



HAL
open science

Effect of environmental variations on bacterial communities dynamics : evolution and adaptation of bacteria as part of a microbial observatory in the Bay of Brest and the Iroise Sea

Clarisse Lemonnier

► **To cite this version:**

Clarisse Lemonnier. Effect of environmental variations on bacterial communities dynamics : evolution and adaptation of bacteria as part of a microbial observatory in the Bay of Brest and the Iroise Sea. Ecosystems. Université de Bretagne occidentale - Brest, 2019. English. NNT : 2019BRES0030 . tel-02953054

HAL Id: tel-02953054

<https://theses.hal.science/tel-02953054v1>

Submitted on 29 Sep 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THESE DE DOCTORAT DE

L'UNIVERSITE
DE BRETAGNE OCCIDENTALE
COMUE UNIVERSITE BRETAGNE LOIRE

ECOLE DOCTORALE N° 598
Sciences de la Mer et du littoral
Spécialité : *Microbiologie*

Par

Clarisse LEMONNIER

Effet des changements environnementaux sur la dynamique des communautés microbiennes : évolution et adaptation des bactéries dans le cadre du développement de l'Observatoire Microbien en Rade de Brest et Mer d'Iroise.

Thèse présentée et soutenue à Plouzané, le 26 avril 2019

Unité de recherche : Laboratoire de Microbiologie des Environnements Extrêmes (UMR 6197) et
Laboratoire des Sciences de l'Environnement Marin (UMR 6539)

Rapporteurs avant soutenance :

Pierre GALAND Directeur de recherche CNRS,
LECOB, Banyuls sur mer

Urania CHRISTAKI Professeur, Université du
Littoral Côte d'Opale

Composition du Jury :

Pierre GALAND Directeur de recherche CNRS,
Rapporteur LECOB, Banyuls sur mer

Urania CHRISTAKI Professeur, Université du Littoral
Rapporteur Côte d'Opale

Fabrice NOT Directeur de recherches CNRS,
Président du Jury Station Biologique de Roscoff

Dominique HERVIOT-HEATH
Examineur Chercheur, Ifremer Centre de Brest

Christine PAILLARD Directeur de recherche CNRS,
Directrice de thèse Institut Universitaire Européen de la
Mer

Loïs MAIGNIEN Maître de conférence, Université de
Co-encadrant de thèse Bretagne Occidentale

Remerciements

Il convient tout d'abord de remercier la région Bretagne, l'UBO et le LabexMER pour avoir financé et permis cette thèse. Merci également aux directeurs du LEMAR, Luis Tito de Morais et du LM2E, Mohammed Jebbar, pour m'avoir accueillie au sein de leurs laboratoires au cours de ces trois années (et un peu plus) de thèse.

Je remercie l'ensemble de mon jury pour avoir accepté d'évaluer ce travail de thèse. Merci à Pierre Galand et Urania Christaki d'avoir endossé le rôle de rapporteur ainsi que Fabrice Not et Dominique Herviot-Heath, celui d'examineur. Vos remarques et votre expérience me seront précieuses.

Merci également aux membres de mon comité interne, Damien Eveillard et Christian Jeanthon, pour avoir suivi l'avancement de ces travaux tout au long de ces trois ans.

Un grand merci à mes deux encadrants de thèse pour avoir mis en place et porté ce projet à la croisée de deux laboratoires. Merci, Christine, pour ta disponibilité et ta curiosité scientifique inépuisable, qui sont d'une grande motivation. Je me souviendrai particulièrement de mon initiation à la pêche aux palourdes dans la vase de la Rade de Brest, qui a un charme différent de la pêche aux bactéries dans l'eau de mer... surtout quand on peut les déguster ensuite ! Loïs, merci de m'avoir fait confiance depuis si longtemps déjà, pour mener à bien ces projets ambitieux, mais passionnants en écologie microbienne. Merci pour m'avoir initiée à ce monde complexe avec une grande patience et surtout merci pour ta disponibilité et ton écoute tout du long de la thèse. C'était une chance de t'avoir en encadrant de thèse.

Cet observatoire microbien n'aurait pas pu se faire sans le soutien précieux de plusieurs personnes. Tout d'abord Adeline, pour m'avoir aidée au tout début de sa mise en place, dans l'achat du matériel de tous les tuyaux, bidons, pompes, filtres, dont les références étaient parfois compliquées à trouver. Une fois le matériel acheté, ce sont des centaines de litres d'eau de mer filtrés et tout autant de PCR/migration/purification/dosages par la suite. Je souhaiterais dire un immense merci à toi, Morgan, pour ton aide indispensable et ta motivation sans faille dans toutes ces manips. C'était un vrai plaisir de travailler à tes côtés et j'espère avoir l'occasion de retravailler avec toi un jour. Enfin, Johanne, merci d'avoir porté cette analyse métagénomique dans les derniers mois de la thèse. Grâce à ton expérience, ton efficacité multitâche, et ta patience sur le refinement manuel des bins, on a pu retrouver le fameux génome d'Amylibacter. Merci !

