.02
	6A	10.6	33.8	200	4.44	0.04	-0.03
	10A	11.4	33.8	180	4.64	0.06	0.02
	14A	11.8	33.8	175	4.78	0.04	0.06
Mesopelagic (500m)	1A	5.7	34.1	500	0.83	0.04	-0.003
	5A	5.9	34.1	500	0.83	0.04	-0.002
	9A	5.7	34.1	500	0.95	0.04	0.01
	13A	5.7	34.1	500	0.76	0.04	-0.003
	17A	5.8	34.1	500	0.89	0.03	-0.003

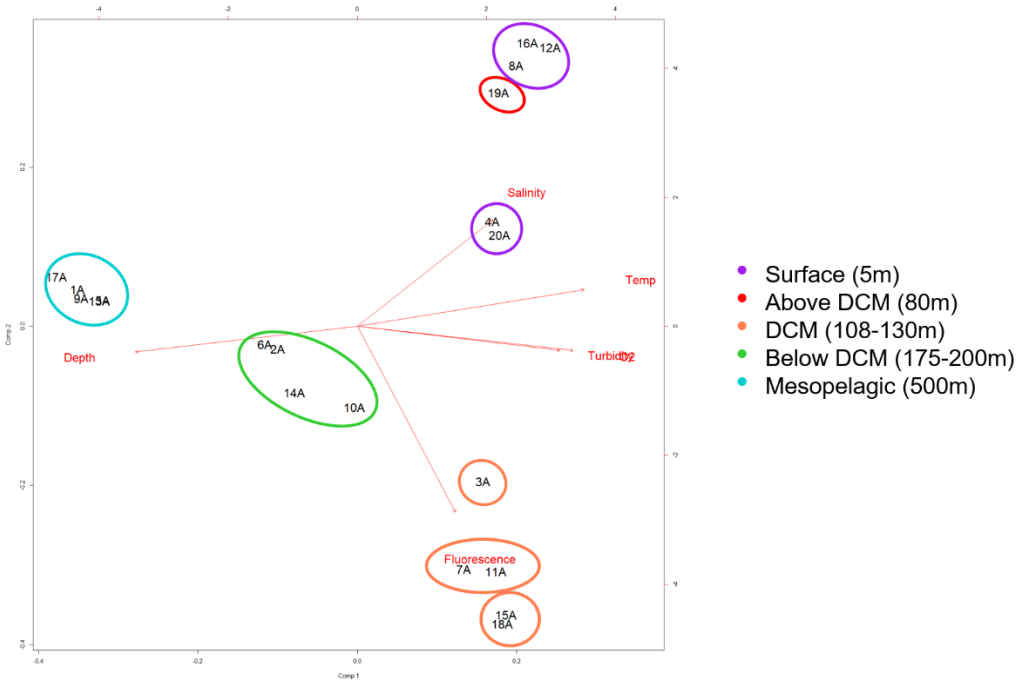


Figure 6: Principal Component Analysis of the samples to study the depth effect.

4.2. Prokaryotic Communities Distribution Across Depth

The total raw sequences (1345438) were quality-filtered, forward and reverse sequences were merged, and chimeras were removed, as well as sequences which had “Eukaryota”, “Chloroplast” and “Mitochondria” in their taxonomic classification. During this process, 566914 sequences were removed. Consequently, the final output was 778524 sequences. The cleaning process for each sample can be seen in table 6.

Table 6: Cleaning Process of the 16S rRNA sequences for each sample (*: the ASVs classified as Eukaryota/Chloroplast/Mitochondria were also removed).

Chlorophyll Depth	Sample ID	Input	Filtered	Merged	Non chimera	Final output*
Surface (5m)	4A	126354	110922	87921	74343	71932
	8A	38147	33093	28698	25229	24682
	12A	42451	36779	30636	25990	25161
	16A	69639	60311	49988	42178	40634
	20A	56439	48396	38827	33061	31881
Above DCM (80m)	19A	73434	64504	55625	49266	48399
DCM (108-130m)	3A	76148	66989	52880	46496	44108
	7A	47128	40784	30237	26243	24277
	11A	56201	49299	41506	36534	34659
	15A	72153	61682	46687	40521	37449
	18A	86160	74219	54846	47905	44623
Below DCM (175-200m)	2A	105949	91796	67750	60208	59927
	6A	30808	25956	17867	16033	16004
	10A	40801	34749	23614	20749	20621
	14A	78690	67297	52210	46010	45654
Mesopelagic (500m)	1A	73436	63165	46994	43416	43408
	5A	55018	47528	36174	33479	33465
	9A	73651	63312	50426	46580	46528
	13A	91066	79285	59762	55541	55495
	17A	51765	44047	31901	29624	29617

4.2.1. Alpha and Beta Diversity

To study the diversity of prokaryotic communities with depth, the Shannon, and Chao-1 indexes, along with the number of observed ASVs were calculated for each sample (fig. 7). In the figure 7 it is visible that samples followed a pattern of diversity. There was an increase of alpha diversity values with depth until “Below DCM” followed by a slight decrease in samples collected at 500m. Moreover, Surface samples registered the lowest values in all measures, particularly in the Shannon index. Being the sample 8A the one that presented the lowest values of the three alpha diversity measures. The Below DCM samples registered the highest values in the three measures, being the sample 2A the one that presented the highest values of the three alpha diversity measures. These results suggested that there is a biodiversity hotspot at Below DCM.

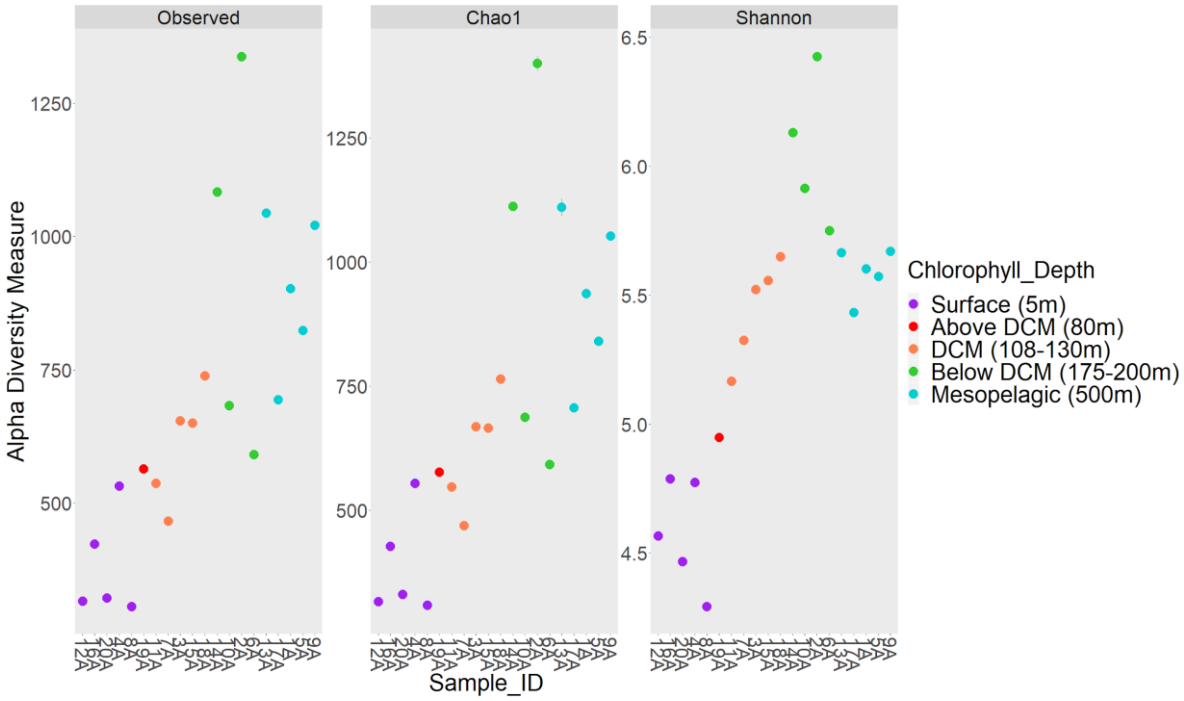


Figure 7: Alpha Diversity of the samples 1A-20A.

In order to study the similarities and differences in the structure of the prokaryotic community between samples, a hierarchical cluster, with logarithm (log) transformation, and multi-dimension scaling plot were performed (fig. 8 and fig. 9).

Regarding the hierarchical cluster (fig. 8), samples were distributed according to depth. Four main clusters were visible, Surface samples and Above DCM sample were grouped in one cluster, closer together with the cluster that grouped DCM samples. On the other hand, Below DCM samples were more related to Mesopelagic samples.

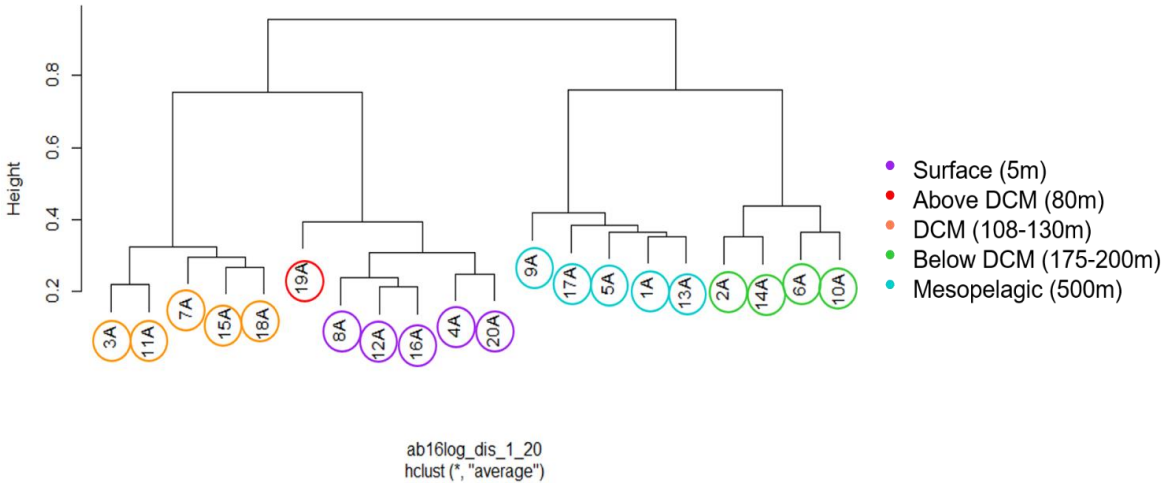


Figure 8: Hierarchical cluster of log transformation of samples 1A-20A.

The multi-dimension scaling plot (fig. 9) also demonstrated an obvious separation of the samples based on depth, similar to the previous analysis, suggesting the existence of different biological environments.

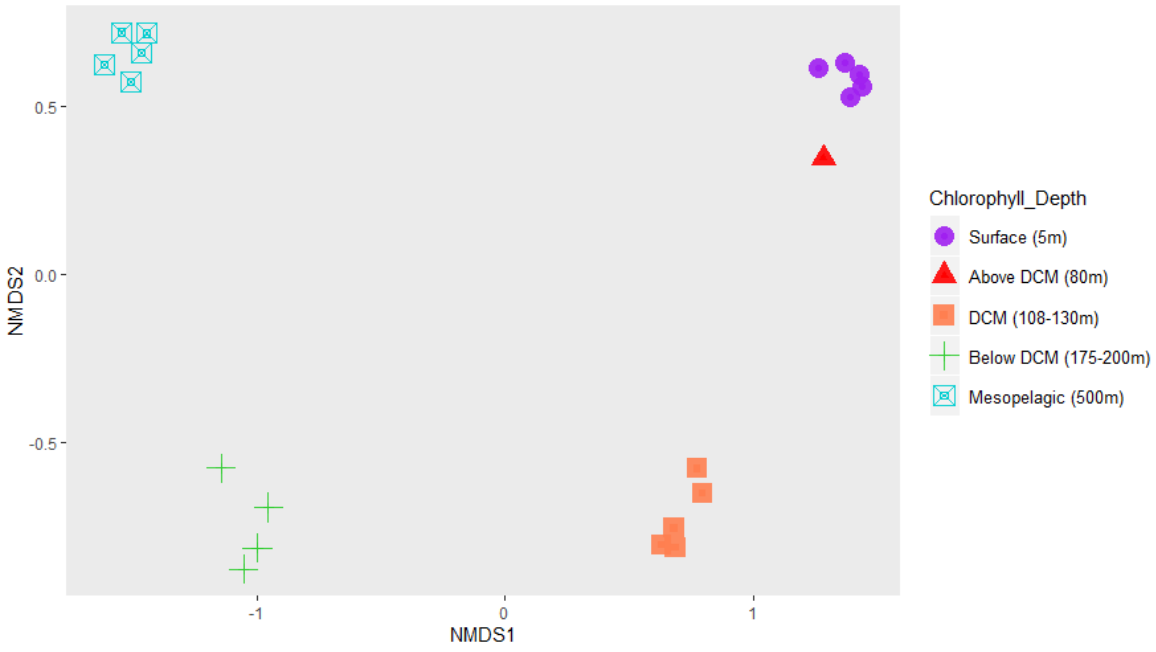


Figure 9: Multi-dimension scaling of the samples of samples 1A-20A.

To verify if depth influenced or not the distribution of prokaryotic communities, a permanova analysis was performed. Consequently, two hypotheses were created, with p-max=0.05.

H0: Depth does not influence the distribution of prokaryotic communities

H1: Depth influences the distribution of prokaryotic communities

The result (table 7) was p-value=0.01 (≤0.05), with F=35.354 with 4 to 19 degrees of freedom. As a result, H0 is rejected. Therefore, depth influenced the distribution of prokaryotic communities. Specifically, 90% (R^2=0.9041), approximately, of the variation of the prokaryotic community distribution was explained by the different depths.

Table 7: Output of permanova analysis of the samples 1A-20A.

	Df	**Sum of Sqs.**	**Mean Sqs.**	**F. Model**	**R²**	**Pr (>F)**
Chlorophyll Depth	4	5.6364	1.40910	35.354	0.9041	0.01
Residuals	15	0.5979	0.03986	0.0959	--	--
Total	19	6.2343	--	--	1.0000	--

4.2.2. Venn Diagram Analysis

Venn diagram analysis revealed that, only 10 out of 6445 ASVs were shared by the five different water column depth (fig. 10). The classification “Mesopelagic” was the one with more exclusive ASVs (1453) and “Above DCM” was the classification with less exclusive ASVs (124). “Mesopelagic” and “Below DCM” were the depth classifications that share more ASVs, 370 in total.

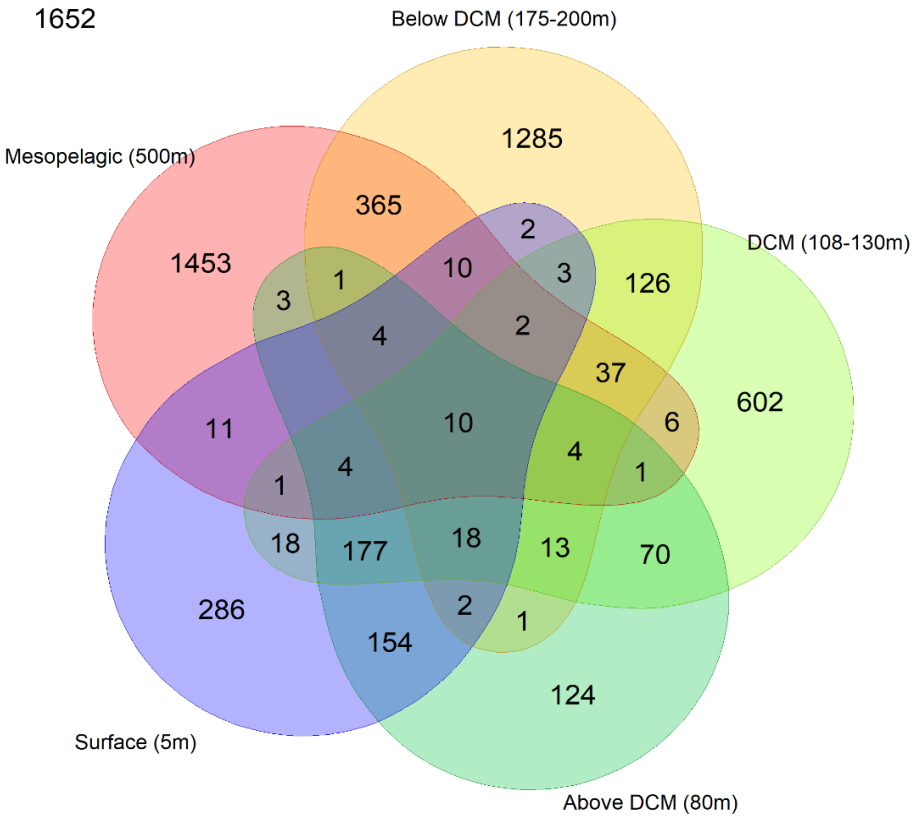


Figure 10: Venn diagram of the samples 1A-20A.

These results might be influenced by the species richness with depth, as previously described, and the different sample values in the different categories. If the relative amount of exclusive ASVs (fig. 11) is considered, the Above DCM sample had a lower number of species present on it and contrary, the Below DCM and Mesopelagic samples had a higher number of species on it, increasing the number of exclusive ASVs present on them. Furthermore, this Venn diagram demonstrated that with depth, the prokaryotic communities become more unique, suggesting that the higher the depth the less biological connectivity within the communities.

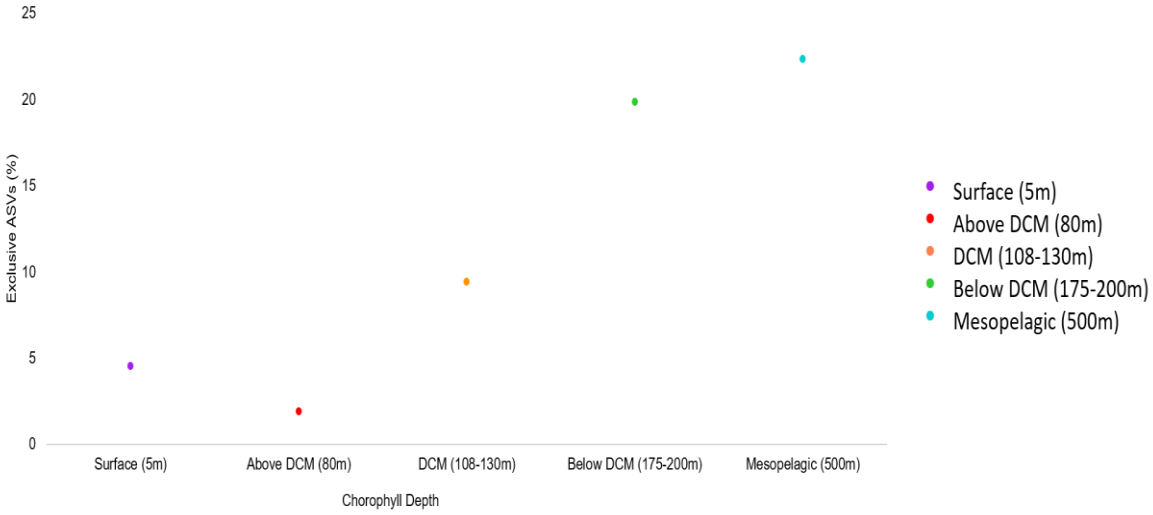


Figure 11: Exclusive ASVs (%) for each depth classification.

4.2.3. Prokaryotic Taxonomic Analysis

Proteobacteria was the most abundant phylum and was present in all of the five water columns depths. However, it is possible to detect depth preference by some phyla. While the abundance of Cyanobacteria and Bacteroidetes decreased with depth, and, Cyanobacteria was only present until 130m (with the exception of the sample 14A), Chloroflexi, Marinimicrobia (SAR 406 clade), Nitrospinae, Planctomycetes, Euryarchaeota, and Thaumarchaeota were more represented in deeper depths.

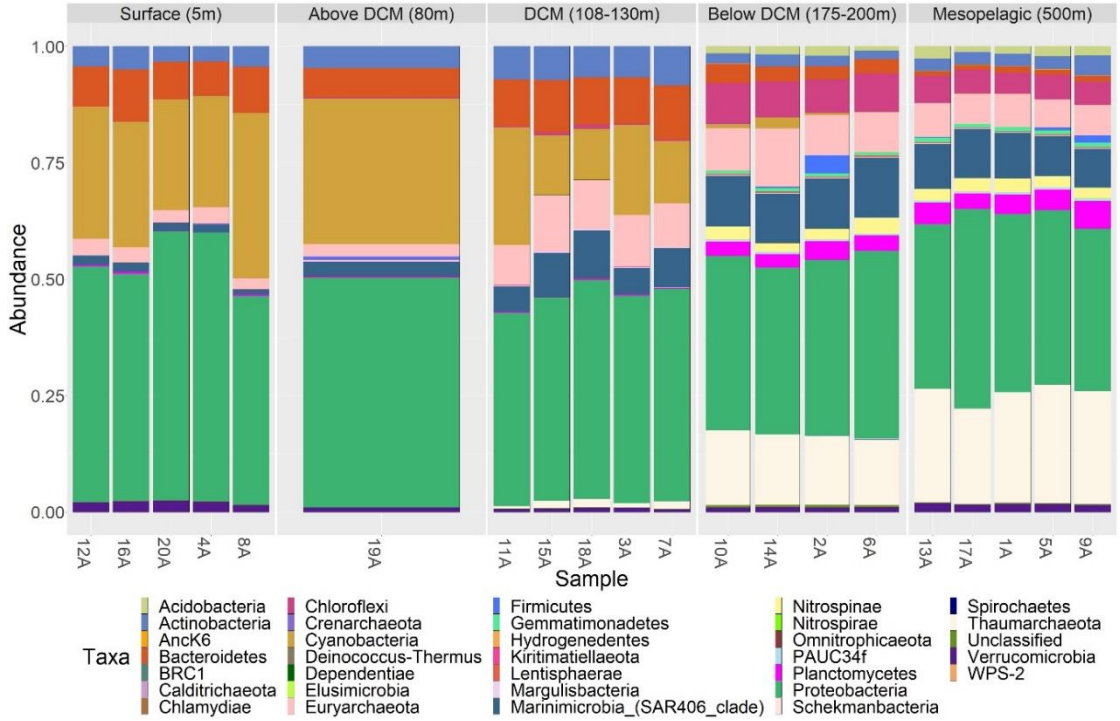


Figure 12: Phyla identification across the water column.

Alphaproteobacteria was the most abundant class and it was present in all five depths. There was also depth preference by the classes (fig. 13). For example, Nitrososphaeria only started to appear in deeper waters. Contrary, the abundance of Oxyphotobacteria decreased with depth. If the classes that belong to the Proteobacteria phylum (Alphaproteobacteria, Deltaproteobacteria, and Gammaproteobacteria) are analysed in detail, different patterns can be found. Although Alphaproteobacteria was present in all depth categories, their abundance started to decrease in Below DCM. On the contrary, Deltaproteobacteria abundance demonstrated a considerable increase with depth. This class was almost absent at Surface and Above DCM. Finally, Gammaproteobacteria abundance was constant across the water column. These different distributions between different classes of the same phylum represent the preference for distinct habitats, in this case, different depths, suggesting once again the presence of different ecological niches.

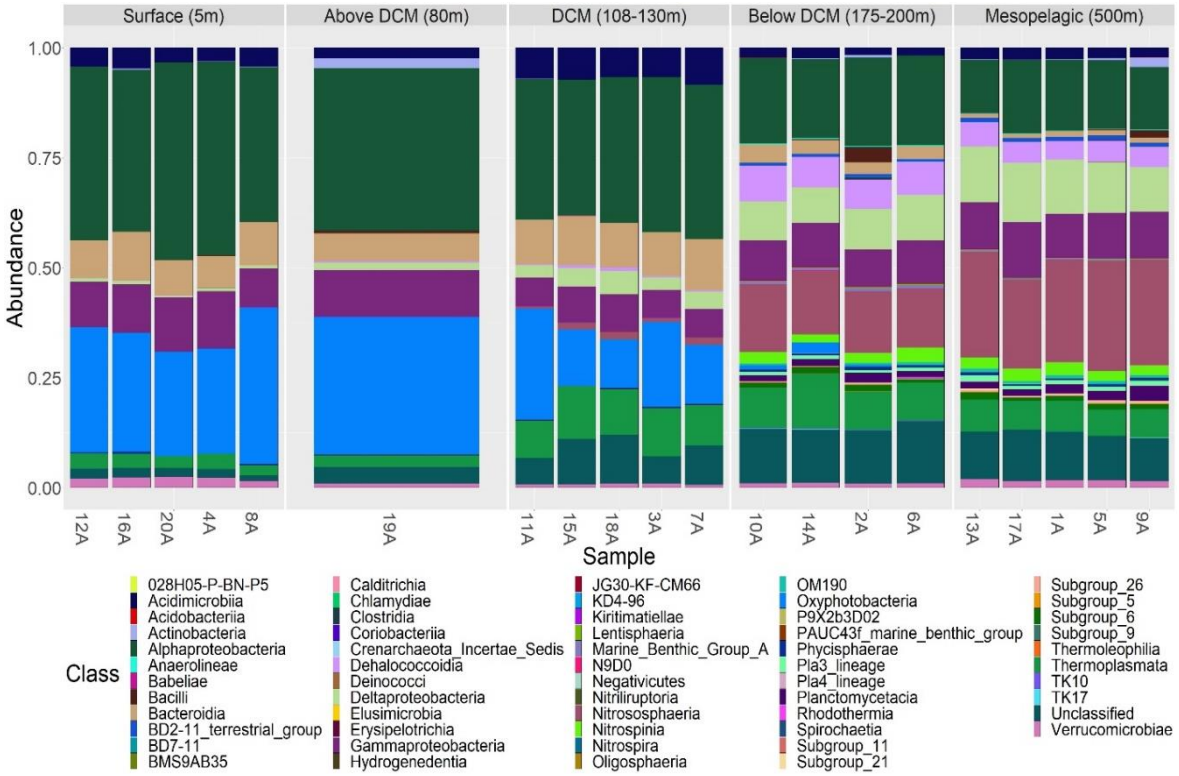


Figure 13: Class identification across the water column.

For Proteobacteria, Cyanobacteria, Thaumarchaeota, Marinimicrobia (SAR406 clade) and Euryarchaeota phyla (fig. 14), the distribution at the genera taxonomy level, were determined (Figs 15-19).

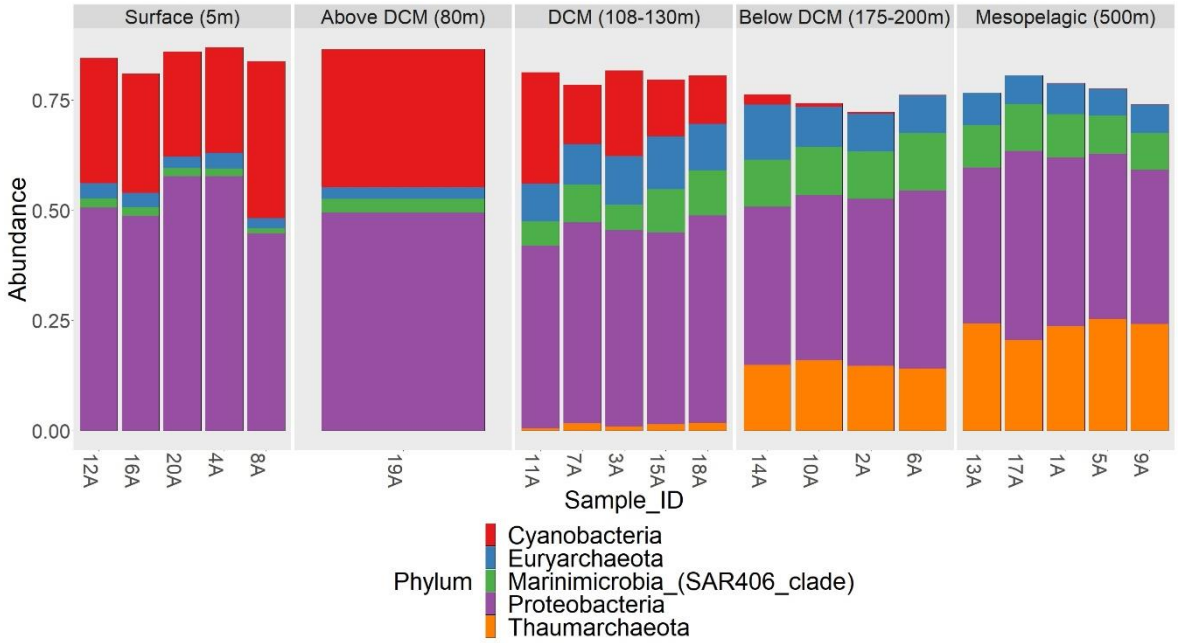


Figure 14: Top 5 phyla identification along the water column.

The Unclassified genus of Proteobacteria corresponded to 62.43%, approximately, of the total of the genera identified in this phylum. Even at this level, the depth preference was very clear for some genera. The genera Ascidiaceihabitans, Candidatus Puniceispirillum, and Clade Ia, belonging all to the class Alphaproteobacteria, were only present in lower depths (until DCM). At higher depths, it was visible the increase of the abundance of the genus SUP05 cluster, from the class Gammaproteobacteria. Additionally, the genus Acinetobacter, from the class Gammaproteobacteria, was only present in the sample 17A (Mesopelagic). A genus that was present across the water column (identified in the five different depth classifications) was Clade Ib, from the class Alphaproteobacteria, with relatively high values of abundance.

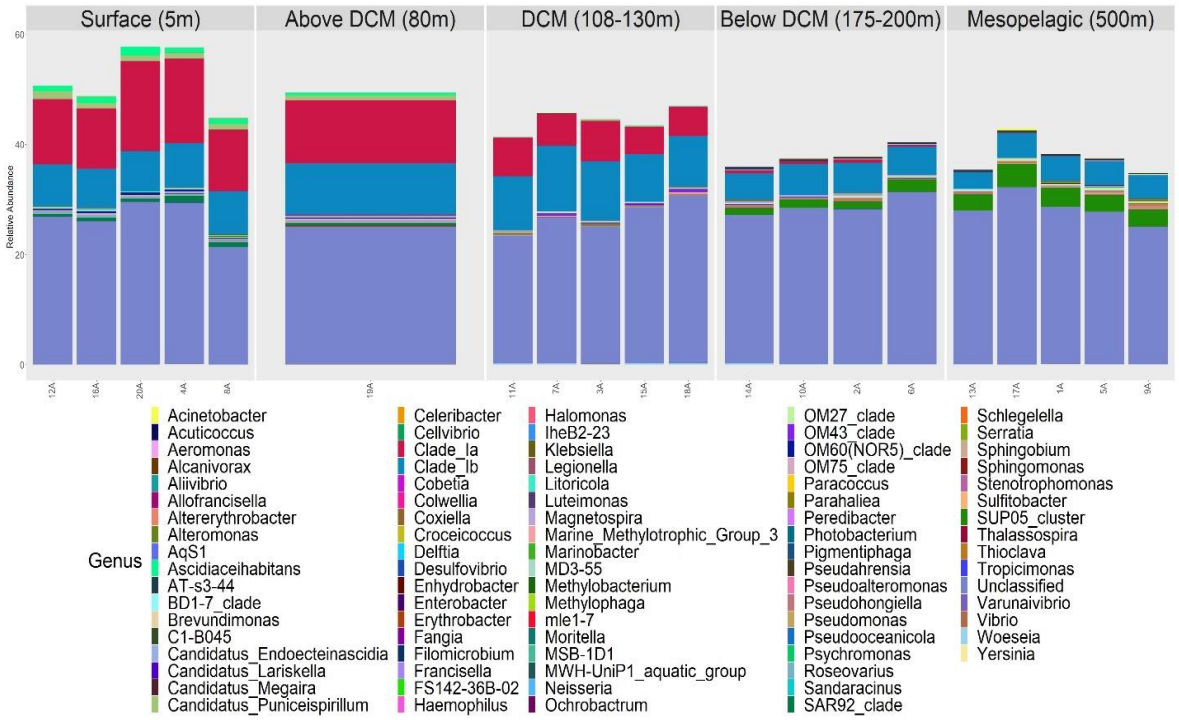


Figure 15: Proteobacteria genera across the water column.

The most common genus of Cyanobacteria was Prochlorococcus MIT9313, from the class Oxyphotobacteria. It was possible to observe that the genus Atelocyanobacterium (UCYN-A), also from the class Oxyphotobacteria, was only present at the surface (fig. 16). The Unclassified genera of Cyanobacteria corresponded to 0.19%, approximately, of the total genera identified in this phylum.

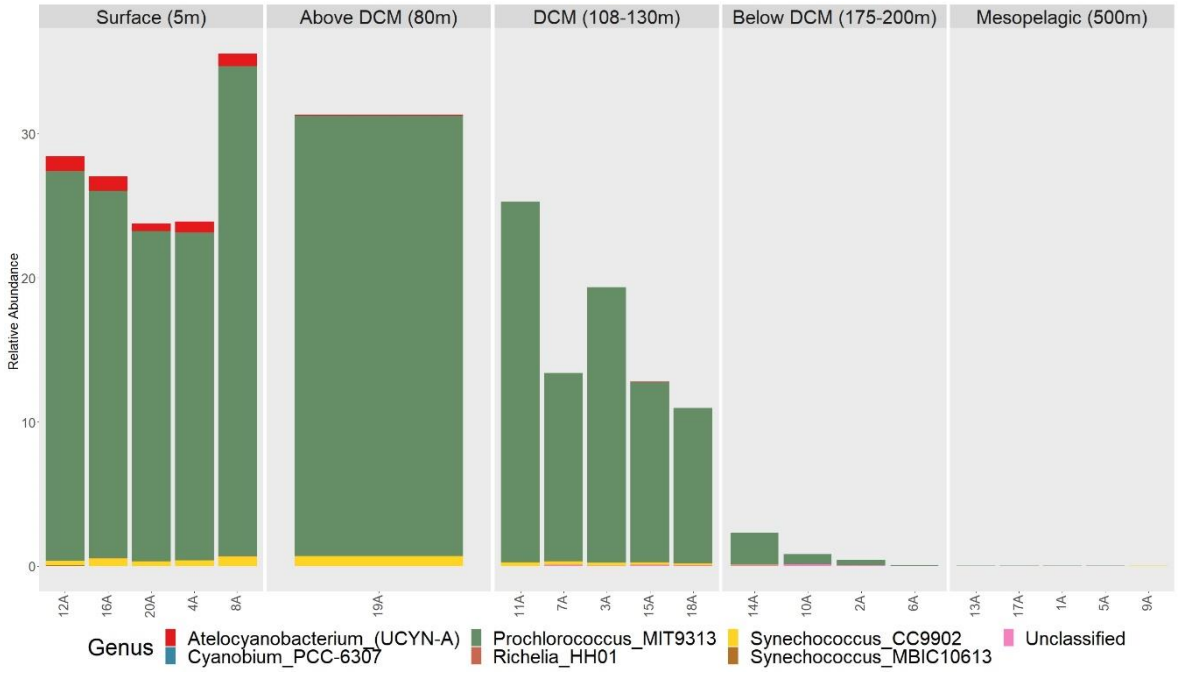


Figure 16: Cyanobacteria genera across the water column.

In the Thaumarchaeota phylum the genus Candidatus Nitrosopelagicus, from the class Nitrososphaeria, was the only one present in “DCM” and it was very abundant in “Below DCM”. The genus Candidatus Nitrosapulimus, also from the class Nitrososphaeria, only appeared in “Below DCM” and “Mesopelagic” but had a lower abundance. The Unclassified genera of Thaumarchaeota corresponded to 80.81%, approximately, of the total genera identified in this phylum.

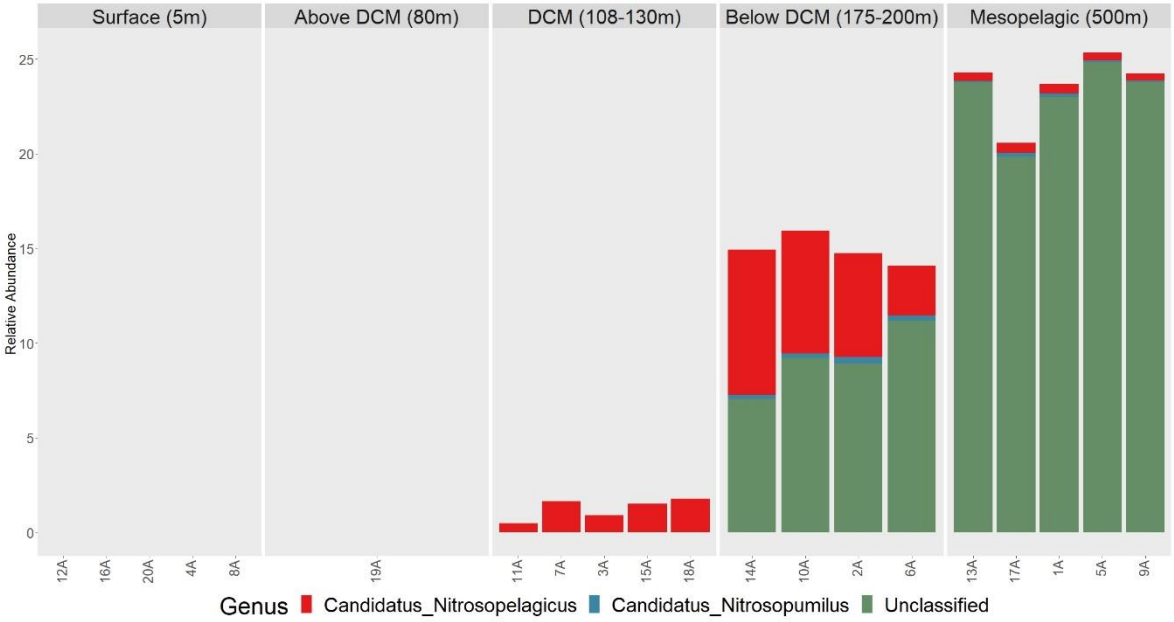


Figure 17: Thaumarchaeota genera across the water column.

All the genera of the phyla Marinimicrobia (SAR406 clade) and Euryarchaeota present in this analysis were registered as Unclassified (Figs. 18 and 19).

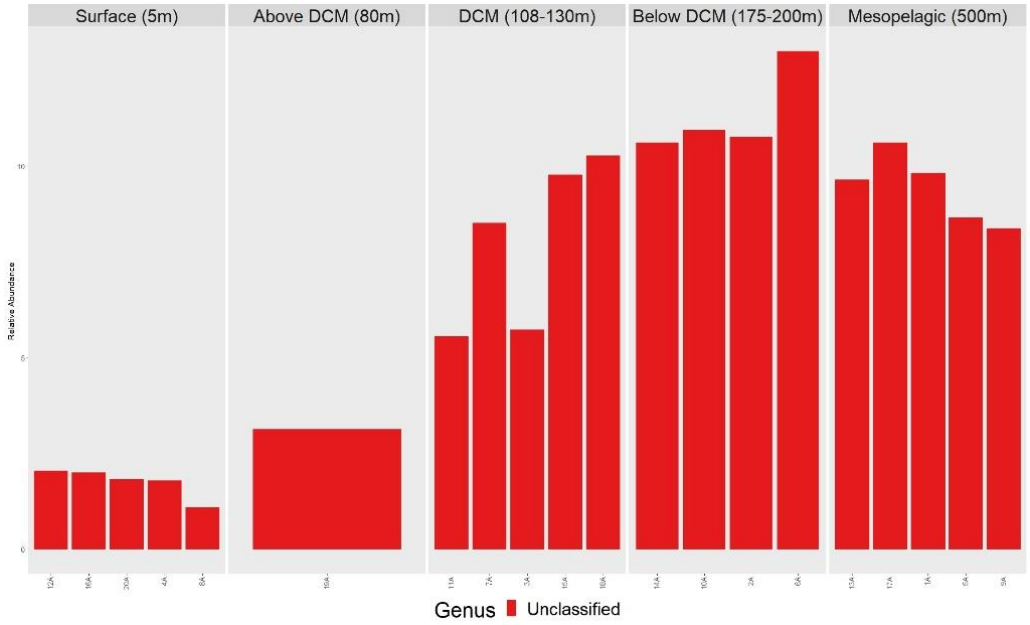


Figure 18: Marinimicrobia (SAR406 clade) genera across the water column.

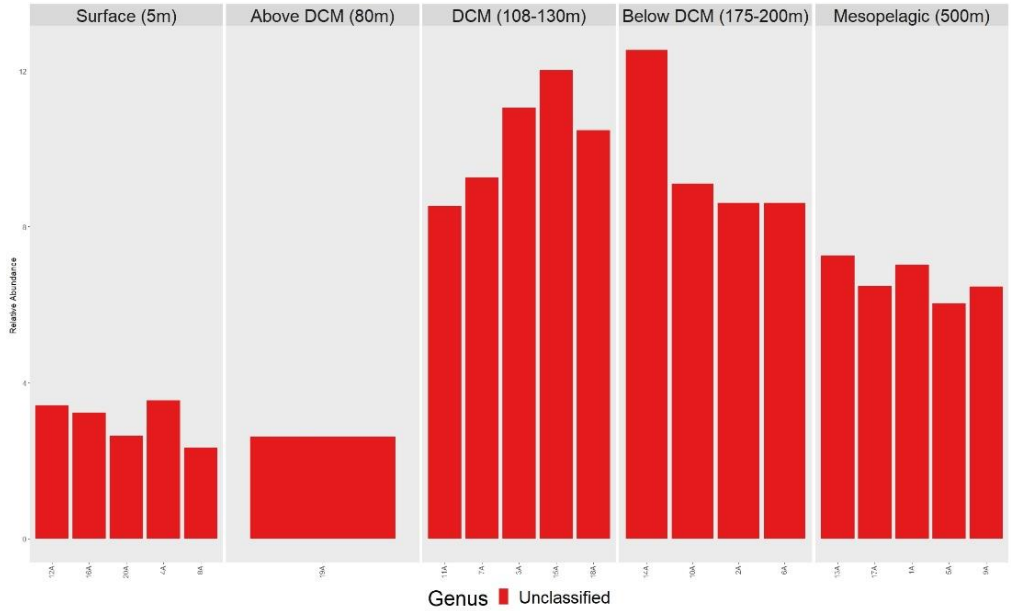


Figure 19: Euryarchaeota genera across the water column.

4.3. Environmental vs Biological Relationships

In order to explore the relationships between the environmental and biological data, two different analyses were performed: a canonical correspondence analysis (CCA) – to explore the relationship between the variables and the samples – and a correlation analysis with the Spearman method – to identify what significant relationships existed between the environmental and biological data.

An envfit test was performed and as a result, it was observed that all the environmental parameters included in the analysis (table 5) were statistically significant for the distribution of the samples (table 8). Surface samples and the sample 19A (classified as Above DCM) revealed a considerable influence of salinity. The samples classified as DCM were more associated with fluorescence. Oxygen and turbidity seemed to have a smaller influence on these samples. The samples classified as Below DCM and Mesopelagic showed a clear correlation with the depth. The temperature had a relatively neutral effect (fig. 20).

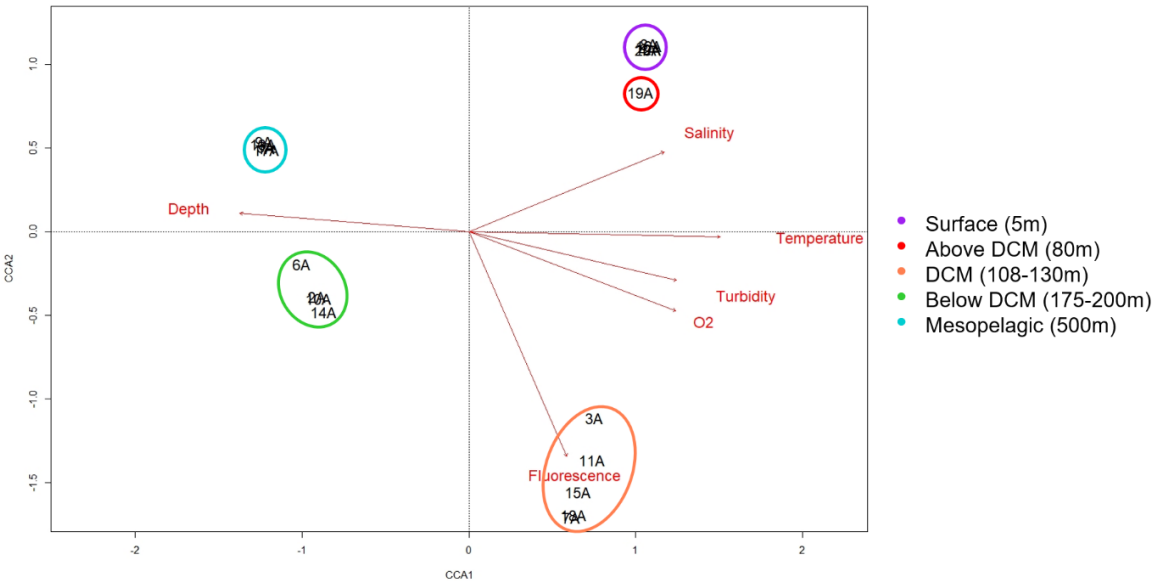


Figure 20: Canonical Correspondence Analysis of the samples 1A-20A.

Table 8: Envfit test results for the depth effect samples.

Environmental Data	***p-value***	***R^2***
Temperature	0.000999	0.7325
Salinity	0.003996	0.4986
Depth	0.000999	0.6306
Oxygen	0.007992	0.4501
Turbidity	0.001998	0.5525
Fluorescence	0.003996	0.5663

Figure 21 shows the correlation analysis, with the Spearman method, between the environmental data (temperature, salinity, depth, oxygen, turbidity, and fluorescence), the alpha diversity indexes (Chao-1 and Shannon) and the relative abundance of the different phyla. In this graphic, it is only marked the significant relationships (p-value≤0.05).

The environmental parameters and the distribution of the prokaryotic community were highly correlated. The different conditions along the water column influenced these abundance shifts by the phyla. For example, the highly abundant Cyanobacteria at lower depths was positively correlated with all environment parameters, with the exception of depth ($r = -0.9$, p-value = 2.6×10^{-8}). Bacteroidetes, additionally, presented similar results with depth being the only negative correlation with the environmental parameters of this phylum ($r = -0.8$, p-value = 6.5×10^{-5}). Contrarily, Thaumarchaeota, very abundant in deeper areas, presented a negative correlation with all environmental parameters, except for depth that presented a strong positive correlation ($r = 0.9$, p-value = 9.98×10^{-11}). Similar relationships, with depth the only positive correlation of the environmental parameters, were Acidobacteria ($r = 0.9$, p-value = 1.8×10^{-6}) and Nitrospinae ($r = 0.9$, p-value = 6.3×10^{-9}). Fluorescence was the environmental parameter that presented the least significative relationships with the different phyla, only presenting two positive significative relationships with Actinobacteria ($r = 0.6$, p-value = 0.006) and Margulisbacteria ($r = 0.6$, p-value = 0.005).

Regarding the alpha diversity indexes, the majority of the phyla demonstrated a positive correlation with these parameters. However, Bacteroidetes (Chao-1: $r = -0.5$, p-value = 0.03; Shannon: $r = -0.5$, p-value = 0.04) and Cyanobacteria (Chao-1: $r = -0.6$, p-value = 0.002; Shannon: $r = -0.7$, p-value = 0.04) were the only phyla which were negatively correlated with these indexes.

Calditrichaeota, Dependentiae, Elusimicrobia, Proteobacteria, Verrucomicrobia, and WPS-2 did not show significative relationships with the parameters analysed.

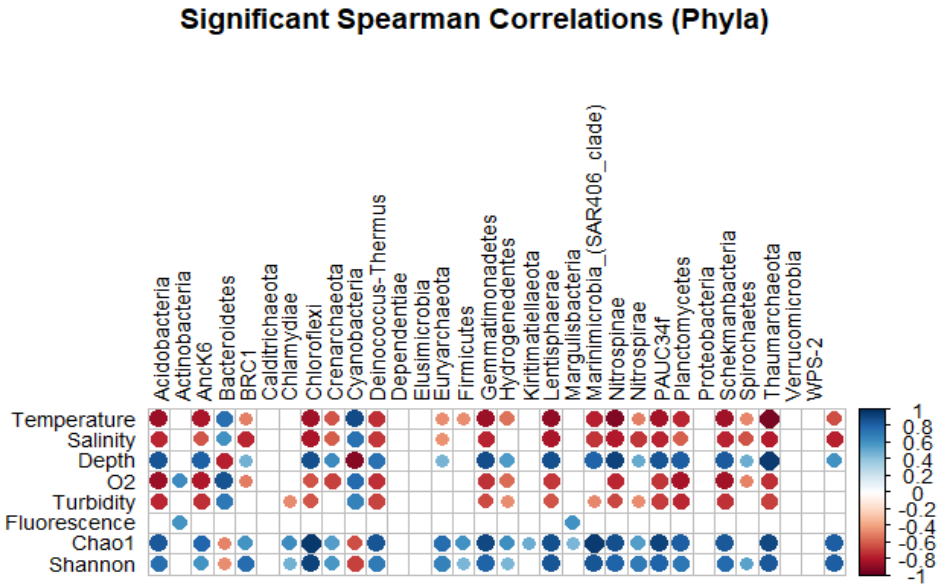


Figure 21: Spearman correlation analysis.

Part B – Water Mass Effect

4.4. Biological and Biogeochemical Gradients Across the Water Masses

4.4.1. Environmental Gradients Across the Water Masses

In tables 9 and 10, the following environmental parameters are presented for Surface and DCM water column depths: temperature, salinity, depth, oxygen, turbidity, and fluorescence, alongside with the water mass classifications. However, some environmental parameters were not available (NA) and were not registered in some samples.

At Surface, the Polar water mass registered the highest value of oxygen (5.26 ml/L) and the Front water mass registered the lowest value (4.84 ml/L). The mean value of oxygen at Polar water mass was 5.26 ml/L, at Front was 5.03 ml/L and at Subtropical was 5.14 ml/L. Turbidity was the most constant and stable parameter recorded in all three water masses, having the same value (0.06 NTU). Fluorescence values were not measurable in all surface samples At DCM, the highest value of oxygen was recorded in the water mass Polar (5.33 ml/L). The mean value of oxygen at Polar water mass was 5.27 ml/L and at Front was 5.22 ml/L. In terms of turbidity, Polar water mass also registered the highest value of this parameter (0.07 NTU). The mean value for turbidity at Polar and Front waters was 0.06 NTU and 0.04 NTU, respectively. Regarding the fluorescence values, Polar was also the water mass that registered the highest values of this parameter (0.67 mg/m3). The mean value of fluorescence at Polar waters was 0.6 mg/m3 and at the Front was 0.48 mg/m3.

If both depths are compared (Surface and DCM), fluorescence demonstrated a considerable increase with depth. Oxygen also demonstrated a small increase with depth. However, turbidity was the parameter that remained stable in both depths.

The distinction between the two water masses was still evident at DCM (fig. 22b), in terms of temperature and salinity. In general, Polar water samples had the lowest range of values of temperature and salinity, and Front samples the highest range of values. The samples 56A and 58A, both classified as Polar, were the exception that presented higher values of temperature. Between the two water masses, a clear directly proportional relationship was not clear as at Surface. Nevertheless, if Front samples were analysed individually this relationship was present: the higher the temperature, the higher the salinity.