Je voudrais également remercier les personnes du suivi SOMLIT, Émilie, Peggy, Jérémy, pour leur bonne humeur lors de ces prélèvements tôt le matin, voir pire, tôt le dimanche matin ! Merci beaucoup également, Johann, pour nos discussions sans fin, à essayer de démêler les fils entre le bazar des bactéries marines et le désordre de la matière organique dissoute...on y arrivera un jour, j'en suis sûre !

A toute l'équipe des campagnes M2BiPAT, Frank, Laurent Memery, Louis Marié, Aude Leynaert, Marc Sourisseau, Raffaele Siano, Pierre Ramond, Sophie Schmitt, Mathilde Cadier. Merci pour ces campagnes intenses mais remplies de bons souvenirs. Merci également pour vos discussions sur vos thématiques respectives, qui m'ont aidée à comprendre un peu mieux le système complexe du front de Ouessant. Merci également à Stéphane l'Helguen, pour ces discussions allant du cycle de l'azote en mer d'Iroise aux choix des tablettes de chocolat !

Merci aux scientifiques et membres d'équipage de l'Akademik Tryoshnikov pour cette belle expérience entre le Cap et l'Allemagne sur fond de Russie.

I want to say a particular thank to Antonio, for your kind help in the network scripts and the manuscrit review. Merci aux membres du LS2N à Nantes pour m'avoir super bien accueillie pendant un temps dans ce monde des réseaux et des mathématiques. Damien, merci pour ces discussions passionnantes et tes conseils sur l'analyse des données M2BiPAT. A Florent pour avoir été un hôte plus que cool pendant cette parenthèse Nantaise. Un grand merci !

Aux membres des deux laboratoires, chercheurs des doctorants. Merci, les doctorants et post-doc du LEMAR pour votre bienveillance et votre sympathie ! Merci François pour ces soirées jeux, qui vont reprendre bientôt, promis ! Merci à tous les membres du LM2E que j'ai pu côtoyer pendant toutes ces années : Alexandre, Claire, Erwan, Karine, Mohammed, Morgane, Nadège, Sophie, Stéphane, Yann. Un merci particulier à Stéphanie, pour ces discussions et ton écoute précieuse, surtout vers la fin de la thèse ! Aux doctorants du LM2E à commencer par les anciennes générations : Matthieu et sa chevelure folle, Charlène, Coraline, Mélanie, Simon, Vincent et Aline, merci pour ces bons moments passés hors des labos. Un immense merci aux doctorants de ce super cru 2015-2018, voir 2019 pour les plus inspirés. Caroline, Florian, Seb, Yang et aux futurs docteurs (ça arrive plus vite qu'on ne le pense) : Blandine, Jordan, Marc, Sarah. Merci infiniment à vous pour ces pauses café, ces tarots, ces discussions, votre bonne humeur qui ont rendu ces dernières années de thèses super chouettes. Merci Ashley pour réussir à comprendre mon anglais le matin alors que je ne suis pas réveillée, et ces échantillonnages inopinés à Ste Anne. Enfin merci et bon courage aux nouveaux venus, Jie et Maxime.

Merci aux stagiaires, qui ont participé aux travaux de cet observatoire. Julien, le tout premier, pour avoir accepté sans trop sourcillé de refaire toute une plaque de banque à quelques semaines de la fin...hum ! Maïly, pour ton humour et ta motivation sans borne pour enchaîner les plaques de qPCR et enfin merci à Malo, le stagiaire qui analysait des données 16S plus vite que son ombre.

Je remercie également les meilleures collègues de bureau qu'il soit : Ewan, merci pour ces pauses thés. Ton humour et ton flegme anglais m'ont beaucoup manqué pendant la dernière année de la thèse. Alexandra, un grand merci pour toutes ces attentions, ton amitié et ton aide pour le côté administratif qui (parfois) me dépasse ! Je te souhaite plein de courage pour cette dernière ligne droite. Merci à Sonia pour ces longues discussions à refaire le monde qui m'ont été précieuses. Merci à Marianna et Gabriella pour votre bonne humeur, vos accents sud-américains et le parfum du maté dans le bureau qui font voyager.