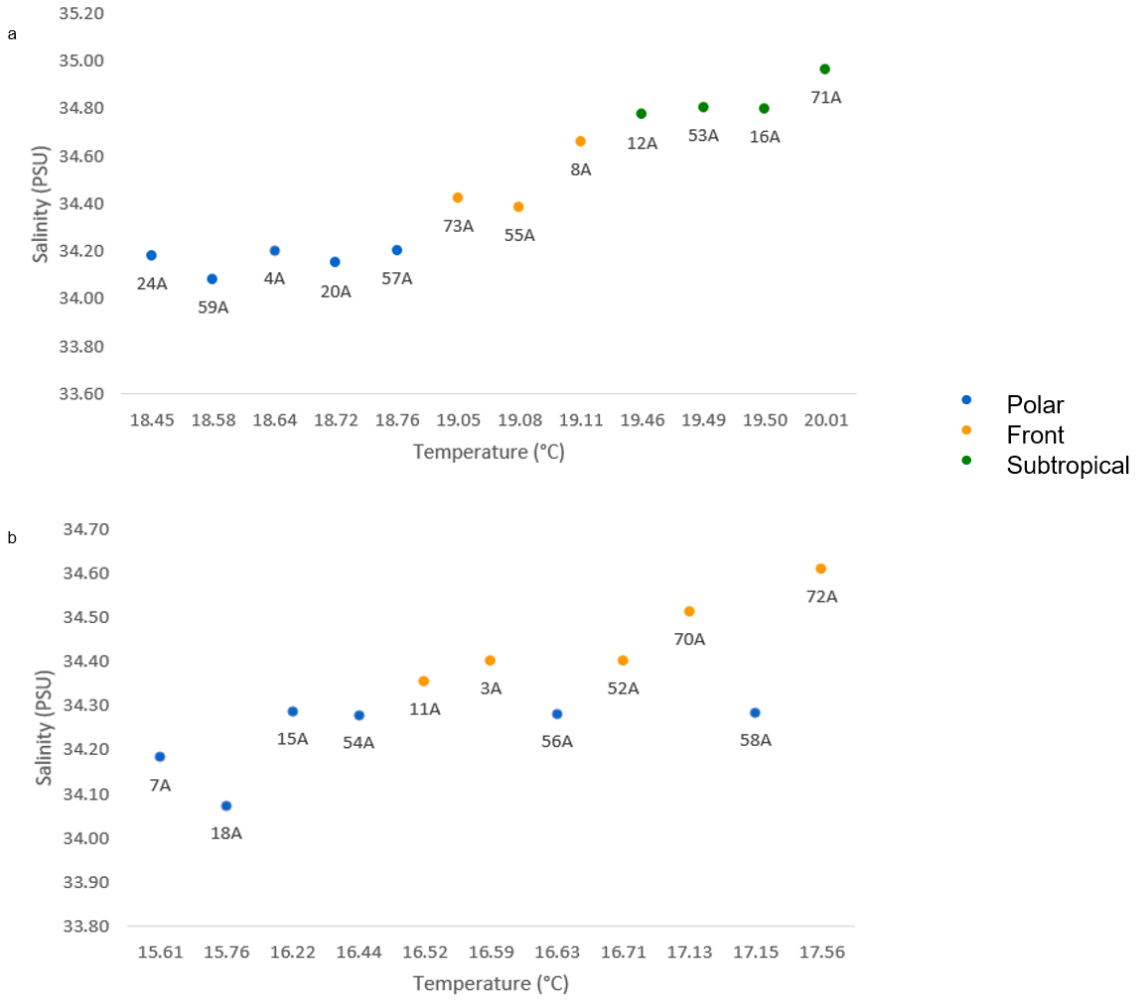


Figure 22: Water masses classification and Temperature-Salinity relationship (Surface-a; DCM-b).

Table 9: Environmental data of the samples collected at Surface.

Water Mass	***Sample ID***	***Temperature (°C)***	***Salinity (PSU)***	***Depth (m)***	***O_2 (ml/L)***	***Turbidity (NTU)***	***Fluorescence (mg/m^3)***
Polar	4A	18.64	34.21	5	5.26	0.06	-0.02
	20A	18.72	34.15	5	5.25	0.06	-0.0032
	24A	18.45	34.18	5	NA	NA	NA
	57A	18.77	34.20	3	NA	NA	NA
	59A	18.58	34.08	3	NA	NA	NA
Front	8A	19.11	34.66	5	5.21	0.06	-0.02
	55A	19.08	34.39	3	NA	NA	NA
	73A	19.05	34.42	3	4.84	0.06	-0.02
Subtropical	12A	19.46	34.78	5	5.16	0.08	-0.03
	16A	19.50	34.8	5	5.16	0.06	-0.02
	53A	19.49	34.81	3	NA	NA	NA
	71A	20.01	34.96	3	5.09	0.03	-0.0032

NA – Not Available

Table 10: Environmental data of the samples collected at DCM.

Water Mass	***Sample ID***	***Temperature (°C)***	***Salinity (PSU)***	***Depth (m)***	***O_2 (ml/L)***	***Turbidity (NTU)***	***Fluorescence (mg/m^3)***
Polar	7A	15.61	34.18	110	5.21	0.06	0.56
	15A	16.22	34.29	122	5.26	0.06	0.67
	18A	15.76	34.07	130	5.33	0.07	0.56
	54A	16.44	34.28	115	NA	NA	NA
	56A	16.63	34.28	115	NA	NA	NA
	58A	17.15	34.29	115	NA	NA	NA
Front	3A	16.59	34.40	120	5.22	0.05	0.63
	11A	16.52	34.36	108	5.25	0.06	0.66
	52A	16.71	34.40	123	NA	NA	NA
	70A	17.13	34.51	107	5.15	0.03	0.49
	72A	17.56	34.61	107	5.27	0.03	0.14

NA – Not Available

4.5. Prokaryotic Distribution Across the Water Masses

The total raw sequences (1733232) were quality-filtered, forward and reverse sequences were merged, and chimeras were removed, as well as sequences which had “Eukaryota”, “Chloroplast” and “Mitochondria” in their taxonomic classification. During this process, 566914 sequences were removed. In addition, after the cleaning process, 675652 sequences were removed. The final output was 1057580 sequences present in the ASV table, as it is visible in tables 11 and 12.

Table 11: Cleaning Process of the ASVs present at Surface (*: the ASVs with Eukaryota/Chloroplast/Mitochondria were also removed from the ASV table).

Water Mass	*Sample ID*	*Input*	*Filtered*	*Merged*	*Non chimera*	*Output**
Polar	4A	126354	110922	87921	74343	74343
	20A	56439	48396	38827	33061	33061
	24A	167446	147108	128772	108571	108571
	57A	73722	63353	54045	46360	46360
	59A	84392	74480	61741	52971	52971
Front	8A	38147	33093	28698	25229	25229
	55A	61149	53910	48679	43494	43494
	73A	63945	53467	45076	39264	39264
Subtropical	12A	42451	36779	30636	25990	25990
	16A	69639	60311	49988	42178	42178
	53A	112195	97237	74060	63342	63342
	71A	79395	68638	55904	48554	48554

Table 12: Cleaning Process of the ASVs present at DCM (*: the ASVs with Eukaryota/Chloroplast/Mitochondria were also removed from the ASV table).

Water Mass	***Sample ID***	***Input***	***Filtered***	***Merged***	***Non chimera***	***Output****
Polar	7A	47128	40784	30237	26243	26243
	15A	72153	61682	46687	40521	40521
	18A	86160	74219	54846	47905	47905
	54A	93900	82444	68938	61373	61373
	56A	74972	64573	52841	44776	44776
	58A	82042	70697	53982	47564	47564
Front	3A	76148	66989	52880	46496	46496
	11A	56201	49299	41506	36534	36534
	52A	56738	48266	38644	34297	34297
	70A	58190	49346	37705	32907	32907
	72A	54326	47231	39951	35607	35607

In fig. 23 (a-surface, b-DCM), it is possible to verify the different sequencing depths of the samples. Nevertheless, all the samples reached the plateau.

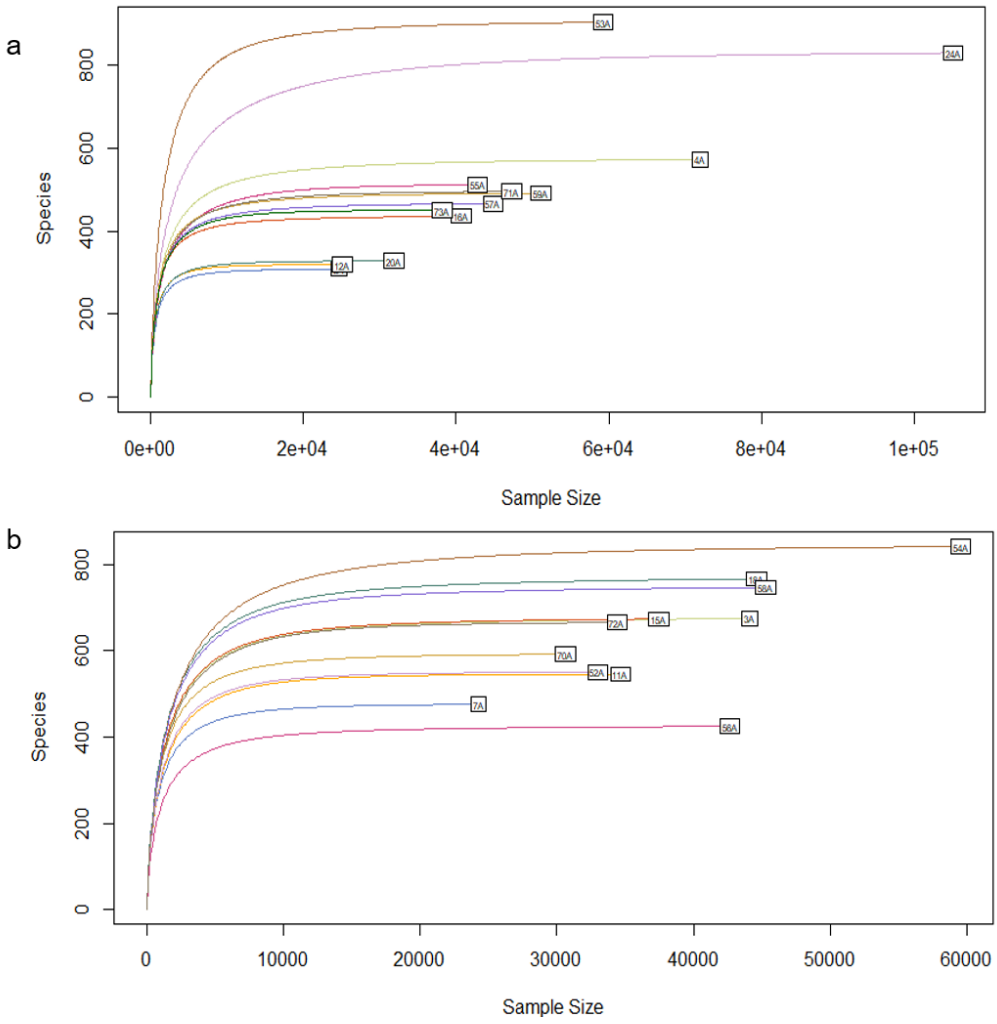


Figure 23: Rarefaction curve of samples collected at Surface (a) and DCM (b).

4.5.1. Alpha and Beta Diversity

To study the alpha diversity across the different water masses (Polar, Front, and Subtropical), the Shannon and Chao-1 indexes, along with the number of observed ASVs were analysed in the different depths (Surface and DCM). Additionally, the Tukey test was performed to compare, with a statistical support, the alpha diversity values between the water masses.

At surface, the lowest value of alpha diversity was registered at the front samples (fig. 24). Subtropical waters presented the highest range of alpha diversity between the three water masses. Additionally, this water mass had an outlier. This outlier had the highest values of alpha diversity, being followed by the Polar outlier. The sample that corresponded to this outlier could be part of a biodiversity hotspot, explaining the big difference between samples. If more samples were collected, it would give a better perspective of these

tendencies. At Surface no significative differences were observed in terms of species richness and the diversity indexes between the three water masses (table 13).

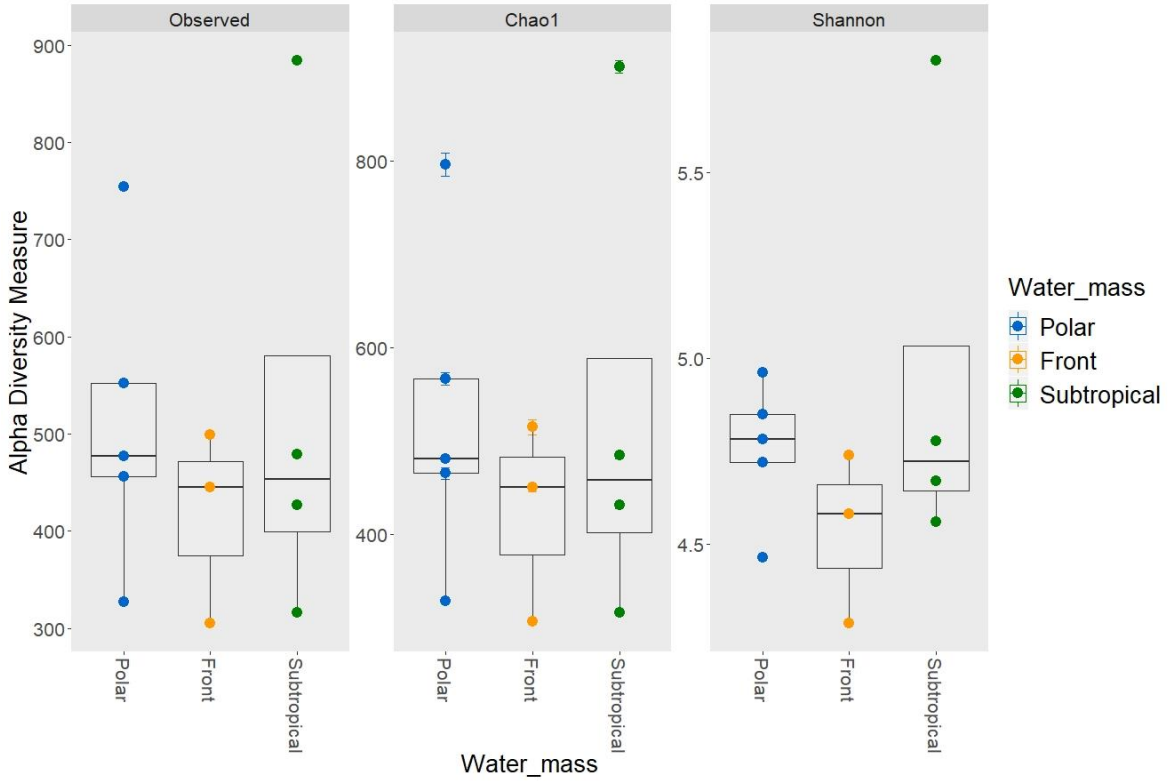


Figure 24: Alpha Diversity of the samples Water Mass-Surface.

Table 13: Tukey test results of alpha diversity and the water masses at Surface.

α diversity	***Water Mass***	***p-value***
Observed	Polar-Front	0.75
	Polar-Subtropical	0.99
	Front-Subtropical	0.71
Chao-1	Polar-Front	0.74
	Polar-Subtropical	0.99
	Front-Subtropical	0.70
Shannon	Polar-Front	0.67
	Polar-Subtropical	0.76
	Front-Subtropical	0.36

At DCM, Polar waters presented a higher variation of alpha diversity measures (fig. 25). This water mass had samples with the highest and lowest values of alpha diversity. Front waters were more uniform, although. Moreover, at this depth, significant differences between the two water masses in terms of species richness and diversity were not observed (table 14). If more samples were collected, it would give a better understanding of the diversity patterns at the DCM depth.

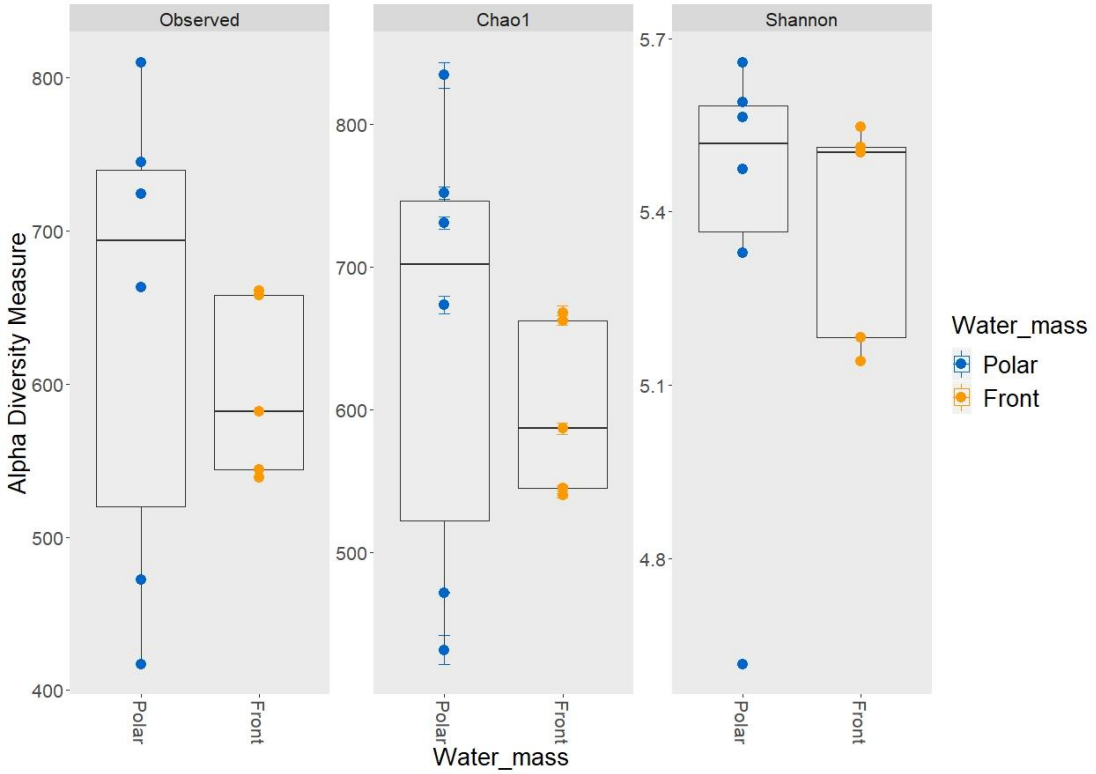


Figure 25: Alpha Diversity of the samples Water Mass-DCM.

Table 14: Tukey test results of alpha diversity and the water masses at DCM.

α diversity	***Water Mass***	***p-value***
Observed	Polar-Front	0.59
Chao-1	Polar-Front	0.55
Shannon	Polar-Front	0.98

In order to study the similarities and differences in the structure of the prokaryotic community between samples, hierarchical clusters, with logarithm (log) transformation, at Surface and DCM were performed. For the Surface samples the sample 53A (classified as Subtropical)

was found to be highly dissimilar from all the other samples (fig. 26). This result was similar to the alpha diversity output. With the aim to have a better comprehension of the sample distribution, a second dendrogram was performed, without the sample 53A (fig. 26b). In this second hierarchical cluster analysis two main branches were also generated At one branch, the samples 55A and 73A (classified as Front) were included. The other main branch was divided into two sub-branches. On one branch, include samples 12A and 16A (classified as Subtropical) and sample 8A (classified as Front). On the other branch, all the Polar samples were together with the exception of sample 71A (classified as Subtropical).

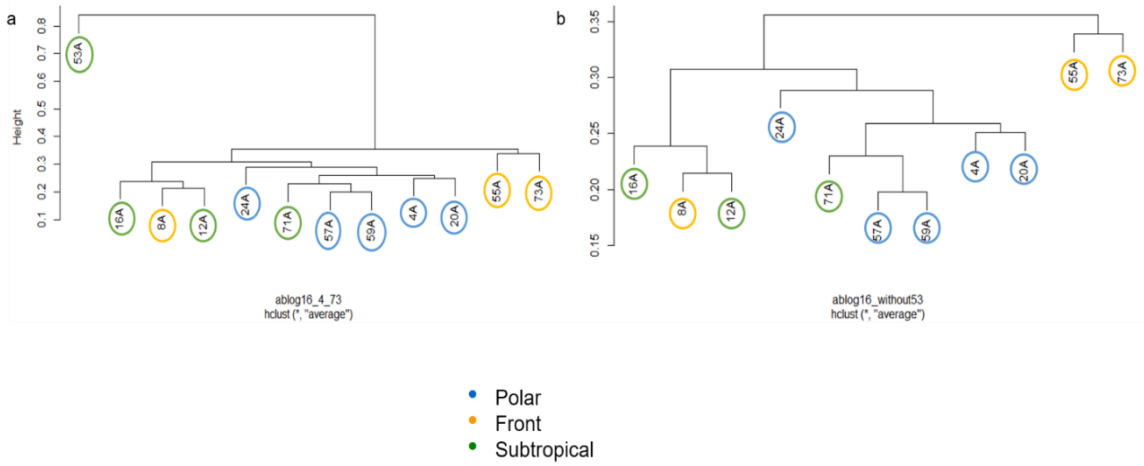


Figure 26: Hierarchical clusters of log transformation with (a) and without (b) the sample 53A of Water Mass-Surface.

At DCM, similar to Surface, the sample 56A (classified as Polar) were found to be highly dissimilar (fig. 27a) and thus a second dendrogram was performed without this sample (fig. 27b). In this second analysis, also two main branches were generated, however, samples from both water masses were present in both branches.

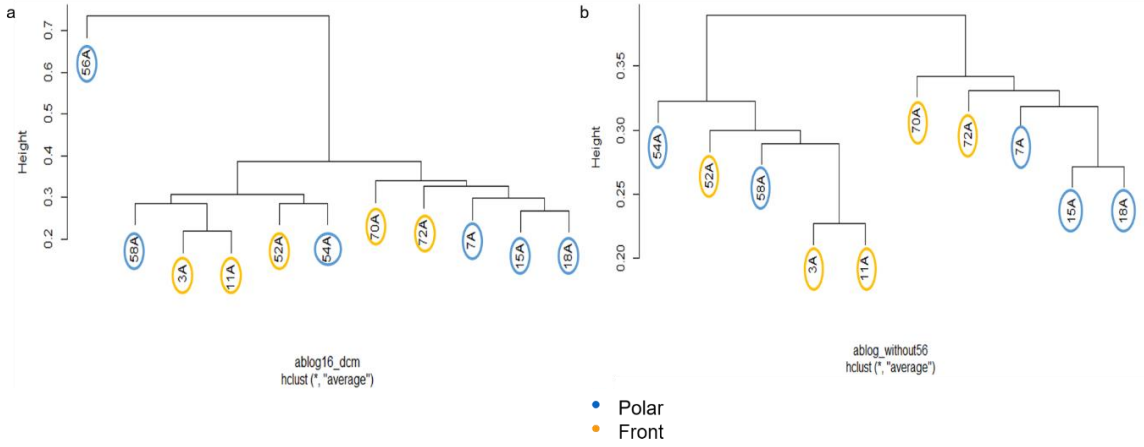


Figure 27: Hierarchical clusters of log transformation with (a) and without (b) the sample 56A of Water Mass-DCM.

To verify if water mass, at both depths, influences, or not the distribution of prokaryotic communities, a permanova analysis was performed. Consequently, two hypotheses were created, with p-max=0.05.

H0: Water mass does not influence the distribution of prokaryotic communities

H1: Water mass influences the distribution of prokaryotic communities

Surface: The result was p-value =0.05 (p0.05), with F=1.4538 with 2 to 11 degrees of freedom (table 15). As a result, H0 was rejected. Therefore, the type of water mass is lightly responsible and is approaching significance for the distribution of the prokaryotic communities at Surface. The type of water mass was responsible for 24% (R^2=0.24417) of the variation of the prokaryotic community.

Table 15: Output of permanova analysis of Water mass-Surface.

	Df	Sum of Sqs	Mean Sqs.	F. Model	R^2	Pr (>F)
Water Mass	2	0.23985	0.119926	1.4538	0.24417	0.05
Residuals	9	0.74245	0.0842494	--	0.75583	--
Total	11	0.98230	--	--	1.0000	--

DCM: The result was p-value =0.68 (>0.05), with F=0.78486 with 2 to 10 degrees of freedom, see table 16. As a result, H0 was valid. Therefore, water mass at DCM did not influence the distribution of the prokaryotic communities. Only 16% (R^2=0.16403) of the variation of the prokaryotic community was explained by the type of water mass.

Table 16: Output of permanova analysis of Water mass-DCM.

	Df	Sum of Sqs	Mean Sqs.	F. Model	R^2	Pr (>F)
Water Mass	2	0.16761	0.083806	0.78486	0.16403	0.68
Residuals	8	0.85423	0.106778	--	0.83597	--
Total	10	1.02184	--	--	1.0000	--

4.5.2. Venn Diagram Analysis

The Venn diagram for surface samples (fig. 28) showed that 411 ASVs were shared by the three different water masses, corresponding to 6.38% of the total ASVs. The water mass Subtropical had 808 exclusive ASVs (12.54%), being the water mass more exclusive. Front was the water mass with less exclusive ASVs, only 165 (2.56%). The water masses Polar and Subtropical showed higher numbers of e shared ASVs, 123 (1.91%) in total. On the other hand, Front and Subtropical were the water masses with less shared ASVs, only 61 (0.95%).

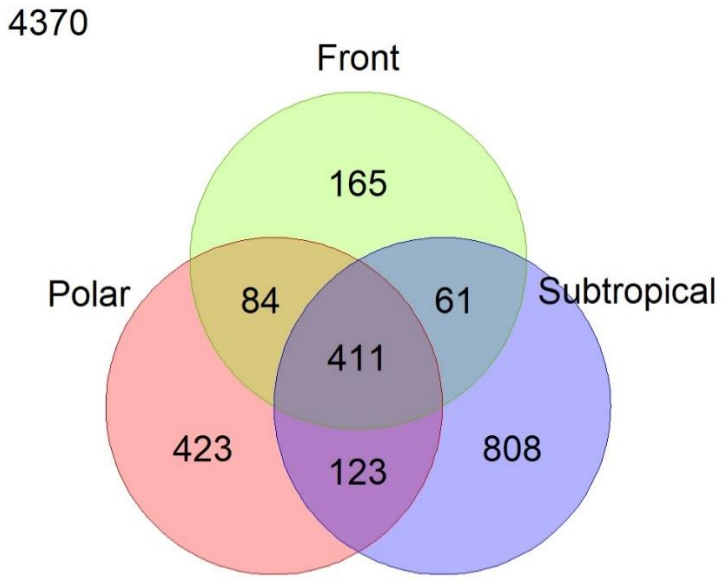


Figure 28: Venn diagram for Surface water masses.

In DCM samples (fig. 29) a total of 937 ASVs were shared between the two different water masses (Front and Polar waters), corresponding to 14.54% of the total ASVs. Front was found to be the least exclusive water mass, with 234 unique ASVs (3.63%), although Polar water presented a total of 656 exclusive ASVs (10.18%).

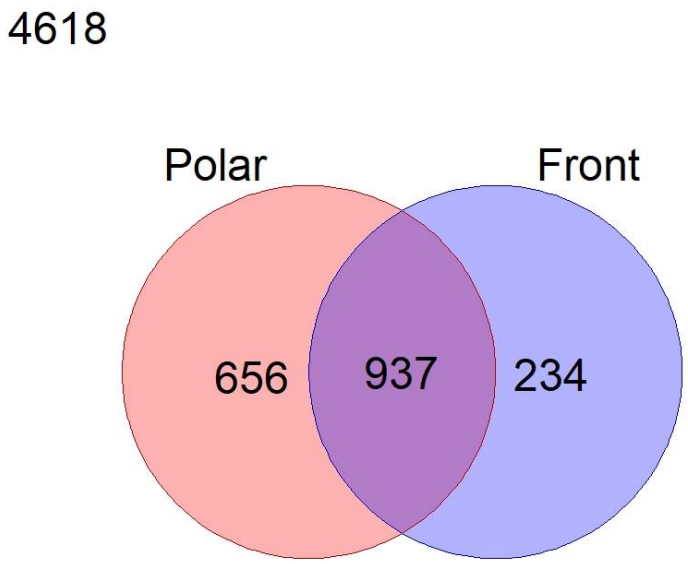


Figure 29: Venn diagram for DCM Water Mass.

In order to explore the distribution of the exclusive ASVs between water masses, plots were performed to compare the values of salinity and temperature with the number of exclusive ASVs. In Surface samples the higher the salinity values, the more exclusive ASVs the Polar and Subtropical samples presented. However, Front samples had an indirect relationship with this parameter: the lower the salinity values, the more exclusive ASVs were presented (fig. 30a). In relation to temperature, the Polar and Subtropical samples that had more exclusive ASVs presented the lowest values of temperature (fig. 30b).

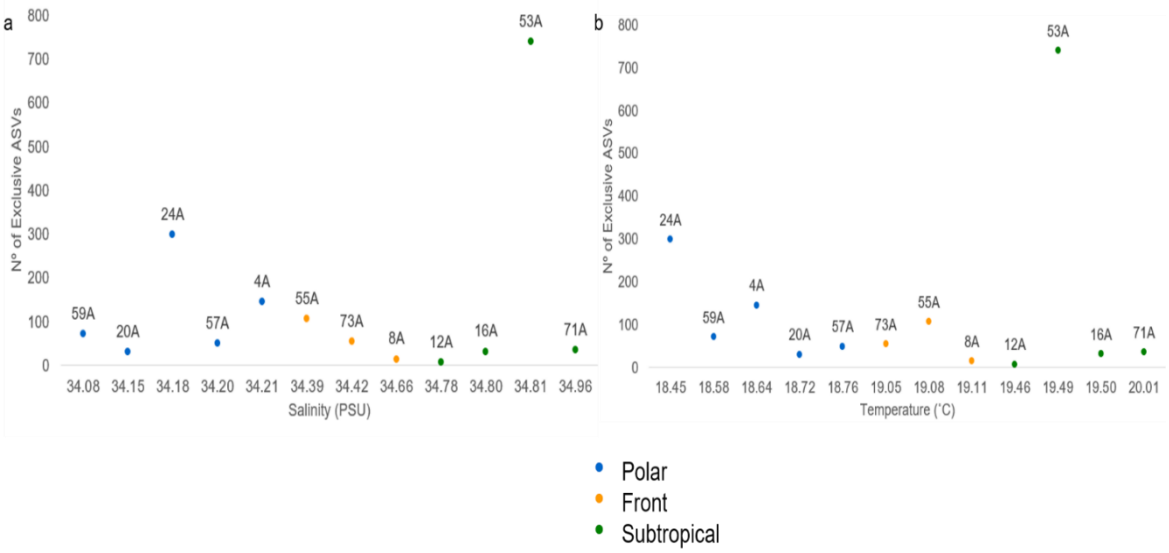


Figure 30: Exclusive ASVs distribution according to salinity (a) and temperature (b) per sample at Surface.

In the DCM samples Polar samples, in general, presented more exclusive ASVs with the increase of salinity. On the other hand, Front samples demonstrated a clear pattern of the number of exclusive ASVs increasing and salinity (fig. 31a). Relatively to temperature, both water masses exhibited an increase in the number of exclusive ASVs with temperature increment (fig. 31b).

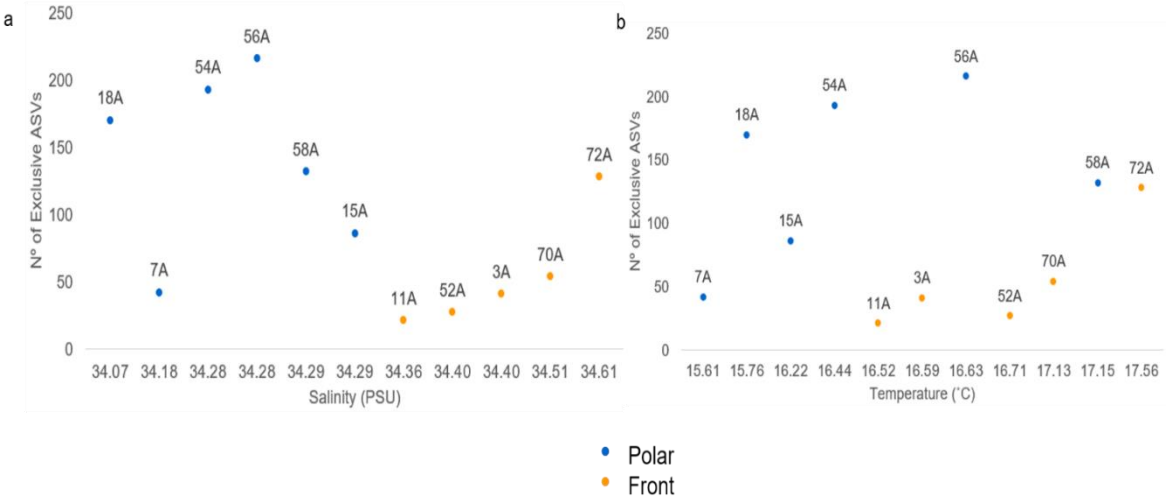


Figure 31: Exclusive ASVs distribution according to salinity (a) and temperature (b) per sample at DCM.

4.5.3. Prokaryotic Taxonomic Analysis

Taxonomic analysis at the phylum level revealed that Surface samples did not show a specific pattern of distribution or a preference for one type of water mass, being the phylum

taxonomic profile very similar between all samples (fig. 32). The most abundant phylum was Proteobacteria, followed by Cyanobacteria. The sample 53A (classified as Subtropical) was the sample that presented more diversity, being the only sample where the phylum Thaumarchaeota was detected.

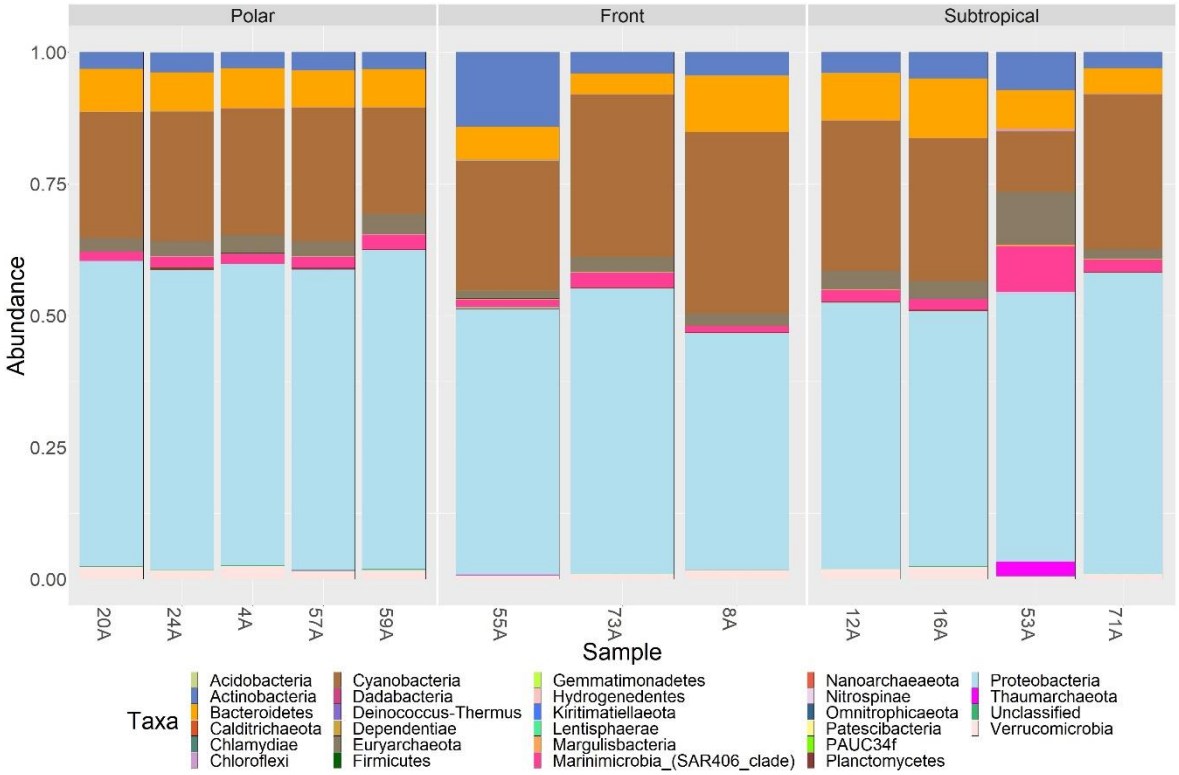


Figure 32: Phyla identification across the samples at Surface.

Phyla distribution at the DCM samples also did not demonstrate a specific pattern of distribution between the water masses. Here Proteobacteria and Cyanobacteria were also the most abundant phyla. Additionally, it was evident an increase of the abundance of Actinobacteria, Euryarchaeota and Marinimicrobia (SAR406 clade) in DCM samples. Thaumarchaeota was also detected in more samples but at low relative abundance.

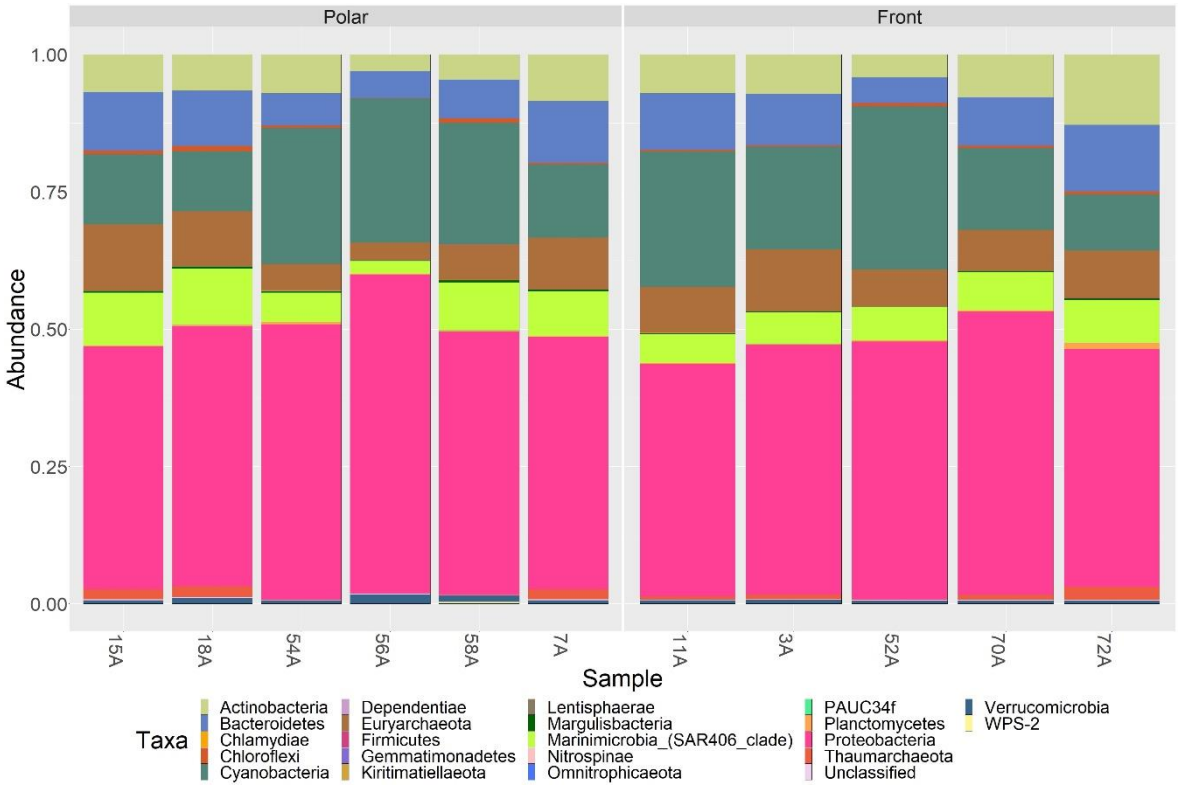


Figure 33: Phyla identification across the samples at DCM.

At Surface samples, the most abundant classes were Alphaproteobacteria (phylum Proteobacteria) and Oxyphotobacteria (phylum Cyanobacteria) without any relevant differences between samples (fig. 34). The sample 53A (classified as Subtropical) was the most diverse.

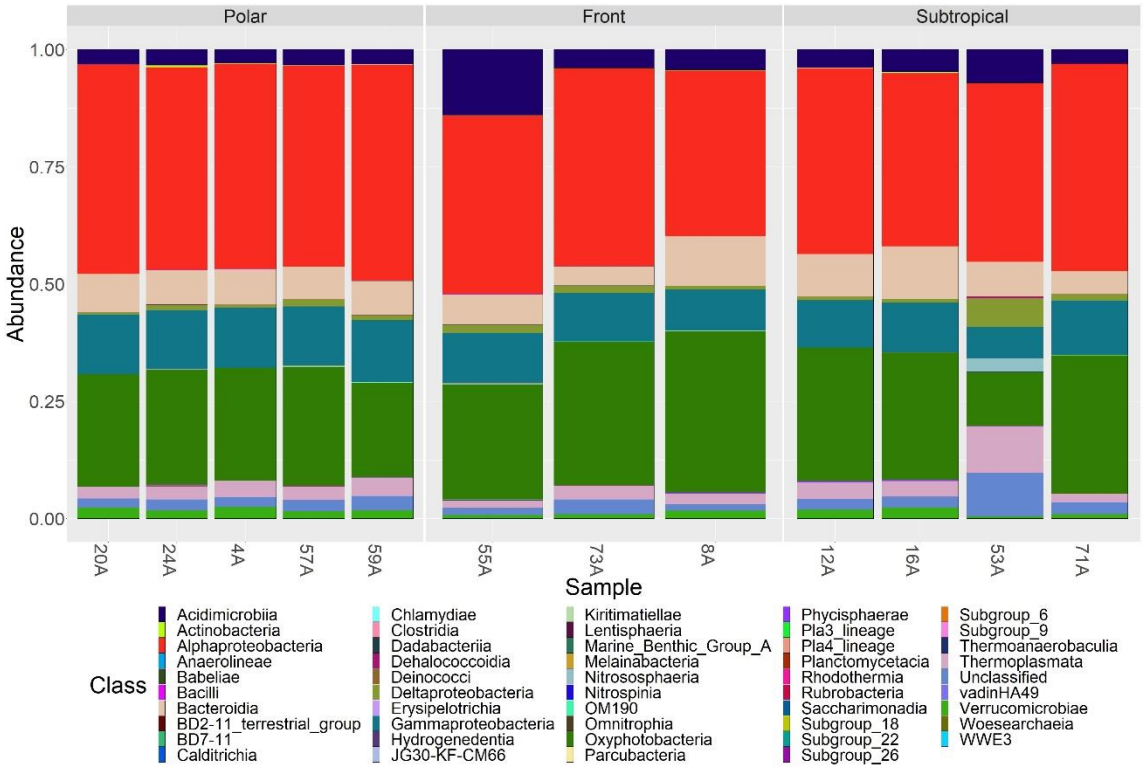


Figure 34: Class identification across the samples at Surface.

At DCM, Alphaproteobacteria (phylum Proteobacteria) and Oxyphotobacteria (phylum Cyanobacteria) dominate in all samples, again without any specific pattern of distribution (fig. 35). The class Thermoplasta (phylum Euryarchaeota) tended to increase in DCM samples compared with surface samples.

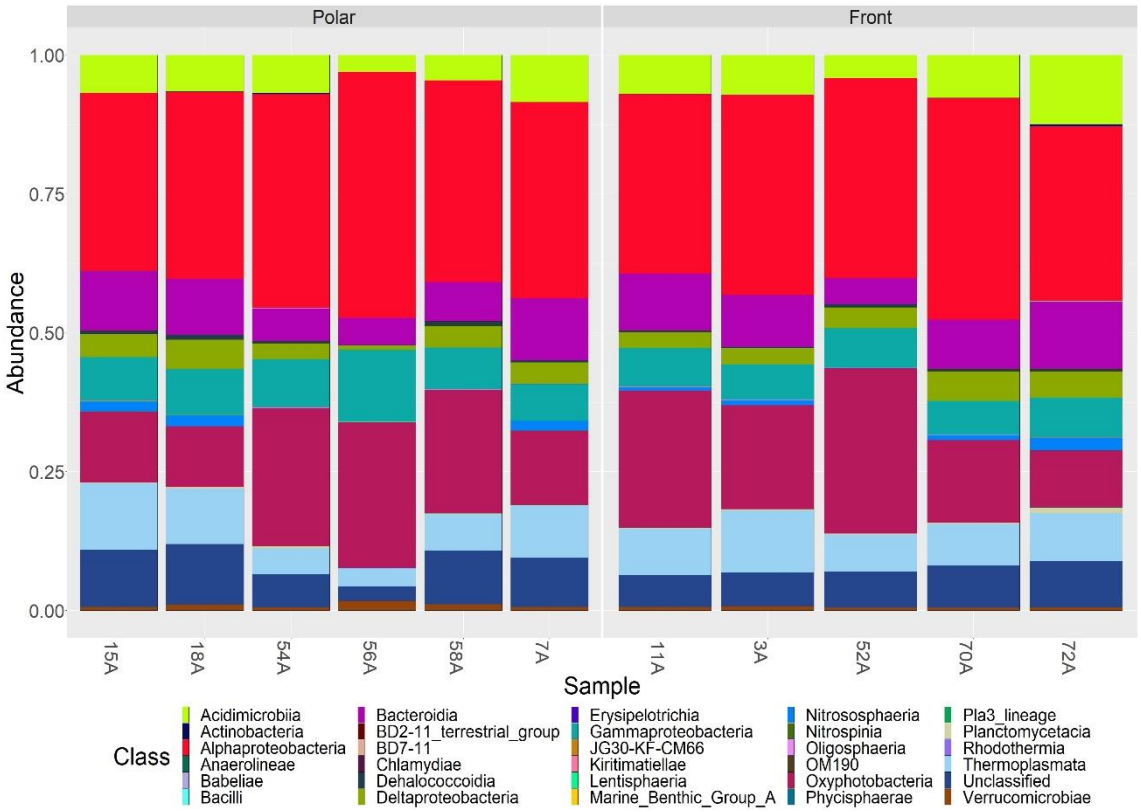


Figure 35: Class identification across the samples at DCM.

In order to understand which genera were common in the different water masses, the following equations were applied:

$$a) \quad \frac{Nº\ of\ shared\ genera\ (without\ NA)}{Total\ genera} * 100$$

$$b) \quad \frac{Nº\ of\ shared\ ASVs}{Total\ ASVs} * 100$$

$$c) \quad \frac{Nº\ of\ sequences\ \ from\ shared\ ASVs}{Total\ sequences} * 100$$

Results from this analysis showed that the most abundant shared genera were Prochlorococcus MIT9313 (phylum Cyanobacteria), followed by Clade Ia, and Clade Ib (phylum Proteobacteria) (fig. 36). At surface, the number of shared genera was relatively low, corresponding to 11.72% between water masses (eq. a). This percentage did not include the classification “NA”. the number of shared ASVs between the three water masses

corresponded to 69.76% of the total ASVs (eq. b). In relation to the shared sequences at surface, these sequences corresponded to 86.65% of the total of the sequences (eq. c) demonstrating minimal difference between the three water masses in terms of genus taxonomic level.

At DCM, even though most of the genera were registered as unclassified, Prochlorococcus MIT9313 was still the genus more abundant. At DCM, the number of shared genera was relatively low, only corresponding to 18.56% (eq. a). This value did not include the classification “NA”. The number of shared ASVs between the two water masses (Polar and Front) corresponded to 85.87% of the total ASVs (eq. b) and a total of 95.65% shared sequences were identified at DCM samples from Front and Polar waters(eq. c), showing also minimal differences between DCM water masses.

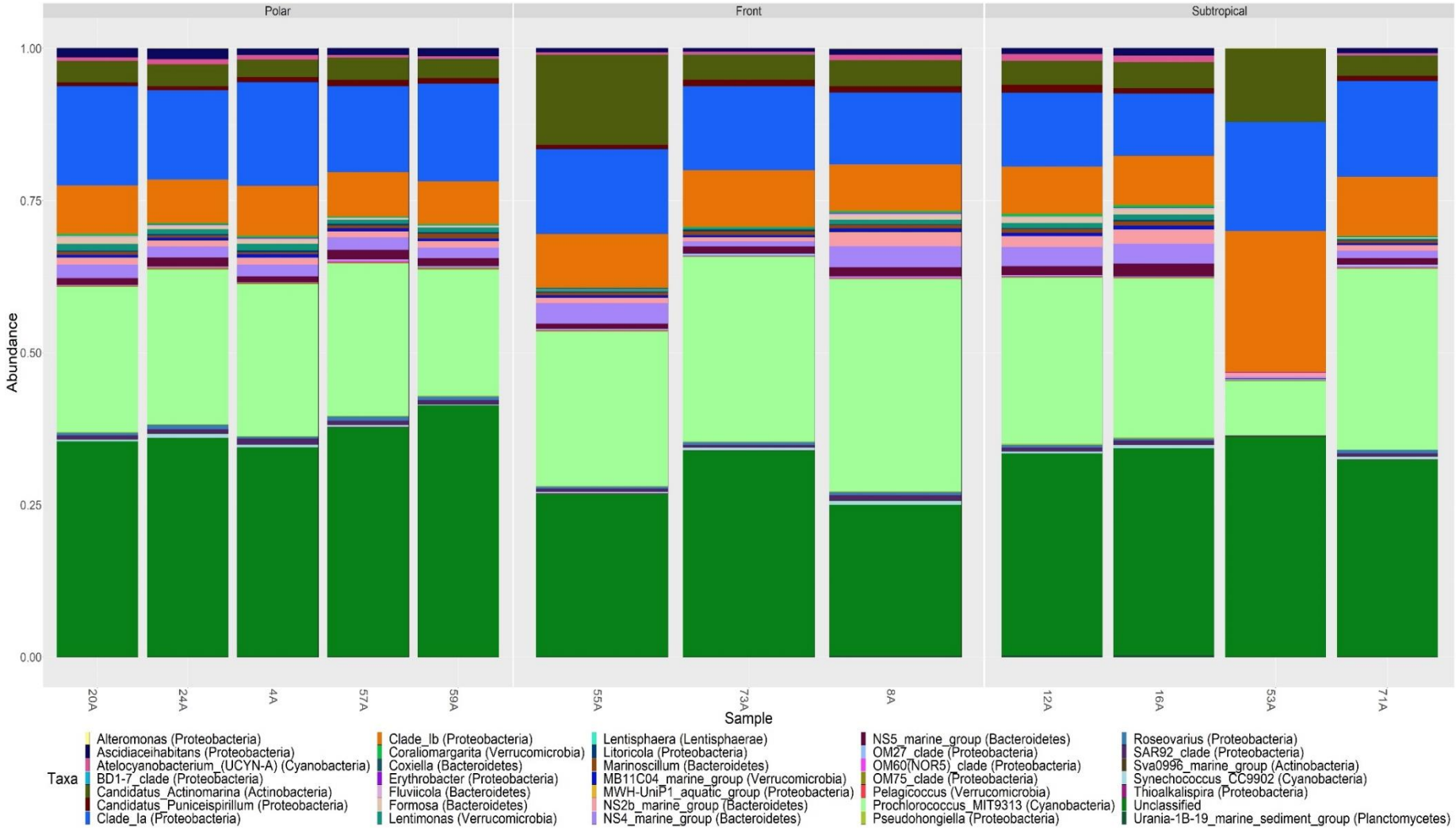


Figure 36: Shared genera between the water masses at Surface samples.