Un merci immense à mes amis, les intechos, qui n'ont pas désespéré de mes...comment dire...difficultés chroniques à me servir de la technologie du téléphone. Hélène, Cathy, Gaëlle, Marie, je vous aime fort, promis je viens vous embêter bientôt ! Anne, je suis avec toi de tout cœur pour ta thèse en sciences non brutes ! C'est fou quand même, on n'aurait jamais parié de se retrouver là, toutes les deux, il y a 10 ans ...

Merci infiniment, Anne et Olaf, pour votre amitié qui m'est précieuse, et pour m'avoir offert de travailler l'introduction de la thèse dans ce bout de Normandie qui m'est si cher.

Un grand merci à ma famille, mes parents leur soutien et leur amour indéfectible. Merci à mes frère et sœurs qui n'ont pas toujours fui quand je parlais bactérie marines et que j'aime énormément. A Mamina, pour m'avoir fait des bons petits plats tout le début de la thèse.

Enfin, merci à toi, Florian, car rien de tout ça n'est possible sans toi.

TABLE OF CONTENTS

GENERAL INTRODUCTION	1
1. MARINE MICROORGANISMS	2
1.1. <i>Marine bacteria: a recent story</i>	2
1.2. <i>Key role of planktonic bacteria in marine ecosystems</i>	3
1.2.1. Role in primary production	4
1.2.2. Role in Organic Matter recycling and the food web: the microbial loop.....	6
1.2.3. Role in global biogeochemical cycles.....	8
1.2.3.1. Carbon	8
1.2.3.2. Nitrogen.....	8
1.2.3.3. Sulfur.....	9
1.3. <i>The revolution of molecular tools in marine microbial ecology</i>	10
1.3.1. Accessing diversity using 16S sequencing.....	11
1.3.1.1. Methods and limits	11
1.3.1.2. An unsuspected diversity in the ocean.....	13
1.3.2. Metagenomics.....	15
1.3.2.1. Methods and limits	15
1.3.2.2. New capacities discovered.....	16
1.4. <i>Marine microbial ecology based on DNA data</i>	17
1.4.1. Bacterial communities characteristics.....	18
1.4.1.1. Free-living and particle-attached bacteria.....	18
1.4.1.2. Rare biosphere of sea water.....	19
1.4.2. Bacterial communities dynamics	19
1.4.2.1. Temporal dynamic	19
1.4.2.2. Spatial dynamic.....	21
1.4.2.3. Everything is everywhere	23
1.4.3. Marine bacteria trophic strategies	24
2. THE IROISE SEA AND THE BAY OF BREST.....	27
2.1. <i>Coastal oceans: a crucial but fragile ecosystem</i>	27
2.2. <i>The Iroise sea and the Ushant Front</i>	28
2.3. <i>The Bay of Brest and the SOMLIT station</i>	30
OVERVIEW AND OBJECTIVES.....	35
CHAPTER I SPATIO-TEMPORAL DYNAMIC OF BACTERIAL COMMUNITIES IN THE IROISE SEA	37
1. RESUME EN FRANÇAIS.....	38
2. ABSTRACT	40
3. INTRODUCTION	41
4. RESULTS	43
4.1.1. Environmental settings	43
4.1.2. Bacterial community dynamics in the Iroise Sea.....	45
4.1.3. Description of modules with a specific dynamic, diversity and correlation with environmental parameters.....	47

5.	DISCUSSION.....	51
6.	CONCLUSION.....	56
7.	MATERIAL AND METHODS.....	57
7.1.	<i>Study site and sampling design.....</i>	57
7.2.	<i>Nutrients, phytoplankton counts and pigments analysis.....</i>	57
7.3.	<i>Bacterioplankton communities sampling.....</i>	58
7.4.	<i>DNA extraction and sequencing.....</i>	58
7.5.	<i>Bioinformatics analysis.....</i>	59
7.6.	<i>Statistical analysis.....</i>	59
7.7.	<i>Network analysis.....</i>	59
8.	SUPPLEMENTARY DATA	61
CHAPTER II. TEMPORAL DYNAMIC OF BACTERIAL COMMUNITIES IN THE BAY OF BREST....		75
1.	RESUME EN FRANÇAIS.....	76
2.	ABSTRACT	78
3.	INTRODUCTION	79
4.	MATERIAL AND METHODS.....	81
4.1.	<i>1. Study site and sampling strategy.....</i>	81
4.2.	<i>2. DNA extraction and sequencing for bacterial diversity.....</i>	82
4.3.	<i>3. Bioinformatics analysis.....</i>	82
4.4.	<i>4. Statistical analysis</i>	83
4.5.	<i>5. Looking for general patterns of seasonality</i>	83
5.	RESULTS	85
5.1.	<i>Environmental settings</i>	85
5.2.	<i>Bacterioplankton community's characteristics</i>	85
5.3.	<i>Seasonality of bacterial communities.....</i>	88
6.	DISCUSSION.....	94
6.1.	<i>Seasonal dynamics of bacterial communities.....</i>	94
6.2.	<i>Break of seasonality: spring blooms</i>	94
6.3.	<i>Late summer communities.....</i>	96
6.4.	<i>Multiple possible drivers in winter.....</i>	96
7.	CONCLUSION.....	98
8.	SUPPLEMENTARY DATA	99
CHAPTER III. INTO THE GENOME OF AN ABUNDANT RHODOBACTERACEAE.....		103
1.	RESUME EN FRANÇAIS.....	104
2.	ABSTRACT	106
3.	INTRODUCTION	107
4.	MATERIAL AND METHODS.....	108
4.1.	<i>Study site and sampling</i>	108