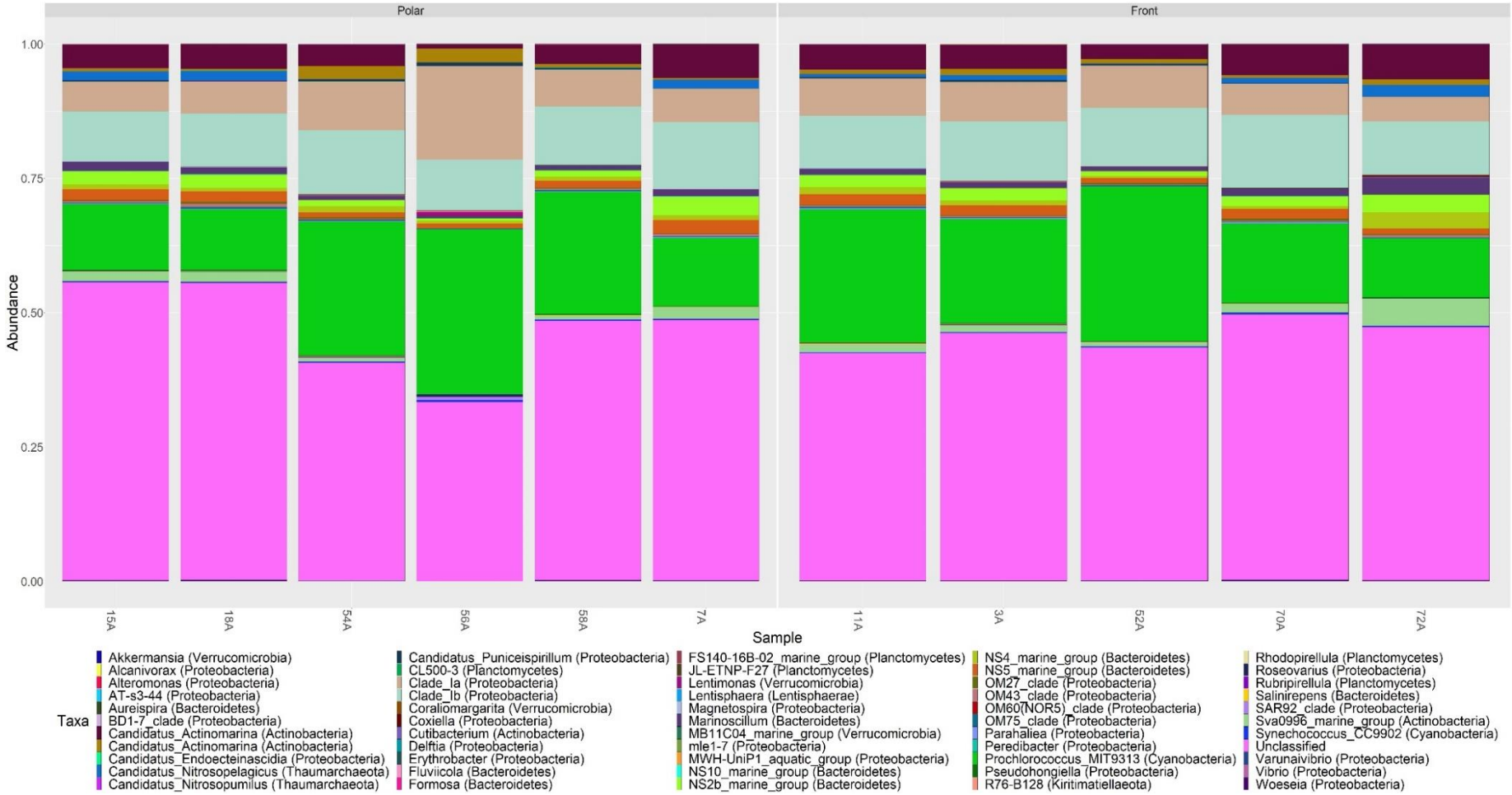


Figure 37: Shared genera between the water masses at DCM samples.

In order to discover which genera were exclusive in the different water masses, the following equations were used:

$$d) \quad \frac{Nº\ of\ exclusive\ genera\ (without\ NA)}{Total\ genera} * 100$$

$$e) \quad \frac{Nº\ of\ exclusive\ ASVs}{Total\ ASVs} * 100$$

$$f) \quad \frac{Nº\ of\ \ sequences\ exclusive\ ASVs}{Total\ sequences} * 100$$

At surface, the number of exclusive genera was higher than the number of shared genera, corresponding to 44.83% (eq. d). This value did not include the classification “NA". The number of exclusive ASVs corresponded to 21.88% of the total ASVs (eq. e). The exclusive sequences represent only 9.69% of the total sequences (eq. f). In comparison to the rest of the water masses, Polar waters presented more diversity in terms of exclusive genera. Contrary, Subtropical showed the least number of exclusive genera (fig. 38).

At DCM, the number of exclusive genera was higher than the number of the shared genera, corresponding to 27.84% of the total genera (eq. d). This value did not include the classification “NA”. The number of exclusive ASVs corresponded to 14.13% of the ASVs total (eq. e). In relation to the exclusive sequences at DCM, it represents only 4.35% of the total sequences (eq. f). Similar to surface, at DCM Polar waters presented higher levels of exclusive genera diversity (fig. 39).

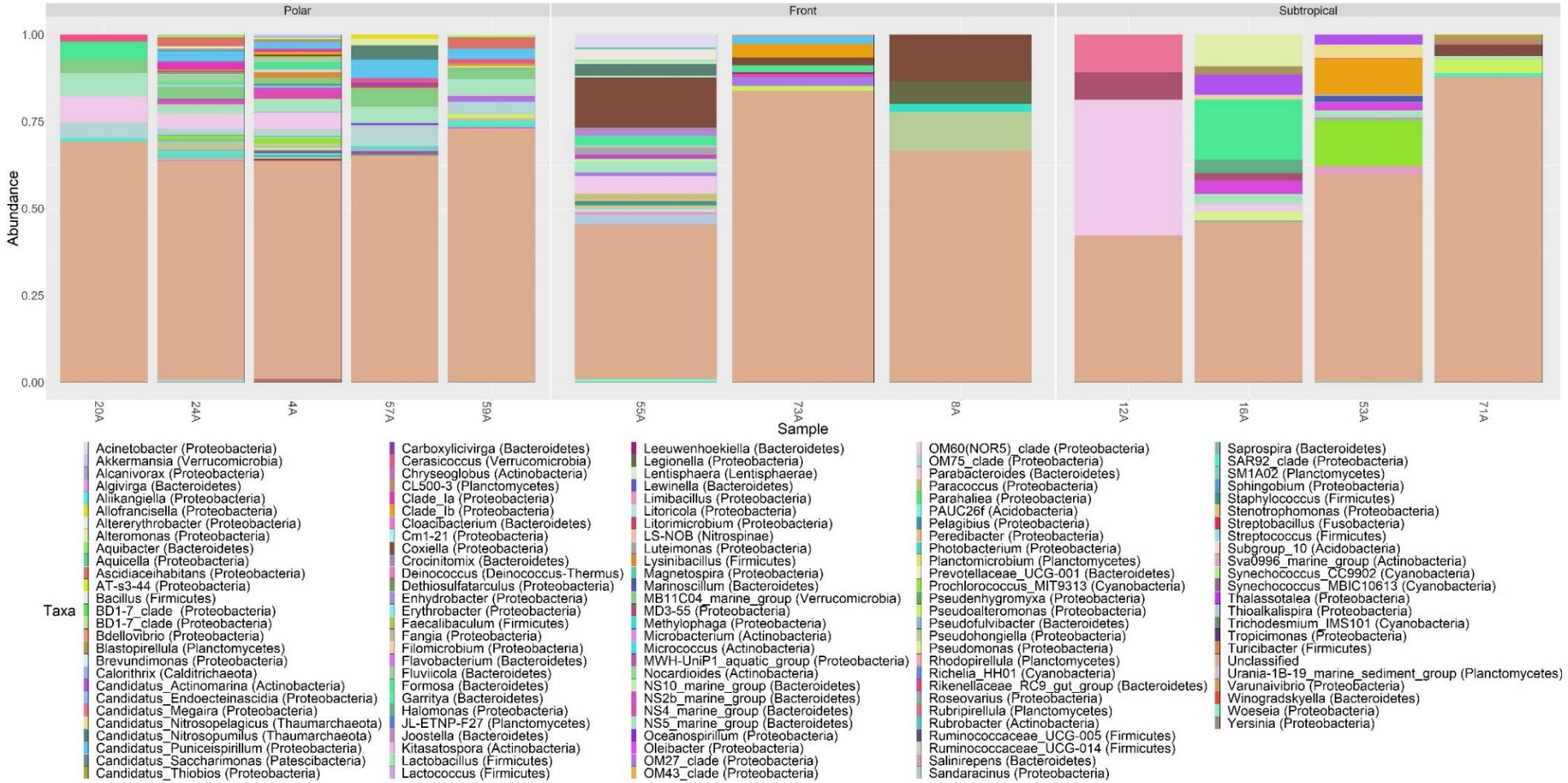


Figure 38: Exclusive genera between the water masses at Surface samples.

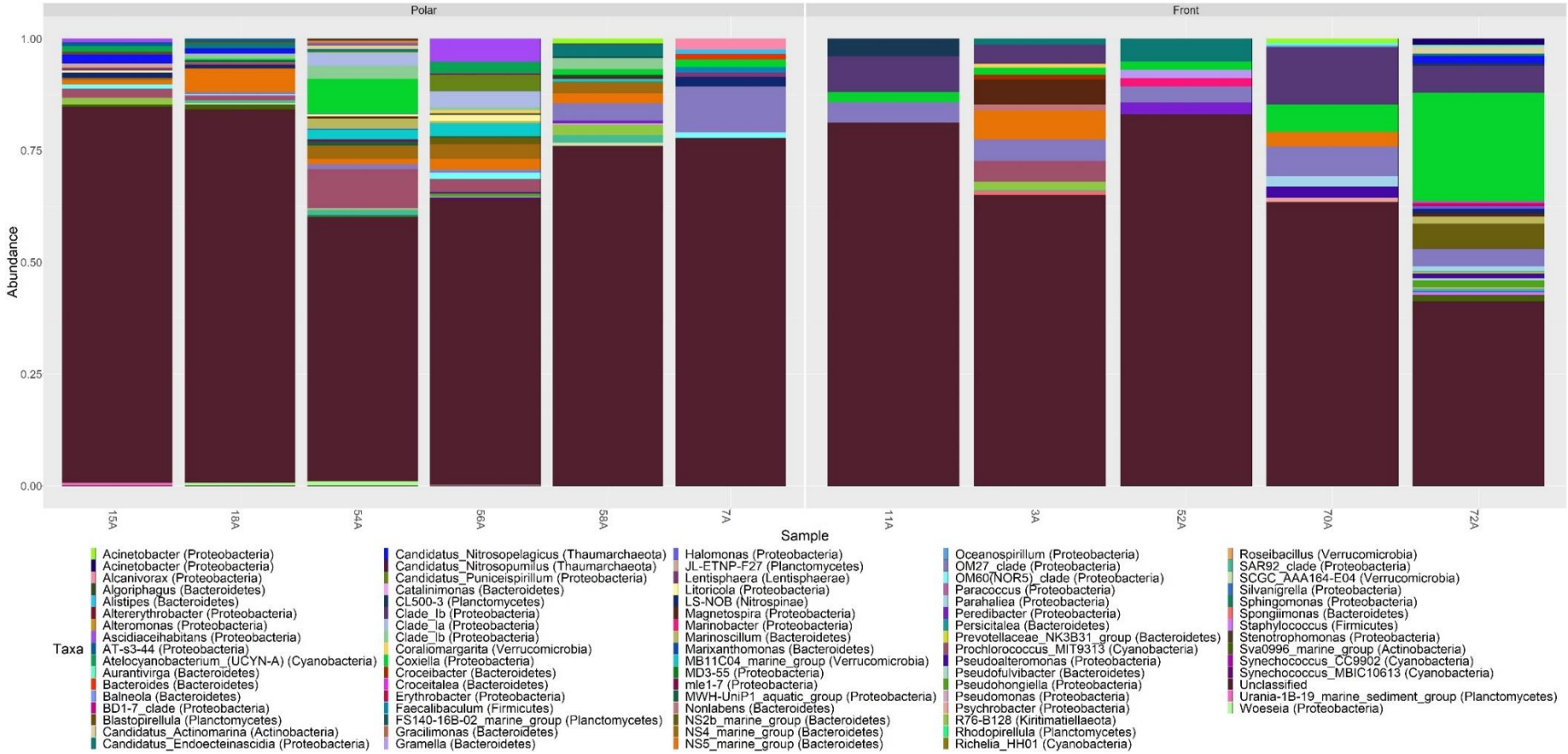


Figure 39: Exclusive genera between the water masses at DCM samples.

4.6. Environmental vs Biological Relationships

In order to explore the relationships between the environmental and biological data a canonical correspondence analysis (CCA) was performed for the ASVs shared between the different water masses and for the ASVs that were exclusive of each water mass.

4.6.1. Shared ASVs

The envfit test (table 17) showed that only oxygen significantly influenced the variance of the surface samples (p≈0.003). Additionally, depth (p≈0.05) was approaching significance. The Polar samples (4A and 20A) and the sample 8A (classified as Front) were positively influenced by these parameters, (fig. 40). The samples 12A and 16A (classified as Subtropical) were minimally influenced by depth and oxygen. Contrary, the samples 71A (classified as Subtropical) and 73A (classified as Front) revealed an inverse relationship with these two parameters. Even though fluorescence was not statistically significant, this parameter demonstrated a strong influence on the sample 71A (classified as Subtropical). The same happens with the other Subtropical samples (12A and 16A) but with turbidity and salinity.

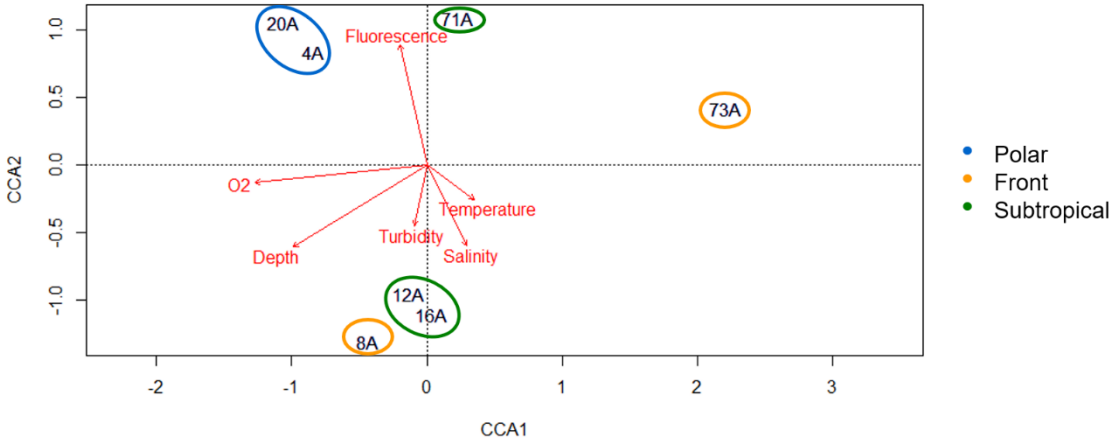


Figure 40: CCA analysis of the shared ASVs at Surface samples.

Table 17: Envfit test results of the shared ASVs at Surface.

Environmental Data	***p-value***	***R^2***
Temperature	0.85015	0.759
Salinity	0.695305	0.1740
Depth	0.046953	0.8145
Oxygen	0.002997	0.9914
Turbidity	0.796204	0.1257
Fluorescence	0.468531	0.3492

The shared ASVs at DCM were not significantly influenced by any of the environmental parameters were included in the analysis (envfit test, p>0.05) (table 18). Although, sample 72A (classified as Front) it seems to be influenced by salinity, temperature, and oxygen and sample 11A (classified as Front) by turbidity and depth (fig. 41).

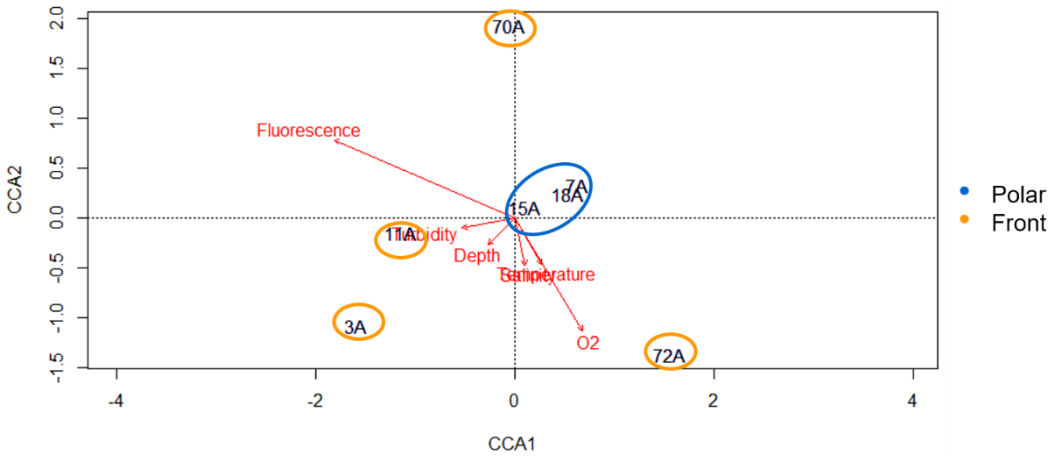


Figure 41: CCA analysis of the shared ASVs at DCM samples.

Table 18: Envfit test results of the shared ASVs at DCM.

Environmental Data	***p-value***	***R^2***
Temperature	0.9071	0.0565
Salinity	0.8741	0.0667
Depth	0.9121	0.0425
Oxygen	0.6913	0.1643
Turbidity	0.952	0.0315

Fluorescence	0.2068	0.5738

4.6.2. Exclusive ASVs

CCA analysis showed that temperature and salinity (p≈0.002) significantly influence the exclusive ASVs distribution. Additionally, with the CCA analysis (fig. 42), it was possible to identify three clusters, which includes samples from each distinct water mass. The Subtropical samples were positively correlated with temperature and salinity (p<0.05). The rest of the samples demonstrated an inverse relationship with these parameters, especially Front samples. Polar samples showed a a major influence on fluorescence.

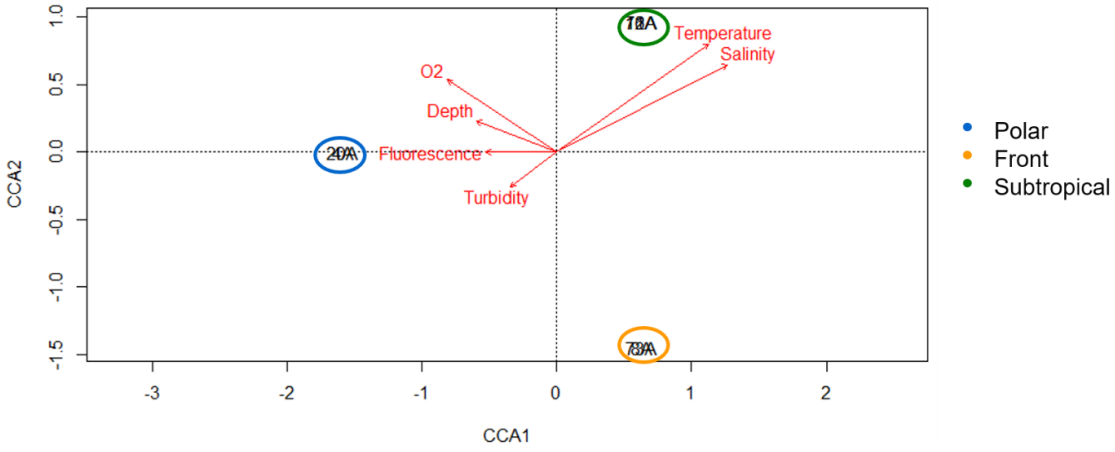


Figure 42: CCA analysis of the exclusive ASVs at Surface samples.

Table 19: Envfit test results of the exclusive ASVs at Surface.

Environmental Data	***p-value***	***R^2***
Temperature	0.001998	0.8633
Salinity	0.001998	0.9141
Depth	1	0.1833
Oxygen	0.34965	0.4263
Turbidity	0.864136	0.0825
Fluorescence	0.96014	0.1250

Contrary to the shared ASVs, envfit test performed for the exclusive ASVs at DCM revealed that the majority of the environmental parameters were statistically significant with the

sample's distribution: temperature (p≈0.005), salinity (p≈0.006), turbidity (p≈0.04) and fluorescence (p≈0.02) (table 20). Polar samples revealed being influenced by oxygen, depth, and turbidity and samples 72A (also classified as Front) showed a positive correlation with salinity and temperature (fig. 43).

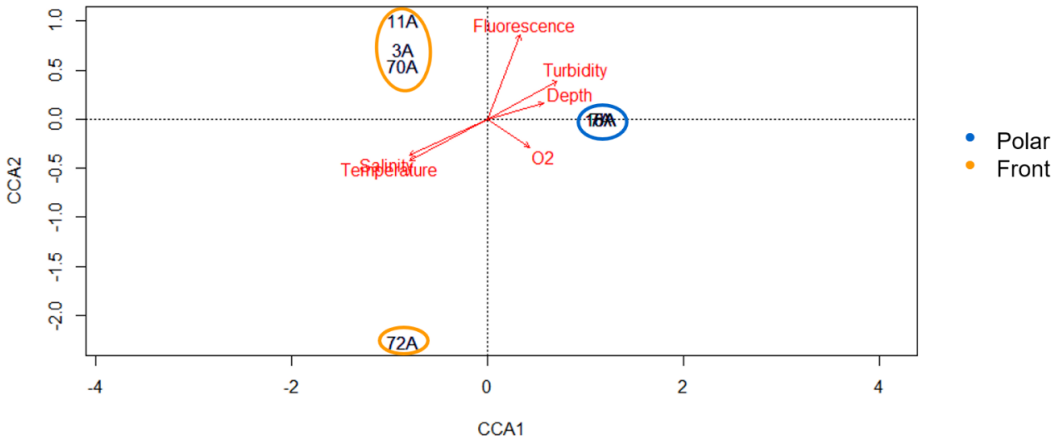


Figure 43: CCA analysis of the exclusive ASVs at DCM samples.

Table 20: Envfit test results of the exclusive ASVs at DCM.

Environmental Data	***p-value***	***R^2***
Temperature	0.004995	0.8882
Salinity	0.005994	0.8522
Depth	0.34965	0.3921
Oxygen	0.555445	0.2697
Turbidity	0.043956	0.7397
Fluorescence	0.020979	0.9439

5. Discussion

5.1. Prokaryotic communities distribution within water column depth

Prokaryotes distribution and structure across the water column are dependent on several physical factors, resulting on the well-known phenomenon of vertical stratification (Martiny et al., 2006; Follows et al., 2007). One of the factors responsible for this stratification and the changes of the prokaryotes community is water column depth (Giovannoni et al., 1996; Giovannoni & Stingl, 2005; DeLong et al., 2006). The results obtained on this study demonstrated a clear environmental depth stratification that characterized the region of the North Pacific Subtropical Front with a clear influence on the prokaryotic communities distribution. These results were coherent with Celussi et al. (2009), Li et al. (2019) and Tobias-Hünefeldt et al. (2019). For example, Celussi et al. (2009), while studying the vertical distribution of the prokaryotes in the Ross Sea reported that epipelagic communities were distinct from the mesopelagic communities. Li et al. (2019) and Tobias-Hünefeldt et al. (2019) also stated that depth differences originated in different communities.

Depth influences physical-chemical parameters, such as light presence, temperature, and nutrient availability (Giovannoni & Stingl, 2005; DeLong et al., 2006). With depth, temperature decreases along with the presence of light, becoming a gradual colder and darker environment (Latasa et al., 2017). These environmental differences influence the development of two major prokaryotic communities adapted to euphotic and aphotic zones. In this study we showed that in the euphotic zone the communities were mainly constituted by phototrophic organisms, such as organisms belonging to the phylum Cyanobacteria. This distribution pattern is similar to the pattern described in Giovannoni & Stingl (2005).

In the aphotic zone the microorganisms present are primarily chemotrophs, such as organisms belonging to the phylum Thaumarchaeota, similar to what was found by Herndl et al. (2008) and Hansman et al. (2009). Originally classified as Crenarchaeota, the phylum Thaumarchaeota is also known as ammonia-oxidizing archaea (AOA) (Pester et al., 2011). This phylum is particularly abundant in oligotrophic oceans and is especially found in the upper mesopelagic layer and in oxygen minimum zones (Karner et al., 2001; Herndl et al., 2005; Coolen et al., 2007). Marine Thaumarchaeota have a crucial role in the biogeochemical dynamics across the water column (Pester et al., 2011). They participate in the remineralization of nitrogen, nitrous oxide production and carbon fixation, which makes these organisms essential to the cycles of carbon and nitrogen (Stewart et al., 2012; Reji et al., 2019). By being chemolithoautotrophs, Thaumarchaeota organisms are able to

fix carbon dioxide (CO_2), with the energy that they get from oxidizing ammonia (NH_3) to nitrite (NH_2^-) (Könneke et al., 2005; Stieglmeier et al., 2014).

Being a phylum designated as a phototroph, Cyanobacteria was limited to the photic zone. The distribution pattern of this phylum was similar with other studies, showing that its distribution is heavily dependent on light availability (Giovannoni & Stingl, 2005; Cantonati et al., 2014; Hamilton et al., 2016). Similar to Schattenhofer et al. (2009), the phylum Bacteroidetes was identified in the five depths analysed but with a higher prevalence in shallower depths.

Cyanobacteria demonstrated an inverse relationship with a Thaumarchaeota. This finding was also observed by Hernández-Avilés et al. (2018). Thaumarchaeota was most abundance in the deeper layers, as previously found (Fernandes et al. 2020) and is designated as a major contributor to the nitrogen cycle including mainly ammonia-oxidizing organisms (Francis et al., 2007; Mincer et al., 2007). Planctomycetes were also recorded at deeper depths and this phylum is known to perform relevant transformations in the global carbon and nitrogen cycles (Glöckner et al., 1999). This increase of Planctomycetes with depth can be due to its metabolic functions and environmental conditions (for example, carbon presence), which lead to the Planctomycetes contribution to the chemolithotrophic growth that involves iron and methane oxidation (Storesund & Øvreås, 2013). The phylum Cyanobacteria is one of the most known and important phototrophic bacteria, presenting a great diversity and immense ecological importance. In the present day, five orders, 150 genera, and 2000 species are acknowledged in this phylum. Cyanobacteria can be found in several habitats, such as lakes, ponds, wetlands, and rivers. Additionally, this phylum plays a significant role in the nitrogen, carbon, and oxygen cycles of aquatic environments. By being present in a wide range of several different habitats, Cyanobacteria is very well adapted to survive in distinct environmental conditions. Where there is light, this phylum can be found, making Cyanobacteria extremely tolerant to UV radiation. In terms of temperature, Cyanobacteria is very adaptable, enduring great ranges of temperature values (Garcia-Pichel, 2009; Vicent, 2009).

5.1.1. Environmental Controls on Depth Prokaryotic Communities Distribution

Cyanobacteria was one of the most abundant phyla identified in samples from the upper layers, dominated by the genera Prochlorococcus and Synechococcus. One critical factor that controls the distribution of these genera is the presence of light. The genus Synechococcus is limited to the euphotic zone while the genus Prochlorococcus is capable

of growing outside this zone, at greater depths, such as 150m or more (Partensky et al., 1999). The presence of Prochlorococcus at greater depths was also observed in our study, while Synechococcus was practically absent at Below DCM and Mesopelagic zones. Additionally, the nitrogen (N) availability also influences the distribution of these two genera. Synechococcus generates a significant amount of energy by reducing NO to NH3 in the presence of light, in order to balance the high cellular N required for their N-rich phycobilisomes (Carr & Mann, 1994; Moore et al., 2002). Partensky et al. (1999) stated also that Synechococcus tends to grow more in nutrient rich waters whereas Prochlorococcus proliferate more in waters where nutrients are minimal. Prochlorococcus also tends to proliferate in more stable waters, while Synechococcus is able to tolerate more unstable mixed water masses (Lindell & Post, 1995).

Regarding the phylum Proteobacteria, the classes Alpha-, Delta-, Gammaproteobacteria presented a different distribution across the water column. Even though Alphaproteobacteria was the most abundant class across the water column, its abundance suffered a decrease in depth. This pattern was also observed in Agogué et al. (2011). Consequently, the most abundant order from this class, SAR11 clade, also presented the same distribution pattern. These organisms present in the euphotic zone are classified as phototrophs and heterotrophs (Han et al., 2015). Thus, Han et al. (2015) introduced the "sinking bacterial hypothesis" to explain the abundance decrease with depth of these organisms. In the deeper layers, the marine environment becomes rougher. Leading to a decrease in temperature and nutrients and does not have the conditions for photosynthesis to be realized. SAR11 clade organisms, and consequently Alphaproteobacteria organisms, are not able to proliferate in these conditions. Contrary to Alphaproteobacteria, Deltaproteobacteria abundance increased with depth. Similar results were also found in Bano & Hollibaugh (2002) and Zaballos et al. (2006). These organisms are classified as chemotrophs. As a result, Deltaproteobacteria take part in the carbon fixation via the cycle of Calvin and in the oxidation of sulfur and methylated compounds (Swan et al., 2011). Gammaproteobacteria abundance is stated to increase with depth (Sogin et al., 2006; Lauro & Bartlett, 2008). However, in our study Gammaproteobacteria abundance was relatively stable across the water column. Yilmaz et al. (2016) also presented a similar distribution pattern. On one hand, the presence of Gammaproteobacteria at higher depths can be justified with the sharply decrease of SAR11 clade, and consequently, Alphaproteobacteria organisms (López-Garcia et al., 2001; Zaballos et al., 2006). Additionally, some genera which belong to this class, for example the genus SUP05, are highly adapted to survive in dysoxic and anoxic waters, as well as to the lower mesopelagic zone (Glaubitz et al., 2009; Lavik et al., 2009; Brown et al., 2014). On the other hand, some orders of this class, for

example SAR86 clade and Oceanospirillales, are frequently found at shallower depths. It is thought that these organisms have similar distributions patterns of the SAR11 clade (Yilmaz et al., 2016). Additionally, these organisms are often associated with phytoplankton blooms (Teeling et al., 2012; Buchan et al., 2014).

The phylum Proteobacteria is considered the largest phylum of bacteria and one of the most diverse, with 460 genera and 1600 species, at least (Kersters et al., 2006; Moon et al., 2018). This phylum is also mentioned as “purple bacteria and relatives” and the organisms that belong to this phylum, in terms of obtaining energy, are classified as phototrophs, heterotrophs as well as chemolithotrophs (Stackebrandt et al., 1988; Pfenning et al., 1997; Schütz et al, 1998; Gupta, 2000; Imhoff et al., 2005). Besides being a diverse phylum, Proteobacteria is also one of the most abundant phyla, it is present from pole to pole in the oceans and can be found across all the water column (Zehr et al., 2000; Huber et al., 2007; Hauptmann et al., 2014; Delmont et al., 2018). At surface, phototrophs and heterotrophs organisms are present due to the existence of dissolved organic matter (González et al., 2000). At higher depths, Proteobacteria is associated with the cycle of carbon, by primary production related with sulphur and methane oxidation, along with heterotrophy (Aristegui et al., 2009; Swan et al., 2011).

5.1.2. Below DCM as a Hotspot of Biodiversity

In this work, prokaryotes diversity increased with depth, with a peak just below the DCM. Data from other analyses were consistent with our results. Meaning, the diversity increases with depth. Some examples are the studies from Treusch et al. (2009), Danovaro et al. (2010), Wilson et al. (2017), Li et al. (2019) and Fernandes et al. (2020). The most common justification for this result was the increase of the stability with depth. Namely, temperature and radiation levels do not vary as much as on the surface (Treusch et al., 2009; Pommier et al., 2010). In addition, Wilson et al. (2017) stated that the reasons for these differences in diversity profiles are related with metabolic reactions. At higher depths, these reactions are based exclusively on chemical redox and are more varied in time, space and in function. These reactions by being more diverse in deeper areas, more diverse organisms are needed to include all these functions.

There are cases, however, where the highest diversity values were recorded at shallower depths. In a soil analysis, the prokaryotes diversity reached its maximum values at surface (Liang et al., 2019). This result was not exclusive to the soils, being found also in the marine environment. A study performed in the North Atlantic Ocean justified the higher diversity at

the photic layer due to the existence of phototrophic organisms and considerable changes in the environmental parameters above the pycnocline (Hahnke et al., 2013). In agreement with these findings, Brown et al. (2009) referred that the existence of radiation for photosynthesis contributes to larger amounts of dissolved organic matter in the euphotic zone.

5.1.3. Connectivity between the different water column depths

Interestingly, in the study area, we found that the number of shared ASVs between the five depths was relatively low. Yet, other analyses demonstrated the same results of having a low number of shared ASVs in a relatively small area. For example, Rasheeda et al. (2017) with the aim to know the microbiota diversity in three different fish species in a fish farm also detected a low number of OTUs of prokaryotes shared by the three fish species. Additionally, an analysis performed in rainbow trout individuals exhibited similar results (Mente et al., 2018). One of the goals of this analysis was to prove that the microbiota structure was influenced by the diet composition of the trout. This study revealed that indeed the prokaryotes present in trout gut and feces were affected by the food that the animals ate (Mente et al., 2018). Besides diet composition, several parameters are known to affect the prokaryotes communities in fish gut. Such as the genetics of the host, the weight of the host and what cycle of life the host is (Tetlock et al., 2012; Smith et al., 2015; Sherril-Mix et al., 2018). The environmental conditions also impact the microbiota composition (Sullam et al., 2012; Parris et al., 2016). For instance, salinity is identified as a crucial factor for defining the fish microbiota (Sullam et al., 2012; Zhang et al., 2016).

Differences in environmental parameters also influenced the ASV exclusivity, and consequently, differences in the prokaryotes community in this study. Lozuponde & Knight (2007) stated that salinity was the major environmental factor for the prokaryotes community structure across the water column. In our study, salinity varied from 34.8 PSU at the surface to 34.1 PSU in the mesopelagic, which may have contributed to the observed habitat partitioning. Additionally, salinity has a direct relationship with water density, which in its turn has an important role in defining the water masses and their prokaryotic community. This environmental factor is also responsible for the water column vertical stratification and also for the prokaryotes communities structure (Riemman et al., 1999; Ghiglione et al., 2008).

Depth is also responsible for changes in temperature, nutrients, and light entrance in the water column. These environmental factors also influence the composition and structure of prokaryotes communities (Fortunato et al., 2012). These organisms have to be prepared for the rapid transformations that happen in the environment that they inhabit, like high variations of temperature and salinity values (Deming & Collins, 2016; Rapp et al., 2018). By not being adapted to different conditions across the water column, the organisms cannot survive in it. This justifies the low number of shared ASVs between the five different depths, which represents five highly unique habitats.

5.2. Prokaryotic communities distribution across the different water masses at Surface and DCM

As previously explained, prokaryotes communities' distribution in the North Subtropical Pacific Ocean Front is influenced by depth, due to its strong vertical stratification, also in agreement with previous studies (Giovannoni & Stingl, 2005). However, the horizontal gradient of ocean fronts caused by the confluence of different water masses can also be responsible for the distribution of the prokaryotic communities. In the area of study, the Polar waters (colder waters with lower salinity) meet the subtropical waters (hotter waters with higher salinity) of the Pacific Ocean (Polovina et al., 2001). Our results on the horizontal prokaryotic communities distribution showed very similar communities between the different water masses at surface and DCM depths, resulting in a non-specific distribution of these organisms through the North Pacific Subtropical Front. Despite these results, several studies determined that water masses were a factor that indeed influenced the distribution of the organisms through the water column (e.g.: Giebel et al., 2009; Celussi et al., 2010 and Agogué et al., 2011).

All water masses have their particular identification and singularity. This identification is done by analysing the distinct physicochemical properties of the water mass, for example, salinity. Productivity and the quantity of organic matter that is present in the water mass can also be taken into consideration. With these specific characteristics, it is understandable to predict that specific prokaryotic communities are found in a specific type of water mass (Arístegui et al., 2009; Celussi et al., 2010). One aspect that influences the specific distribution of prokaryotes communities is the variation of the physicochemical parameters between the water masses. Temperature variations are commonly pointed as the physicochemical gradient that influences the most the prokaryotes distribution (Yu et al., 2015). Besides temperature, salinity is another physicochemical parameter that has a strong influence on the water mass characteristics and on the prokaryotes communities

which inhabit it. Together, these two parameters determine the density of the water. Consequently, these gradients create a barrier in the water column. Consequently, these different density values separate the different types of water masses and a robust oceanographic barrier is created within the same habitat. Subsequently, the different environmental conditions in the water masses can reduce the capacity of the mobility of the organisms between the water masses (de Vargas et al., 1999; Pommier et al., 2007; Agogué et al., 2011; Arístegui et al., 2009; Galand et al., 2009b). As a result, different communities of prokaryotes are often found. In our study, the salinity and temperature gradients across Polar, subtropical waters and the front area are not very sharp, in comparison with other studied fronts (Pommier et al. 2007; Yu et al. 2015). Thus, the oceanographic barrier, in terms of water characteristics, in the study area could not be as resistant as in other fronts, resulting in similar prokaryotes communities at the taxonomic levels evaluated. Additionally, Galand et al. (2010) presented two theories to explain the distribution of the prokaryotic communities between the water masses. The first theory is "barriers to dispersal" where water density creates an oceanographic barrier, as earlier explained (Spencer-Cervato & Thierstein, 1997). The second theory is "high dispersal" that assumes that the organisms move between the different types of water masses without restrictions. However, the maintenance of the community only happens in beneficial areas (Darling et al., 2004; Sexton & Norris, 2008). The second theory may help explain the results observed in this study, microorganisms dispersed freely and were tolerant to the relatively small physicochemical variations between the water masses.

Even though the prokaryotic communities were very similar between the water masses, we found that surface samples presented some prokaryotic communities differentiation. This could be explained by the fact that the surface is a more dynamic area, being more influenced by the environmental pressure and the organisms spend less time at this depth. While in deeper areas the organisms are not subjected to such strong environmental pressure, being a more stable area and the communities spend more time at this depth (Treusch et al., 2009; Ghiglione et al., 2012; Yu et al., 2015).

The North Pacific Subtropical Front is located in the Pacific Ocean, as the name suggests. This ocean is designated as the biggest ocean in the world. Its largeness is displayed in big spatial differences in the mixed layer depth, thermocline depth, upwelling velocity, and other physical features (Longhurst, 2007). Like others analysis, this thesis aims to comprehend the dynamics of the North Pacific Subtropical Front, given that this area is essential for several species, which use NPSF as a migration route and a local to nourish, as well as for fisheries (Polovina et al., 2001; Bograd et al., 2004).

This study area corresponded to a small part of the NPSF, located between 28ºN and 40ºN. This front divides the low chlorophyll subtropical gyres and the high chlorophyll subarctic gyres in the Pacific Ocean (Polovina et al., 2001). NPSF generally reaches its maximum strength in the late winter (January) and in the early spring (March) (Bograd et al., 2004). Nevertheless, the sample collection occurred in the spring (beginning of June) and the front was located between 28ºN and 29ºN. However, according to Chai et al. (2003), in that time of year, NPSF reaches its maximum strength (0.2/0.3mg/m3) between 34ºN-35ºN. The distance of the study area to the local front maximum strength and time of the sample collection could be a cause for the environmental gradients, and consequently, the water masses, not as strong as expected due to the lack of strength of the front. As earlier stated, temperature, salinity and density influence the efficiency of the water masses and consequently, the prokaryotic communities which inhabit it (Yu et al., 2015; Wilson et al., 2017). By not being robust oceanic barriers, the composition of the prokaryotic communities was very similar, in this analysis, between the water masses, especially in the higher taxonomic levels.

5.2.1. Prokaryotic diversity across the different water masses

In this study, the water masses did not present significant differences of species richness and on the diversity indexes evaluated, neither at the Surface or DCM. Alves Junior et al (2014) also had identical results regarding the similarity of prokaryotes communities. The physiochemical gradients of the area analysed were not strong enough to create a robust oceanic barrier and it was suggested, from a microbial ecology point of view, that the water masses in question could be designated as just one water mass. Celussi et al. (2010) also stated that water masses were not as different in terms of phylogeny, and the phylum Proteobacteria was the most abundant in the different water masses. Proteobacteria was the most dominant phylum in this study as well, and it was possible that this cosmopolitan phylum could overshadow the rest of the organisms.

On the other hand, several studies presented results in which the water masses had distinct prokaryotes communities and distinct alpha diversity values (e.g.: Teira et al., 2006; Galand et al., 2009a; Agogué et al., 2011; Han et al., 2015; Techtmann et al., 2015; Zoccarato et al., 2016; Bauer et al., 2018; Fu et al., 2019; Qin et al., 2019). These distinctions in the communities are mainly justified by sharp gradients of environmental factors, especially salinity (Hamilton et al., 2008; Varela et al., 2008; West et al., 2016; Zoccarato et al., 2016; Du et al., 2019), which did not vary much in our study area. Another environmental factor responsible for the prokaryotes communities dynamics is the temperature (Varela et al.,

2008; Zoccarato et al., 2016; Fu et al., 2019), which also did not change much between water masses in the North Pacific Subtropical Front. Even though it is stated that diversity increases with temperature (Furhman et al., 2008).

Besides the lack of differences in community composition and diversity observed in this study between the three water masses, the number of exclusive ASVs was also relatively low for each water mass. Additionally, environmental factors can also influence the presence and distribution of exclusive ASVs in the different water masses. With the CCA analysis, it was possible to observe that salinity and temperature influence the distribution of the exclusive ASVs within the water masses at both depths. As stated previously, salinity and temperature, and consequently density, can act as an oceanic barrier for the prokaryotes and inhibit their movements (Aristegui et al., 2009). But the water masses identified in this analysis were not as resistant and, therefore, the oceanic barrier was not as efficient, explaining the lower exclusive ASVs within Polar waters, Subtropical waters, and Front Waters.

5. Conclusions and Future Perspectives

The North Pacific Subtropical Front is a relevant marine ecosystem, but its dynamics and biological composition is still poorly understood. The results present in this thesis constitute an important dataset for future analysis in this area.

The use of next-generation sequencing demonstrated to be a relevant technique to analyse the composition, abundance, and diversity of the prokaryotic communities, leading to the gain of knowledge about the microbiomes present in the NPSF.

Regarding the depth effect, four distinct microbiomes were identified across the water column and an evident vertical stratification of the water column was observed. The depth factor strongly influenced the distribution of the prokaryotic communities, as temperature, salinity, oxygen and nutrient availability sharply changed with depth. Proteobacteria was the most abundant phylum. Additionally, some phyla demonstrated a preference for surface environments, for example, Cyanobacteria, and other phyla demonstrated a preference for deeper environments, for example, Thaumarchaeota. With this analysis, it was also possible to show that diversity increased with depth at the NPSF, supporting the theory that more stable waters promote the occurrence of highly diverse prokaryotic communities.

Regarding the water mass effect, the physicochemical gradients, especially temperature and salinity, did not differ much between the different types of water masses. Consequently, the water masses did not act as robust oceanic barriers. By not being efficient oceanic barriers, the water masses presented similar microbiomes between themselves. In addition, the majority of the ASVs were shared between the water masses which prevented the diversity to change within water masses.

Finally, this thesis demonstrated that the prokaryotic communities in the North Pacific Subtropical Front were strongly influenced by the vertical gradients, within 500 m depth and suffered little influence regarding the horizontal gradients in the North Pacific Subtropical Front.

In terms of future perspectives, it would be interesting to investigate if the prokaryotic communities responded in the same way in a different time of the year and to verify if the

eukaryotic communities would demonstrate the same or different pattern of distribution to the same effects analysed on this study.

6. Scientific Communications (based on the work of this thesis)

Lopes, E., Semedo, M., Tomasino, M. P., Magalhães, C. 2020. Prokaryotic Dynamics Across the North Pacific Subtropical Front. (ref. 16756) In: Book of Abstracts of the 13th Meeting of Young Researchers at U. Porto (IJUP013). 12-14/02/2020, Porto, Portugal. ISBN: 978-989-746-253-5 (Poster communication)

Lopes, E., Semedo, M., Tomasino, M. P., Magalhães, C. 2020. Prokaryotic Dynamics Across the North Pacific Subtropical Front. In Book of Abstracts of the 1st Blue Think Conference 2020: Share Science, Spread Knowledge. 17-09-2020, Matosinhos, Portugal. ISBN: 978-989-54965-0-1 (Poster communication)

The book of abstracts of “FEMS Online Conference on Microbiology 2020” is not available yet but part of this work will be presented at this conference by a poster presentation, with the title “Depth Distribution of Prokaryotic Communities in the North Pacific Subtropical Front”, and the scientific program can be found here (https://fems2020belgrade.com/scientific-program/).

7. References

Acha, E. M., Piola, A., Iribarne, O., & Mianzan, H. (2015). Ecological processes at marine fronts: oases in the ocean. *Springer*.

Agogué, H., Lamy, D., Neal, P. R., Sogin, M. L., & Herndl, G. J. (2011). Water mass-specificity of bacterial communities in the North Atlantic revealed by massively parallel sequencing. *Molecular Ecology*, *20*(2), 258–274. https://doi.org/10.1111/j.1365-294X.2010.04932.x

Aksenov, Y., Bacon, S., Coward, A. C., & Holliday, N. P. (2010). Polar outflow from the Arctic Ocean: A high resolution model study. *Journal of Marine Systems*, *83*(1–2), 14–37. https://doi.org/10.1016/j.jmarsys.2010.06.007

Alves Junior, N., Meirelles, P. M., de Oliveira Santos, E., Dutilh, B., Silva, G. G. Z., Paranhos, R., … Thompson, F. (2014). Microbial community diversity and physical–chemical features of the Southwestern Atlantic Ocean. *Archives of Microbiology*, *197*(2), 165–179. https://doi.org/10.1007/s00203-014-1035-6

Arístegui, J., Gasol, J. M., Duarte, C. M., & Herndl, G. J. (2009). Microbial oceanography of the dark ocean's pelagic realm. *Limnology and Oceanography*, *54*(5), 1501–1529. https://doi.org/10.4319/lo.2009.54.5.1501

Apprill, A., Mcnally, S., Parsons, R., & Weber, L. (2015). Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquatic Microbial Ecology*, *75*(2), 129–137. https://doi.org/10.3354/ame01753

Ayers, J. M., & Lozier, M. S. (2010). Physical controls on the seasonal migration of the North Pacific transition zone chlorophyll front. *Journal of Geophysical Research: Oceans*, *115*(5). https://doi.org/10.1029/2009JC005596

Azam, F., & Malfatti, F. (2007). Microbial structuring of marine ecosystems. *Nature Reviews Microbiology*. Nature Publishing Group. https://doi.org/10.1038/nrmicro1747

Bano, N., & Hollibaugh, J. T. (2002). Phylogenetic composition of bacterioplankton assemblages from the Arctic Ocean. *Applied and Environmental Microbiology*, *68*(2), 505–518. https://doi.org/10.1128/AEM.68.2.505-518.2002

Bauer, S. L. M., Stensland, A., Daae, F. L., Sandaa, R. A., Thorseth, I. H., Steen, I. H., & Dahle, H. (2018). Water masses and depth structure prokaryotic and T4-like viral communities around hydrothermal systems of the Nordic Seas. *Frontiers in Microbiology*, *9*(MAY). https://doi.org/10.3389/fmicb.2018.01002

Beja, O., Aravind, L., Koonin, E. V., Suzuki, M. T., Hadd, A., Nguyen, L. P., … DeLong, E. F. (2000). Bacterial rhodopsin: Evidence for a new type of phototrophy in the sea. *Science*, *289*(5486), 1902–1906. https://doi.org/10.1126/science.289.5486.1902

Belkin, I.M., 2002. Front. In: Nybakken, J.W., Broenkow, W.W., Vallier, T.L. (Eds.), *Interdisciplinary Encyclopedia of Marine Sciences*. Grolier, Danbury, CT, pp. 433–436.

Belkin, I. M., & Cornillon, P. C. (2007). Fronts in the world ocean's large marine ecosystems. *Ices Cm*, (March), 1–33.

Belkin, I. M., Cornillon, P. C., & Sherman, K. (2009). Fronts in the world ocean's large marine ecosystems. *Progress in Oceanography*, *81 (1–4)*, 223–236.

Belkin, I. M., Hunt, G. L., Hazen, E. L., Zamon, J. E., Schick, R. S., Prieto, R., … Zhang, W. (2014, September 1). Fronts, fish, and predators. *Deep-Sea Research Part II: Topical Studies in Oceanography*. Elsevier Ltd. https://doi.org/10.1016/j.dsr2.2014.07.009

Bograd, S. J., Foley, D. G., Schwing, F. B., Wilson, C., Laurs, R. M., Polovina, J. J., … Brainard, R. E. (2004). On the seasonal and interannual migrations of the transition zone chlorophyll front. *Geophysical Research Letters*, *31*(17). https://doi.org/10.1029/2004GL020637

Bost, C. A., Cotté, C., Bailleul, F., Cherel, Y., Charrassin, J. B., Guinet, C., … Weimerskirch, H. (2009). The importance of oceanographic fronts to marine birds and mammals of the southern oceans. *Journal of Marine Systems*, *78*(3), 363–376. https://doi.org/10.1016/j.jmarsys.2008.11.022

Brown, M. V., Ostrowski, M., Grzymski, J. J., & Lauro, F. M. (2014). A trait-based perspective on the biogeography of common and abundant marine bacterioplankton clades. *Marine Genomics*. Elsevier. https://doi.org/10.1016/j.margen.2014.03.002

Brown, M. V., Philip, G. K., Bunge, J. A., Smith, M. C., Bissett, A., Lauro, F. M., … Donachie, S. P. (2009). Microbial community structure in the North Pacific Ocean. *ISME Journal*, *3*(12), 1374–1386. https://doi.org/10.1038/ismej.2009.86

Buchan, A., LeCleir, G. R., Gulvik, C. A., & González, J. M. (2014). Master recyclers: features and functions of bacteria associated with phytoplankton blooms. Nature Reviews. Microbiology. https://doi.org/10.1038/nrmicro3326

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, *13*(7), 581–583. https://doi.org/10.1038/nmeth.3869

Cantonati, M., Guella, G., Komárek, J., & Spitale, D. (2014). Depth distribution of epilithic cyanobacteria and pigments in a mountain lake characterized by marked water-level fluctuations. *Freshwater Science*, *33*(2), 537–547. https://doi.org/10.1086/675930

Cao, S., Zhang, F., He, J., Ji, Z., & Zhou, Q. (2020). Water masses influence bacterioplankton community structure in summer Kongsfjorden. *Extremophiles*, *24*(1), 107–120. https://doi.org/10.1007/s00792-019-01139-y

Capone, D. G., Burns, J. A., Montoya, J. P., Subramaniam, A., Mahaffey, C., Gunderson, T., … Carpenter, E. J. (2005). Nitrogen fixation by Trichodesmium spp.: An important source of new nitrogen to the tropical and subtropical North Atlantic Ocean. *Global Biogeochemical Cycles*, *19*(2), 1–17. https://doi.org/10.1029/2004GB002331

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Huntley, J., Fierer, N., … Knight, R. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME Journal*, *6*(8), 1621–1624. https://doi.org/10.1038/ismej.2012.8

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., … Knight, R. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(SUPPL. 1), 4516–4522. https://doi.org/10.1073/pnas.1000080107

Carr, N. G., & Mann, N. H. (1994). The Oceanic Cyanobacterial Picoplankton. In *The Molecular Biology of Cyanobacteria* (pp. 27–48). Springer Netherlands. https://doi.org/10.1007/978-94-011-0227-8_2

Carral, L., de Lara Rey, J., Alvarez-Feal, J. C., & Carral Couce, J. (2017). Winch control gear for CTD sampling with a system to compensate vertical motion heave when manoeuvring in rough seas. *Ocean Engineering*. Elsevier Ltd. https://doi.org/10.1016/j.oceaneng.2017.03.005

Celussi, M., Bergamasco, A., Cataletto, B., Umani, S. F., & Del Negro, P. (2010). Water masses bacterial community structure and microbial activities in the Ross Sea, Antarctica. *Antarctic Science*, *22*(4), 361–370. https://doi.org/10.1017/S0954102010000192

Celussi, M., Cataletto, B., Fonda Umani, S., & Del Negro, P. (2009). Depth profiles of bacterioplankton assemblages and their activities in the Ross Sea. *Deep-Sea Research Part I: Oceanographic Research Papers*, *56*(12), 2193–2205. https://doi.org/10.1016/j.dsr.2009.09.001

Chai, F., Jiang, M., Barber, R. T., Dugdale, R. C., & Chao, Y. (2003). Interdecadal variation of the Transition Zone Chlorophyll Front: A physical-biological model simulation between 1960 and 1990. *Journal of Oceanography*, *59*(4), 461–475. https://doi.org/10.1023/A:1025540632491

CHAO, A. (1984). Nonparametric estimation of the number of classes in a population. *Scandinavian Journal of Statistics*, *11*(4), 265–270.

Chen, Q., Hu, H., Ren, X., & Yang, X. Q. (2019). Numerical Simulation of Midlatitude Upper-Level Zonal Wind Response to the Change of North Pacific Subtropical Front Strength. *Journal of Geophysical Research: Atmospheres*, *124*(9), 4891–4912. https://doi.org/10.1029/2018JD029589

Comeau, A. M., Douglas, G. M., & Langille, M. G. I. (2017). Microbiome Helper: a Custom and Streamlined Workflow for Microbiome Research. *MSystems*, *2*(1). https://doi.org/10.1128/msystems.00127-16

Coolen, M. J. L., Abbas, B., Van Bleijswijk, J., Hopmans, E. C., Kuypers, M. M. M., Wakeham, S. G., & Sinninghe Damsté, J. S. (2007). Putative ammonia-oxidizing Crenarchaeota in suboxic waters of the Black Sea: A basin-wide ecological study using 16S ribosomal and functional genes and membrane lipids. *Environmental Microbiology*, *9*(4), 1001–1016. https://doi.org/10.1111/j.1462-2920.2006.01227.x

Danovaro, R., Company, J. B., Corinaldesi, C., D'Onghia, G., Galil, B., Gambi, C., … Tselepides, A. (2010). Deep-sea biodiversity in the Mediterranean Sea: The known, the unknown, and the unknowable. *PLoS ONE*, *5*(8). https://doi.org/10.1371/journal.pone.0011832

Darling, K. F., Kucera, M., Pudsey, C. J., & Wade, C. M. (2004). Molecular evidence links cryptic diversification in polar planktonic protists to Quaternary climate dynamics. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(20), 7657–7662. https://doi.org/10.1073/pnas.0402401101

de Sousa, A. G. G., Tomasino, M. P., Duarte, P., Fernández-Méndez, M., Assmy, P., Ribeiro, H., … Magalhães, C. (2019). Diversity and Composition of Pelagic Prokaryotic and Protist Communities in a Thin Arctic Sea-Ice Regime. *Microbial Ecology*, *78*(2), 388–408. https://doi.org/10.1007/s00248-018-01314-2

De Vargas, C., Norris, R., Zaninetti, L., Gibb, S. W., & Pawlowski, J. (1999). Molecular evidence of cryptic speciation in planktonic foraminifers and their relation to oceanic provinces. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(6), 2864–2868. https://doi.org/10.1073/pnas.96.6.2864

Del Giorgio, P. A., & Cole, J. J. (1998). Bacterial growth efficiency in natural aquatic systems. *Annual Review of Ecology and Systematics*, *29*, 503–541. https://doi.org/10.1146/annurev.ecolsys.29.1.503

Debroas, D., Humbert, J. F., Enault, F., Bronner, G., Faubladier, M., & Cornillot, E. (2009). Metagenomic approach studying the taxonomic and functional diversity of the bacterial community in a mesotrophic lake (Lac du Bourget - France). *Environmental Microbiology*, *11*(9), 2412–2424. https://doi.org/10.1111/j.1462-2920.2009.01969.x

Delmont, T. O., Quince, C., Shaiber, A., Esen, Ö. C., Lee, S. T., Rappé, M. S., … Eren, A. M. (2018). Nitrogen-fixing populations of Planctomycetes and Proteobacteria are abundant in surface ocean metagenomes. *Nature Microbiology*, *3*(7), 804–813. https://doi.org/10.1038/s41564-018-0176-9

DeLong, E. F. (2005). Microbial community genomics in the ocean. *Nature Reviews Microbiology*. https://doi.org/10.1038/nrmicro1158

DeLong, E. F., Preston, C. M., Mincer, T., Rich, V., Hallam, S. J., Frigaard, N. U., … Karl, D. M. (2006). Community genomics among stratified microbial assemblages in the ocean's interior. *Science*, *311*(5760), 496–503. https://doi.org/10.1126/science.1120250

Deming, J. W., & Eric Collins, R. (2016). Sea ice as a habitat for Bacteria, Archaea and viruses. In *Sea Ice: Third Edition* (pp. 326–351). Wiley Blackwell. https://doi.org/10.1002/9781118778371.ch13

Dinsdale, E. A., Edwards, R. A., Hall, D., Angly, F., Breitbart, M., Brulc, J. M., … Rohwer, F. (2008). Functional metagenomic profiling of nine biomes. *Nature*, *452*(7187), 629–632. https://doi.org/10.1038/nature06810

Dore, J. E., Brum, J. R., Tupas, L. M., & Karl, D. M. (2002). Seasonal and interannual variability in sources of nitrogen supporting export in the oligotrophic subtropical North Pacific Ocean. *Limnology and Oceanography*, *47*(6), 1595–1607. https://doi.org/10.4319/lo.2002.47.6.1595

Du, P., Jiang, Z. B., Wang, Y. M., Qian, P., Hu, H. J., Liao, Y. B., … Zhang, D. M. (2019). Spatial heterogeneity of the planktonic protistan community in a semi-closed eutrophic bay, China. *Journal of Plankton Research*, *41*(3), 223–239. https://doi.org/10.1093/plankt/fbz012

Dugdale, R. C., & Goering, J. J. (1967). UPTAKE OF NEW AND REGENERATED FORMS OF NITROGEN IN PRIMARY PRODUCTIVITY. *Limnology and Oceanography*, *12*(2), 196–206. https://doi.org/10.4319/lo.1967.12.2.0196

Dusa, A. (2020). venn: Draw Venn Diagrams. R package version 1.9. https://CRAN.R-project.org/package=venn

Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., & Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics*, *27*(16), 2194–2200. https://doi.org/10.1093/bioinformatics/btr381

Eiler, A., Zaremba-Niedzwiedzka, K., Martínez-García, M., Mcmahon, K. D., Stepanauskas, R., Andersson, S. G. E., & Bertilsson, S. (2014). Productivity and salinity structuring of the microplankton revealed by comparative freshwater metagenomics. *Environmental Microbiology*, *16*(9), 2682–2698. https://doi.org/10.1111/1462-2920.12301

Eppley, R. W., & Peterson, B. J. (1979). Particulate organic matter flux and planktonic new production in the deep ocean. *Nature*, *282*(5740), 677–680. https://doi.org/10.1038/282677a0

Fernandes, G. L., Shenoy, B. D., & Damare, S. R. (2020). Diversity of Bacterial Community in the Oxygen Minimum Zones of Arabian Sea and Bay of Bengal as Deduced by Illumina Sequencing. *Frontiers in Microbiology*, *10*. https://doi.org/10.3389/fmicb.2019.03153

Fiala, M., Sournia, A., Claustre, H., Marty, J. C., Prieur, L., & Vétion, G. (1994). Gradients of phytoplankton abundance, composition and photosynthetic pigments across the Almeria-Oran front (SW Mediterranean Sea). *Journal of Marine Systems*, *5*(3–5), 223–233. https://doi.org/10.1016/0924-7963(94)90048-5

Follows, M., Dutkiewicz, S., Grant, S., & Chisholm, S. (2007). Emergent Biogeography of Microbial Communities in a Model Ocean. *Science, 315 (5820),* 1843 - 1846. https://doi.org/10.1126/science.1138544

Fortunato, C. S., Herfort, L., Zuber, P., Baptista, A. M., & Crump, B. C. (2012). Spatial variability overwhelms seasonal patterns in bacterioplankton communities across a river to ocean gradient. *ISME Journal*, *6*(3), 554–563. https://doi.org/10.1038/ismej.2011.135

Francis, C. A., Beman, J. M., & Kuypers, M. M. M. (2007, May). New processes and players in the nitrogen cycle: The microbial ecology of anaerobic and archaeal ammonia oxidation. *ISME Journal*. https://doi.org/10.1038/ismej.2007.8

Frank, A. H., Garcia, J. A. L., Herndl, G. J., & Reinthaler, T. (2016). Connectivity between surface and deep waters determines prokaryotic diversity in the North Atlantic Deep Water. *Environmental Microbiology*, *18*(6), 2052–2063. https://doi.org/10.1111/1462-2920.13237

Fu, Y., Rivkin, R. B., & Lang, A. S. (2019). Effects of vertical water mass segregation on bacterial community structure in the Beaufort Sea. *Microorganisms*, *7*(10). https://doi.org/10.3390/microorganisms7100385

Fuhrman, J. A., Steele, J. A., Hewson, I., Schwalbach, M. S., Brown, M. V., Green, J. L., & Brown, J. H. (2008). A latitudinal diversity gradient in planktonic marine bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(22), 7774–7778. https://doi.org/10.1073/pnas.0803070105

Galand, P. E., Casamayor, E. O., Kirchman, D. L., & Lovejoy, C. (2009b). Ecology of the rare microbial biosphere of the Arctic Ocean. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(52), 22427–22432. https://doi.org/10.1073/pnas.0908284106

Galand, P. E., Lovejoy, C., Hamilton, A. K., Ingram, R. G., Pedneault, E., & Carmack, E. C. (2009a). Archaeal diversity and a gene for ammonia oxidation are coupled to oceanic circulation. *Environmental Microbiology*, *11*(4), 971–980. https://doi.org/10.1111/j.1462-2920.2008.01822.x

Galand, P. E., Potvin, M., Casamayor, E. O., & Lovejoy, C. (2010). Hydrography shapes bacterial biogeography of the deep Arctic Ocean. *ISME Journal*, *4*(4), 564–576. https://doi.org/10.1038/ismej.2009.134

Garcia-Pichel, F. (2009). Cyanobacteria. In Encyclopedia of Microbiology (pp. 107-124). Elsevier Inc.. https://doi.org/10.1016/B978-012373944-5.00250-9

Ghiglione, J. F., Galand, P. E., Pommier, T., Pedrós-Alió, C., Maas, E. W., Bakker, K., … Murray, A. E. (2012). Pole-to-pole biogeography of surface and deep marine bacterial communities. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(43), 17633–17638. https://doi.org/10.1073/pnas.1208160109

Ghiglione, J. F., Palacios, C., Marty, J. C., Mével, G., Labrune, C., Conan, P., … Goutx, M. (2008). Role of environmental factors for the vertical distribution (0–1000 m) of marine bacterial communities in the NW Mediterranean Sea. *Biogeosciences Discussions*, *5*(3), 2131–2164. https://doi.org/10.5194/bgd-5-2131-2008

Giebel, H. A., Brinkhoff, T., Zwisler, W., Selje, N., & Simon, M. (2009). Distribution of Roseobacter RCA and SAR11 lineages and distinct bacterial communities from the subtropics to the Southern Ocean. *Environmental Microbiology*, *11*(8), 2164–2178. https://doi.org/10.1111/j.1462-2920.2009.01942.x

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Huntley, J., Fierer, N., … Knight, R. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME Journal*, *6*(8), 1621–1624. https://doi.org/10.1038/ismej.2012.8

Giovannoni, S. J., Rappé, M. S., Vergin, K. L., & Adair, N. L. (1996). 16S rRNA genes reveal stratified open ocean bacterioplankton populations related to the green non-sulfur

bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(15), 7979–7984. https://doi.org/10.1073/pnas.93.15.7979

Giovannoni, S. J., & Stingl, U. (2005). Molecular diversity and ecology of microbial plankton. *Nature*. Nature Publishing Group. https://doi.org/10.1038/nature04158

Glaubitz, S., Lueders, T., Abraham, W. R., Jost, G., Jürgens, K., & Labrenz, M. (2009). 13C-isotope analyses reveal that chemolithoautotrophic Gamma- and Epsilonproteobacteria feed a microbial food web in a pelagic redoxcline of the central Baltic Sea. *Environmental Microbiology*, *11*(2), 326–337. https://doi.org/10.1111/j.1462-2920.2008.01770.x

Glöckner, F. O., Fuchs, B. M., & Amann, R. (1999). Bacterioplankton compositions of lakes and oceans: A first comparison based on fluorescence in situ hybridization. *Applied and Environmental Microbiology*, *65*(8), 3721–3726. https://doi.org/10.1128/aem.65.8.3721-3726.1999

Gupta, R. S. (2000). The phylogeny of proteobacteria: relationships to other eubacterial phyla and eukaryotes. *FEMS Microbiology Reviews*, *24*(4), 367–402. https://doi.org/10.1111/j.1574-6976.2000.tb00547.x

Hagström, Å., Ammerman, J., Henrichs, S., & Azam, F. (1984). Bacterioplankton growth in seawater: II. Organic matter utilization during steady-state growth in seawater cultures. *Marine Ecology Progress Series*, *18*, 41–48. https://doi.org/10.3354/meps018041

Hahnke, R. L., Probian, C., Fuchs, B. M., & Harder, J. (2013). Variations in pelagic bacterial communities in the North Atlantic Ocean coincide with water bodies. *Aquatic Microbial Ecology*, *71*(2), 131–140. https://doi.org/10.3354/ame01668

Hamilton, T. L., Bryant, D. A., & Macalady, J. L. (2016). The role of biology in planetary evolution: Cyanobacterial primary production in low-oxygen Proterozoic oceans. *Environmental Microbiology*. Blackwell Publishing Ltd. https://doi.org/10.1111/1462-2920.13118

Hamilton, A. K., Lovejoy, C., Galand, P. E., & Ingram, R. G. (2008). Water masses and biogeography of picoeukaryote assemblages in a cold hydrographically complex system. *Limnology and Oceanography*, *53*(3), 922–935. https://doi.org/10.4319/lo.2008.53.3.0922

Han, D., Ha, H. K., Hwang, C. Y., Lee, B. Y., Hur, H. G., & Lee, Y. K. (2015). Bacterial communities along stratified water columns at the Chukchi Borderland in the western Arctic Ocean. *Deep-Sea Research Part II: Topical Studies in Oceanography*, *120*, 52–60. https://doi.org/10.1016/j.dsr2.2015.01.018

Handelsman, J., Rondon, M. R., Brady, S. F., Clardy, J., & Goodman, R. M. (1998). Molecular biological access to the chemistry of unknown soil microbes: A new frontier for natural products. *Chemistry and Biology*, *5*(10). https://doi.org/10.1016/S1074-5521(98)90108-9

Hansman, R. L., Griffin, S., Watson, J. T., Druffel, E. R. M., Ingalls, A. E., Pearson, A., & Aluwihare, L. I. (2009). The radiocarbon signature of microorganisms in the mesopelagic ocean. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(16), 6513–6518. https://doi.org/10.1073/pnas.0810871106

Harrell, J. F. E., with contributions from Charles Dupont and many others. (2020). Hmisc: Harrell Miscellaneous. R package version 4.3-1. https://CRAN.R-project.org/package=Hmisc

Hauptmann, A. L., Stibal, M., Bælum, J., Sicheritz-Pontén, T., Brunak, S., Bowman, J. S., … Blom, N. (2014). Bacterial diversity in snow on North Pole ice floes. *Extremophiles*, *18*(6), 945–951. https://doi.org/10.1007/s00792-014-0660-y

Hernández-Avilés, J. S., Callieri, C., Bertoni, R., Morabito, G., Leoni, B., Lepori, F., … Salmaso, N. (2018). Prokaryoplankton and phytoplankton community compositions in five large deep perialpine lakes. *Hydrobiologia*, *824*(1), 71–92. https://doi.org/10.1007/s10750-018-3586-z

Herndl, G. J., Agogué, H., Baltar, F., Reinthaler, T., Sintes, E., & Varela, M. M. (2008). Regulation of aquatic microbial processes: the 'microbial loop' of the sunlit surface waters and the dark ocean dissected. *Aquatic Microbial Ecology*, *53*(Sp. Iss. 1), 59–68. https://doi.org/10.3354/ame01225

Herndl, G. J., Reinthaler, T., Teira, E., Van Aken, H., Veth, C., Pernthaler, A., & Pernthaler, J. (2005). Contribution of Archaea to total prokaryotic production in the deep atlantic ocean. *Applied and Environmental Microbiology*, *71*(5), 2303–2309. https://doi.org/10.1128/AEM.71.5.2303-2309.2005

Hopkins, J., Challenor, P., & Shaw, A. G. P. (2010). A new statistical modeling approach to ocean front detection from SST satellite images. *Journal of Atmospheric and Oceanic Technology*, *27*(1), 173–191. https://doi.org/10.1175/2009JTECHO684.1

Howell, E. A., Kobayashi, D. R., Parker, D. M., Balazs, G. H., & Polovina, J. J. (2008). TurtleWatch: A tool to aid in the bycatch reduction of loggerhead turtles *Caretta caretta* in the Hawaii-based pelagic longline fishery. *Endangered Species Research*, *5*(2–3), 267–278. https://doi.org/10.3354/esr00096

Huber, J. A., Mark Welch, D. B., Morrison, H. G., Huse, S. M., Neal, P. R., Butterfield, D. A., & Sogin, M. L. (2007). Microbial population structures in the deep marine biosphere. *Science*, *318*(5847), 97–100. https://doi.org/10.1126/science.1146689

Imhoff, J. F., Hiraishi, A., & Süling, J. (2005). Anoxygenic Phototrophic Purple Bacteria. In *Bergey's Manual® of Systematic Bacteriology* (pp. 119–132). Springer US. https://doi.org/10.1007/0-387-28021-9_15

Ji, Y., Ashton, L., Pedley, S. M., Edwards, D. P., Tang, Y., Nakamura, A., … Yu, D. W. (2013). Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology Letters*, *16*(10), 1245–1257. https://doi.org/10.1111/ele.12162

Johnson, K. S., Riser, S. C., & Karl, D. M. (2010). Nitrate supply from deep to near-surface waters of the North Pacific subtropical gyre. *Nature*, *465*(7301), 1062–1065. https://doi.org/10.1038/nature09170

Kahru, M., Di Lorenzo, E., Manzano-Sarabia, M., & Mitchell, B. G. (2012). Spatial and temporal statistics of sea surface temperature and chlorophyll fronts in the California Current. *Journal of Plankton Research*, *34*(9), 749–760. https://doi.org/10.1093/plankt/fbs010

Karl, D. M., Hebel, D. V., Björkman, K., & Letelier, R. M. (1998). The role of dissolved organic matter release in the productivity of the oligotrophic North Pacific Ocean. *Limnology and Oceanography*, *43*(6), 1270–1286. https://doi.org/10.4319/lo.1998.43.6.1270

Karner, M. B., Delong, E. F., & Karl, D. M. (2001). Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature*, *409*(6819), 507–510. https://doi.org/10.1038/35054051

Kersters, K., De Vos, P., Gillis, M., Swings, J., Vandamme, P., & Stackebrandt, E. (2006). Introduction to the Proteobacteria. In *The Prokaryotes* (pp. 3–37). Springer New York. https://doi.org/10.1007/0-387-30745-1_1

Kobashi, F., & Xie, S.-P. (2011). Interannual variability of the North Pacific Subtropical Countercurrent: role of local ocean–atmosphere interaction. In *New Developments in Mode-Water Research* (pp. 121–134). Springer Japan. https://doi.org/10.1007/978-4-431-54162-2_9

Könneke, M., Bernhard, A. E., De La Torre, J. R., Walker, C. B., Waterbury, J. B., & Stahl, D. A. (2005). Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature*, *437*(7058), 543–546. https://doi.org/10.1038/nature03911

Kopf, A., Bicak, M., Kottmann, R., Schnetzer, J., Kostadinov, I., Lehmann, K., … Glöckner, F. O. (2015). The ocean sampling day consortium. *GigaScience*. BioMed Central Ltd. https://doi.org/10.1186/s13742-015-0066-5

Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., & Schloss, P. D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the miseq illumina sequencing platform. *Applied and Environmental Microbiology*, *79*(17), 5112–5120. https://doi.org/10.1128/AEM.01043-13

Latasa, M., Cabello, A. M., Morán, X. A. G., Massana, R., & Scharek, R. (2017). Distribution of phytoplankton groups within the deep chlorophyll maximum. *Limnology and Oceanography*, *62*(2), 665–685. https://doi.org/10.1002/lno.10452

Lauro, F. M., & Bartlett, D. H. (2008). Prokaryotic lifestyles in deep sea habitats. *Extremophiles*. Springer Japan. https://doi.org/10.1007/s00792-006-0059-5

Lavik, G., Stührmann, T., Brüchert, V., Van Der Plas, A., Mohrholz, V., Lam, P., … Kuypers, M. M. M. (2009). Detoxification of sulphidic African shelf waters by blooming chemolithotrophs. *Nature*, *457*(7229), 581–584. https://doi.org/10.1038/nature07588

Le Boeuf, B. J., Crocker, D. E., Costa, D. P., Blackwell, S. B., Webb, P. M., & Houser, D. S. (2000). Foraging ecology of northern elephant seals. *Ecological Monographs*, *70*(3), 353–382. https://doi.org/10.1890/0012-9615(2000)070[0353:FEONES]2.0.CO;2

Letelier, R. M., Karl, D. M., Abbott, M. R., & Bidigare, R. R. (2004). Light driven seasonal patterns of chlorophyll and nitrate in the lower euphotic zone of the North Pacific Subtropical

Gyre. *Limnology and Oceanography*, *49*(2), 508–519. https://doi.org/10.4319/lo.2004.49.2.0508

Li, W., Wang, M., Burgaud, G., Yu, H., & Cai, L. (2019). Fungal Community Composition and Potential Depth-Related Driving Factors Impacting Distribution Pattern and Trophic Modes from Epi- to Abyssopelagic Zones of the Western Pacific Ocean. *Microbial Ecology*, *78*(4), 820–831. https://doi.org/10.1007/s00248-019-01374-y

Liang, X., Wagner, R. E., Zhuang, J., DeBruyn, J. M., Wilhelm, S. W., Liu, F., … Radosevich, M. (2019). Viral abundance and diversity vary with depth in a southeastern United States agricultural ultisol. *Soil Biology and Biochemistry*, *137*. https://doi.org/10.1016/j.soilbio.2019.107546

Lindell, D., & Post, A. F. (1995). Ultraphytoplankton succession is triggered by deep winter mixing in the Gulf of Aqaba (Eilat), Red Sea. *Limnology and Oceanography*, *40*(6), 1130–1141. https://doi.org/10.4319/lo.1995.40.6.1130

Longhurst, A. R. (2007). Chapter 11 - The Pacific Ocean. *Ecological Geography of the Sea (Second Edition)*, 327–441. https://doi.org/http://dx.doi.org/10.1016/B978-012455521-1/50012-7

López-Garcia, P., López-López, A., Moreira, D., & Rodriguez-Valera, F. (2001). Diversity of free-living prokaryotes from a deep-sea site at the Antarctic Polar Front. *FEMS Microbiology Ecology*, *36*(2–3), 193–202. https://doi.org/10.1016/S0168-6496(01)00133-7

Lozupone, C. A., & Knight, R. (2007). Global patterns in bacterial diversity. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(27), 11436–11440. https://doi.org/10.1073/pnas.0611525104

Mann, K. H., & Lazier, J. R. N. (2013). *Dynamics of Marine Ecosystems: Biological-Physical Interactions in the Oceans: Third Edition*. *Dynamics of Marine Ecosystems: Biological-Physical Interactions in the Oceans: Third Edition* (pp. 1–496). Blackwell Publishing Ltd. https://doi.org/10.1002/9781118687901

Martin, J. H., Coale, K. H., Johnson, K. S., Fitzwater, S. E., Gordon, R. M., Tanner, S. J., … Tindale, N. W. (1994). Testing the iron hypothesis in ecosystems of the equatorial Pacific Ocean. *Nature*, *371*(6493), 123–129. https://doi.org/10.1038/371123a0

Martiny, J. B. H., Bohannan, B. J. M., Brown, J. H., Colwell, R. K., Fuhrman, J. A., Green, J. L., … Staley, J. T. (2006). Microbial biogeography: Putting microorganisms on the map. *Nature Reviews Microbiology*. https://doi.org/10.1038/nrmicro1341

McMurdie, P. J., & Holmes, S. (2013). Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE*, *8*(4). https://doi.org/10.1371/journal.pone.0061217

Mente, E., Nikouli, E., Antonopoulou, E., Martin, S. A. M., & Kormas, K. A. (2018). Core versus diet-associated and postprandial bacterial communities of the rainbow trout (*Oncorhynchus mykiss*) midgut and faeces. *Biology Open*, *7*(6). https://doi.org/10.1242/bio.034397

Mincer, T. J., Church, M. J., Taylor, L. T., Preston, C., Karl, D. M., & DeLong, E. F. (2007). Quantitative distribution of presumptive archaeal and bacterial nitrifiers in Monterey Bay and the North Pacific Subtropical Gyre. *Environmental Microbiology*, *9*(5), 1162–1175. https://doi.org/10.1111/j.1462-2920.2007.01239.x

Moore, L. R., Post, A. F., Rocap, G., & Chisholm, S. W. (2002). Utilization of different nitrogen sources by the marine cyanobacteria Prochlorococcus and Synechococcus. *Limnology and Oceanography*, *47*(4), 989–996. https://doi.org/10.4319/lo.2002.47.4.0989

Moon, C. D., Young, W., Maclean, P. H., Cookson, A. L., & Bermingham, E. N. (2018). Metagenomic insights into the roles of Proteobacteria in the gastrointestinal microbiomes of healthy dogs and cats. *Microbiology Open*. Blackwell Publishing Ltd. https://doi.org/10.1002/mbo3.677

Morán, X. A. G., & Calvo-Díaz, A. (2009). Single-cell vs. bulk activity properties of coastal bacterioplankton over an annual cycle in a temperate ecosystem. *FEMS Microbiology Ecology*, *67*(1), 43–56. https://doi.org/10.1111/j.1574-6941.2008.00601.x

Nakamura, H., & Kazmin, A. S. (2003). Decadal changes in the North Pacific oceanic frontal zones as revealed in ship and satellite observations. *Journal of Geophysical Research: Oceans*, *108*(3). https://doi.org/10.1029/1999jc000085

Nakamura, H., Lin, G., & Yamagata, T. (1997). Decadal Climate Variability in the North Pacific during the Recent Decades. *Bulletin of the American Meteorological Society*, *78*(10), 2215–2225. https://doi.org/10.1175/1520-0477(1997)078<2215:DCVITN>2.0.CO;2

Oh, S., Caro-Quintero, A., Tsementzi, D., DeLeon-Rodriguez, N., Luo, C., Poretsky, R., & Konstantinidis, K. T. (2011). Metagenomic insights into the evolution, function, and complexity of the planktonic microbial community of Lake Lanier, a temperate freshwater ecosystem. *Applied and Environmental Microbiology*, *77*(17), 6000–6011. https://doi.org/10.1128/AEM.00107-11

Ohman, M. D., Powell, J. R., Picheral, M., & Jensen, D. W. (2012). Mesozooplankton and particulate matter responses to a deep-water frontal system in the southern California Current System. *Journal of Plankton Research*, *34*(9), 815–827. https://doi.org/10.1093/plankt/fbs028

Okkonen, S. R., Ashjian, C. J., Campbell, R. G., Clarke, J. T., Moore, S. E., & Taylor, K. D. (2011). Satellite observations of circulation features associated with a bowhead whale feeding “hotspot” near Barrow, Alaska. *Remote Sensing of Environment*, *115*(8), 2168–2174. https://doi.org/10.1016/j.rse.2011.04.024

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., … Wagner, H. (2019). vegan: Community Ecology Package. R package version 2.5-2. *Cran R*, *1*, 2. Retrieved from https://cran.r-project.org/package=vegan

Olson, D., Hitchcock, G., Mariano, A., Ashjian, C., Peng, G., Nero, R., & Podesta, G. (1994). Life on the Edge: Marine Life and Fronts. *Oceanography*, *7*(2), 52–60. https://doi.org/10.5670/oceanog.1994.03

Parada, A. E., Needham, D. M., & Fuhrman, J. A. (2016). Every base matters: Assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environmental Microbiology*, *18*(5), 1403–1414. https://doi.org/10.1111/1462-2920.13023

Parris, D. J., Brooker, R. M., Morgan, M. A., Dixson, D. L., & Stewart, F. J. (2016). Whole gut microbiome composition of damselfish and cardinalfish before and after reef settlement. *PeerJ*, *2016*(8). https://doi.org/10.7717/PEERJ.2412

Partensky, F., Blanchot, J., & Vaulot, D. (1999). Differential distribution and ecology of Prochlorococcus and Synechococcus in oceanic waters: a review. In L. Charpy & A. W. D. Larkum (Eds.), *Marine Cyanobacteria* (pp. 457–475). Bulletin de l’Institut Océanographique de Monaco. Numéro spécial 19.

Pavan-Kumar, A., Gireesh-Babu, P., & Lakra, W. S. (2015). DNA Metabacoding: a new approach for rapid biodiversity assessment. *Journal of Cell Science and Molecular Biology*, *2*(1), 111. Retrieved from http://opensciencepublications.com/fulltextarticles/JCMB-2350-0190-2-111.html

Pester, M., Schleper, C., & Wagner, M. (2011). The Thaumarchaeota: An emerging view of their phylogeny and ecophysiology. *Current Opinion in Microbiology*, *14*(3), 300–306. https://doi.org/10.1016/j.mib.2011.04.007

Pichel, W. G., Churnside, J. H., Veenstra, T. S., Foley, D. G., Friedman, K. S., Brainard, R. E., … Clemente-Colón, P. (2007). Marine debris collects within the North Pacific Subtropical Convergence Zone. *Marine Pollution Bulletin*, *54*(8), 1207–1211. https://doi.org/10.1016/j.marpolbul.2007.04.010

Polovina, J. J., Howell, E., Kobayashi, D. R., & Seki, M. P. (2001). The transition zone chlorophyll front, a dynamic global feature defining migration and forage habitat for marine resources. *Progress in Oceanography*. Elsevier Ltd. https://doi.org/10.1016/S0079-6611(01)00036-2

Polovina, J. J., Kobayashi, D. R., Parker, D. M., Seki, M. P., & Balazs, G. H. (2000). Turtles on the edge: Movement of loggerhead turtles (*Caretta caretta*) along oceanic fronts, spanning longline fishing grounds in the central North Pacific, 1997-1998. *Fisheries Oceanography*, *9*(1), 71–82. https://doi.org/10.1046/j.1365-2419.2000.00123.x

Pommier, T., Canbäck, B., Riemann, L., Boström, K. H., Simu, K., Lundberg, P., … Hagström, Å. (2007). Global patterns of diversity and community structure in marine bacterioplankton. *Molecular Ecology*, *16*(4), 867–880. https://doi.org/10.1111/j.1365-294X.2006.03189.x

Pommier, T., Neal, P. R., Gasol, J. M., Coll, M., Acinas, S. G., & Pedrós-Alió, C. (2010). Spatial patterns of bacterial richness and evenness in the NW Mediterranean Sea explored by pyrosequencing of the 16S rRNA. *Aquatic Microbial Ecology*, *61*(3), 221–233. https://doi.org/10.3354/ame01484

Rapp, J. Z., Fernández-Méndez, M., Bienhold, C., & Boetius, A. (2018). Effects of ice-algal aggregate export on the connectivity of bacterial communities in the central Arctic Ocean. *Frontiers in Microbiology*, *9*(MAY). https://doi.org/10.3389/fmicb.2018.01035

Rasheeda, M. K., Rangamaran, V. R., Srinivasan, S., Ramaiah, S. K., Gunasekaran, R., Jaypal, S., … Ramalingam, K. (2017). Comparative profiling of microbial community of three economically important fishes reared in sea cages under tropical offshore environment. *Marine Genomics*, *34*, 57–65. https://doi.org/10.1016/j.margen.2017.04.003

Reji, L., Tolar, B. B., Smith, J. M., Chavez, F. P., & Francis, C. A. (2019). Differential co-occurrence relationships shaping ecotype diversification within Thaumarchaeota populations in the coastal ocean water column. *ISME Journal*, *13*(5), 1144–1158. https://doi.org/10.1038/s41396-018-0311-x

Ribeiro, H., de Sousa, T., Santos, J. P., Sousa, A. G. G., Teixeira, C., Monteiro, M. R., … Magalhães, C. (2018). Potential of dissimilatory nitrate reduction pathways in polycyclic aromatic hydrocarbon degradation. *Chemosphere*, *199*, 54–67. https://doi.org/10.1016/j.chemosphere.2018.01.171

Riemann, L., Nielsen, T. G., Kragh, T., Richardson, K., Parner, H., Jakobsen, H. H., & Munk, P. (2011). Distribution and production of plankton communities in the subtropical convergence zone of the Sargasso Sea. I. Phytoplankton and bacterioplankton. *Marine Ecology Progress Series*, *426*, 57–70. https://doi.org/10.3354/meps09001

Riemann, L., F. Steward, G., Fandino, L. B., Campbell, L., Landry, M. R., & Azam, F. (1999). Bacterial community composition during two consecutive NE Monsoon periods in the Arabian Sea studied by denaturing gradient gel electrophoresis (DGGE) of rRNA genes. *Deep-Sea Research Part II: Topical Studies in Oceanography*, *46*(8–9), 1791–1811. https://doi.org/10.1016/S0967-0645(99)00044-2

Roden, G. I. (1972). Temperature and salinity fronts at the boundaries of the subarctic-subtropical transition zone in the western Pacific. *Journal of Geophysical Research*, *77*(36), 7175–7187. https://doi.org/10.1029/jc077i036p07175

Roden, G. I. (1980). On the Subtropical Frontal Zone North of Hawaii During Winter. *Journal of Physical Oceanography*, *10*(3), 342–362. https://doi.org/10.1175/1520-0485(1980)010<0342:otsfzn>2.0.co;2

RStudio. (2019). R.studio Team. *Rstudio Team*, 1. Retrieved from https://rstudio.com/

Sagarminaga, Y., & Arrizabalaga, H. (2014). Relationship of Northeast Atlantic albacore juveniles with surface thermal and chlorophyll-a fronts. *Deep-Sea Research Part II: Topical Studies in Oceanography*, *107*, 54–63. https://doi.org/10.1016/j.dsr2.2013.11.006

Samo, T. J., Pedler, B. E., Ball, G. I., Pasulka, A. L., Taylor, A. G., Aluwihare, L. I., … Landry, M. R. (2012). Microbial distribution and activity across a water mass frontal zone in the California Current Ecosystem. *Journal of Plankton Research*, *34*(9), 802–814. https://doi.org/10.1093/plankt/fbs048

Sauzède, R., Martinez, E., Maes, C., Pasqueron de Fommervault, O., Poteau, A., Mignot, A., … Laurent, V. (2020). Enhancement of phytoplankton biomass leeward of Tahiti as observed by Biogeochemical-Argo floats. *Journal of Marine Systems*, *204*. https://doi.org/10.1016/j.jmarsys.2019.103284

Schattenhofer, M., Fuchs, B. M., Amann, R., Zubkov, M. V., Tarran, G. A., & Pernthaler, J. (2009). Latitudinal distribution of prokaryotic picoplankton populations in the Atlantic Ocean. *Environmental Microbiology*, *11*(8), 2078–2093. https://doi.org/10.1111/j.1462-2920.2009.01929.x

Schütz, M., Klughammer, C., Griesbeck, C., Quentmeier, A., Friedrich, C. G., & Hauska, G. (1998). Sulfide-quinone reductase activity in membranes of the chemotrophic bacterium *Paracoccus denitrificans* GB17. *Archives of Microbiology*, *170*(5), 353–360. https://doi.org/10.1007/s002030050653

Seki Michael, P., Polovina Jeffrey, J., Kobayashi Donald, R., Bidigare Robert, R., & Mitchum Gary, T. (2002). An oceanographic characterization of swordfish (*Xiphias gladius*) longline fishing grounds in the springtime subtropical North Pacific. *Fisheries Oceanography*. https://doi.org/10.1046/j.1365-2419.2002.00207.x

Sempéré, R., Dafner, E., Van Wambeke, F., Lefève, D., Magen, C., Allègre, S., … Prieur, L. (2003). Distribution and cycling of total organic carbon across the Almeria-Oran front in the Mediterranean Sea: Implications for carbon cycling in the western basin. *Journal of Geophysical Research: Oceans*, *108*(11). https://doi.org/10.1029/2002jc001475

Sexton, P. F., & Norris, R. D. (2008). Dispersal and biogeography of marine plankton: Long-distance dispersal of the foraminifer *Truncorotalia truncatulinoides*. *Geology*, *36*(11), 899–902. https://doi.org/10.1130/G25232A.1

Shaffer, S. A., Tremblay, Y., Awkerman, J. A., Henry, R. W., Teo, S. L. H., Anderson, D. J., … Costa, D. P. (2005). Comparison of light- and SST-based geolocation with satellite telemetry in free-ranging albatrosses. *Marine Biology*, *147*(4), 833–843. https://doi.org/10.1007/s00227-005-1631-8

Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, *27*(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Sherrill-Mix, S., McCormick, K., Lauder, A., Bailey, A., Zimmerman, L., Li, Y., … Bushman, F. D. (2018). Allometry and ecology of the bilaterian gut microbiome. *MBio*, *9*(2). https://doi.org/10.1128/mBio.00319-18

Shiozaki, T., Furuya, K., Kodama, T., & Takeda, S. (2009). Contribution of N2 fixation to new production in the western North Pacific Ocean along 155°E. *Marine Ecology Progress Series*, *377*, 19–32. https://doi.org/10.3354/meps07837

Smith, S. L., & Lane, P. V. Z. (1991). The jet off Point Arena, California: its role in aspects of secondary production in the copepod *Eucalanus californicus* Johnson. *Journal of Geophysical Research*, *96*(C8). https://doi.org/10.1029/91jc00906

Smith, C. C. R., Snowberg, L. K., Gregory Caporaso, J., Knight, R., & Bolnick, D. I. (2015). Dietary input of microbes and host genetic variation shape among-population differences in stickleback gut microbiota. *ISME Journal*, *9*(11), 2515–2526. https://doi.org/10.1038/ismej.2015.64

Sogin, M. L., Morrison, H. G., Huber, J. A., Welch, D. M., Huse, S. M., Neal, P. R., … Herndl, G. J. (2006). Microbial diversity in the deep sea and the underexplored "rare biosphere." *Proceedings of the National Academy of Sciences of the United States of America*, *103*(32), 12115–12120. https://doi.org/10.1073/pnas.0605127103

Sournia, A. (1994). Pelagic biogeography and fronts. *Progress in Oceanography*, *34*(2–3), 109–120. https://doi.org/10.1016/0079-6611(94)90004-3

Spencer-Cervato, C., & Thierstein, H. R. (1997). First appearance of *Globorotalia truncatulinoides*: Cladogenesis and immigration. *Marine Micropaleontology*, *30*(4), 267–291. https://doi.org/10.1016/S0377-8398(97)00004-2

Stackebrandt, E., Murray, R. G. E., & Truper, H. G. (1988). Proteobacteria classis nov., a name for the phylogenetic taxon that includes the "purple bacteria and their relatives." *International Journal of Systematic Bacteriology*, *38*(3), 321–325. https://doi.org/10.1099/00207713-38-3-321

Stewart, F. J., Ulloa, O., & Delong, E. F. (2012). Microbial metatranscriptomics in a permanent marine oxygen minimum zone. *Environmental Microbiology*, *14*(1), 23–40. https://doi.org/10.1111/j.1462-2920.2010.02400.x

Stieglmeier, M., Klingl, A., Alves, R., Rittmann, S., Melcher, M., Leisch, N., & Schleper, C. (2014). *Nitrososphaera viennensis* gen. nov., sp. nov., an aerobic and mesophilic, ammonia-oxidizing archaeon from soil and a member of the archaeal phylum Thaumarchaeota. *International journal of systematic and evolutionary microbiology*, *64*(Pt 8), 2738–2752. https://doi.org/10.1099/ijs.0.063172-0

Storesund, J. E., & Øvreås, L. (2013). Diversity of planctomycetes in iron-hydroxide deposits from the arctic mid ocean ridge (AMOR) and description of bythopirellula goksoyri gen. nov., sp. nov., a novel Planctomycete from deep sea iron-hydroxide deposits. *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*, *104*(4), 569–584. https://doi.org/10.1007/s10482-013-0019-x

Sullam, K. E., Essinger, S. D., Lozupone, C. A., O'Connor, M. P., Rosen, G. L., Knight, R., … Russell, J. A. (2012). Environmental and ecological factors that shape the gut bacterial communities of fish: A meta-analysis. *Molecular Ecology*, *21*(13), 3363–3378. https://doi.org/10.1111/j.1365-294X.2012.05552.x

Swan, B. K., Martinez-Garcia, M., Preston, C. M., Sczyrba, A., Woyke, T., Lamy, D., … Stepanauskas, R. (2011). Potential for chemolithoautotrophy among ubiquitous bacteria lineages in the dark ocean. *Science*, *333*(6047), 1296–1300. https://doi.org/10.1126/science.1203690

Qin, C., Zhu, W., Ma, H., Duan, D., Zuo, T., Xi, S., & Pan, W. (2019). Are habitat changes driving protist community shifts? A case study in Daya Bay, China. *Estuarine, Coastal and Shelf Science*, *227*. https://doi.org/10.1016/j.ecss.2019.106356

Qiu, C., & Kawamura, H. (2012). Study on SST front disappearance in the subtropical North Pacific using microwave SSTs. *Journal of Oceanography*, *68*(3), 417–426. https://doi.org/10.1007/s10872-012-0106-z

Qiu, C., Kawamura, H., Mao, H., & Wu, J. (2014). Mechanisms of the disappearance of sea surface temperature fronts in the subtropical North Pacific Ocean. *Journal of Geophysical Research: Oceans*, *119*(7), 4389–4398. https://doi.org/10.1002/2014JC010142

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., … Glöckner, F. O. (2013). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research*, *41*(D1). https://doi.org/10.1093/nar/gks1219

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., & Willerslev, E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology*, *21*(8), 2045–2050. https://doi.org/10.1111/j.1365-294X.2012.05470.x

Taylor, J. R., & Ferrari, R. (2011). Ocean fronts trigger high latitude phytoplankton blooms. *Geophysical Research Letters*, *38*(23). https://doi.org/10.1029/2011GL049312

Techtmann, S. M., Fortney, J. L., Ayers, K. A., Joyner, D. C., Linley, T. D., Pfiffner, S. M., & Hazen, T. C. (2015). The unique chemistry of Eastern Mediterranean water masses selects for distinct microbial communities by depth. *PLoS ONE*, *10*(3). https://doi.org/10.1371/journal.pone.0120605

Teeling, H., Fuchs, B. M., Becher, D., Klockow, C., Gardebrecht, A., Bennke, C. M., … Amann, R. (2012). Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science*, *336*(6081), 608–611. https://doi.org/10.1126/science.1218344

Teira, E., Lebaron, P., Van Aken, H., & Herndl, G. J. (2006). Distribution and activity of Bacteria and Archaea in the deep water masses of the North Atlantic. *Limnology and Oceanography*, *51*(5), 2131–2144. https://doi.org/10.4319/lo.2006.51.5.2131

Tetlock, A., Yost, C. K., Stavrinides, J., & Manzon, R. G. (2012). Changes in the gut microbiome of the sea lamprey during metamorphosis. *Applied and Environmental Microbiology*, *78*(21), 7638–7644. https://doi.org/10.1128/AEM.01640-12

Thomson, R. E., & Emery, W. J. (2014). *Data Analysis Methods in Physical Oceanography: Third Edition*. *Data Analysis Methods in Physical Oceanography: Third Edition* (pp. 1–716). Elsevier Inc. https://doi.org/10.1016/C2010-0-66362-0

Traganza, E. D., Redalije, D. G., & Garwood, R. W. (1987). Chemical flux, mixed layer entrainment and phytoplankton blooms at upwelling fronts in the California coastal zone. *Continental Shelf Research*, *7*(1), 89–105. https://doi.org/10.1016/0278-4343(87)90066-5

Treusch, A. H., Vergin, K. L., Finlay, L. A., Donatz, M. G., Burton, R. M., Carlson, C. A., & Giovannoni, S. J. (2009). Seasonality and vertical structure of microbial communities in an ocean gyre. *ISME Journal*, *3*(10), 1148–1163. https://doi.org/10.1038/ismej.2009.60

Varela, M. M., Van Aken, H. M., Sintes, E., & Herndl, G. J. (2008). Latitudinal trends of Crenarchaeota and Bacteria in the meso- and bathypelagic water masses of the Eastern North Atlantic. *Environmental Microbiology*, *10*(1), 110–124. https://doi.org/10.1111/j.1462-2920.2007.01437.x

Videau, C., Sournia, A., Prieur, L., & Fiala, M. (1994). Phytoplankton and primary production characteristics at selected sites in the geostrophic Almeria-Oran front system (SW Mediterranean Sea). *Journal of Marine Systems*, *5*(3–5), 235–250. https://doi.org/10.1016/0924-7963(94)90049-3

Vincent, W. F. (2009). Cyanobacteria. In *Encyclopedia of Inland Waters* (pp. 226–232). Elsevier Inc. https://doi.org/10.1016/B978-012370626-3.00127-7

Taiyun Wei and Viliam Simko (2017). R package "corrplot": Visualization of a Correlation Matrix (Version 0.84). Available from https://github.com/taiyun/corrplot

Wang, L. Y., Hu, H. B., Yang, X. Q., & Ren, X. J. (2016). Atmospheric eddy anomalies associated with the wintertime North Pacific subtropical front strength and their influences on the seasonal-mean atmosphere. *Science China Earth Sciences*, *59*(10), 2022–2036. https://doi.org/10.1007/s11430-016-5331-7

West, N. J., Lepère, C., Manes, C. L. de O., Catala, P., Scanlan, D. J., & Lebaron, P. (2016). Distinct spatial patterns of SAR11, SAR86, and Actinobacteria diversity along a transect in the ultra-oligotrophic South Pacific Ocean. *Frontiers in Microbiology*, *7*(MAR). https://doi.org/10.3389/fmicb.2016.00234

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. *Springer*-Verlag New York.

Wilson, B., Müller, O., Nordmann, E. L., Seuthe, L., Bratbak, G., & Øvreås, L. (2017). Changes in marine prokaryote composition with season and depth over an Arctic polar year. *Frontiers in Marine Science*, *4*(APR). https://doi.org/10.3389/fmars.2017.00095

Woodsona, C. B., & Litvinb, S. Y. (2015). Ocean fronts drive marine fishery production and biogeochemical cycling. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(6), 1710–1715. https://doi.org/10.1073/pnas.1417143112

Woo, M., & Pattiaratchi, C. (2008). Hydrography and water masses off the western Australian coast. *Deep-Sea Research Part I: Oceanographic Research Papers*, *55*(9), 1090–1104. https://doi.org/10.1016/j.dsr.2008.05.005

Yang, C., Wang, X., Miller, J. A., De Blécourt, M., Ji, Y., Yang, C., … Yu, D. W. (2014). Using metabarcoding to ask if easily collected soil and leaf-litter samples can be used as a general biodiversity indicator. *Ecological Indicators*, *46*, 379–389. https://doi.org/10.1016/j.ecolind.2014.06.028

Yilmaz, P., Yarza, P., Rapp, J. Z., & Glöckner, F. O. (2016). Expanding the world of marine bacterial and archaeal clades. *Frontiers in Microbiology*, *6*(JAN). https://doi.org/10.3389/fmicb.2015.01524

Yu, Z., Yang, J., Liu, L., Zhang, W., & Amalfitano, S. (2015). Bacterioplankton community shifts associated with epipelagic and mesopelagic waters in the Southern Ocean. *Scientific Reports*, *5*. https://doi.org/10.1038/srep12897

Zaballos, M., López-López, A., Ovreas, L., Bartual, S. G., D'Auria, G., Alba, J. C., … Rodríguez-Valera, F. (2006). Comparison of prokaryotic diversity at offshore oceanic locations reveals a different microbiota in the Mediterranean Sea. *FEMS Microbiology Ecology*, *56*(3), 389–405. https://doi.org/10.1111/j.1574-6941.2006.00060.x

Zainuddin, M., Saitoh, K., & Saitoh, S. I. (2008). Albacore (*Thunnus alalunga*) fishing ground in relation to oceanographic conditions in the western North Pacific Ocean using remotely sensed satellite data. *Fisheries Oceanography*, *17*(2), 61–73. https://doi.org/10.1111/j.1365-2419.2008.00461.x

Zehr, J. P., Carpenter, E. J., & Villareal, T. A. (2000). New perspectives on nitrogen-fixing microorganisms in tropical and subtropical oceans. *Trends in Microbiology*. https://doi.org/10.1016/S0966-842X(99)01670-4

Zhang, M., Sun, Y., Liu, Y., Qiao, F., Chen, L., Liu, W. T., … Li, E. (2016). Response of gut microbiota to salinity change in two euryhaline aquatic animals with reverse salinity preference. *Aquaculture*, *454*, 72–80. https://doi.org/10.1016/j.aquaculture.2015.12.014

Zhang, L., Xu, H., Ma, J., Shi, N., & Deng, J. (2019). North Pacific subtropical sea surface temperature frontogenesis and its connection with the atmosphere above. *Earth System Dynamics*, *10*(2), 261–270. https://doi.org/10.5194/esd-10-261-2019

Zoccarato, L., Pallavicini, A., Cerino, F., Fonda Umani, S., & Celussi, M. (2016). Water mass dynamics shape Ross Sea protist communities in mesopelagic and bathypelagic layers. *Progress in Oceanography*, *149*, 16–26. https://doi.org/10.1016/j.pocean.2016.10.003

8. Supplementary Information

Supplementary information is on a virtual platform and is given as ‘figshare private links’ to be accessed online through the figshare web service.

Supplementary information 1: Classification of prokaryotic taxa across the sample collection at phylum, class, order, family, genus, species and ASVs levels, as well as, its absolute abundance. (DOI: 10.6084/m9.figshare.13138703)

Supplementary information 2: Classification of prokaryotic taxa across the sample collection at phylum, class, order, family, genus, species and ASVs levels, as well as, its relative abundance. (DOI: 10.6084/m9.figshare.13138706)