4.2.	<i>Libraries preparation for metagenomes</i>	108
4.3.	<i>Reconstruction of Metagenome Assembled Genomes (MAGs)</i>	108
4.3.1.	<i>Sequences quality filtering and assembly</i>	108
4.3.2.	<i>Binning and MAG characterization</i>	109
4.3.3.	<i>Phylogenomic</i>	110
1.	RESULTS AND DISCUSSION	111
1.1.	<i>Recovery of MAGs and identification of the dominant MAG affiliated to the Rhodobacteraceae family</i>	111
1.2.	<i>Genomic characteristics of MAG ID03_00003</i>	111
1.3.	<i>Gene content and putative metabolism</i>	115
1.4.	<i>Taxonomic affiliation and Phylogeny</i>	118
1.5.	<i>Importance of MAG ID03_00003 in the Bay of Brest and the Iroise sea</i>	119
2.	CONCLUSION	120
3.	SUPPLEMENTARY DATA	121
	DISCUSSION AND PERSPECTIVES	125
1.	THE BREST BAY TIMES SERIES IN THE CONTEXT OF COASTAL GENOMIC OBSERVATORIES	126
1.1.	<i>How to bridge the gap between physical and chemical oceanography and microbial ecology in sub-mesoscale structures?</i>	129
2.	FROM ABIOTIC TO BIOTIC DRIVERS	133
3.	A PERSPECTIVE: BETTER UNDERSTANDING OF ANTHROPIC PRESSURES ON COASTAL ECOSYSTEMS	134
	GENERAL CONCLUSION	137
	REFERENCES	138

LIST OF FIGURES

FIGURE 1 GLOBAL DISTRIBUTION OF (A) PROCHLOROCOCCUS AND (B) SYNECHOCOCCUS MEAN ANNUAL ABUNDANCE AT THE SEA SURFACE.	5
FIGURE 2 GLOBAL VIEW OF THE PELAGIC FOOD WEB.....	7
FIGURE 3 GRAPH SHOWING THE AUGMENTATION OF PAPER CITATIONS FOLLOWING SEQUENCING TECHNIQUES EVOLUTION IN THE FIELD OF MICROBIOLOGY.....	12
FIGURE 4 MAP INDICATING THE DIFFERENT SAMPLING STATIONS OF THE EXPEDITION TARA OCEAN.....	22
FIGURE 5 LOCALIZATION OF THE IROISE SEA AND SATELLITE SEA SURFACE TEMPERATURE HIGHLIGHTING THE PRESENCE OF THE USHANT FRONT.....	29
FIGURE 6 MAP OF THE BAY OF BREST.	31
FIGURE 7 TIDE CURRENTS IN THE BAY OF BREST.....	32
FIGURE 8 ENVIRONMENTAL AND BACTERIAL COMMUNITIES CHARACTERISTICS OF THE IROISE SEA FOR THE DIFFERENT CRUISES.	44
FIGURE 9 VISUALIZATION OF THE DIFFERENT MODULES AND THEIR MAIN CHARACTERISTICS.....	46
FIGURE 10 MODULES DYNAMIC ACROSS THE DIFFERENT CAMPAIGNS, STATIONS AND DEPTH.....	48
FIGURE 11 MODULES TAXONOMIES AND COEFFICIENT OF VARIATION OF THEIR OTUS A.	50
FIGURE 12 SEASONAL VARIATIONS OF THE ENVIRONMENT IN THE BAY OF BREST.....	86
FIGURE 13 SEASONALITY OF THE BACTERIOPLANKTON COMMUNITIES.	87
FIGURE 14 VARIATION OF THE DIFFERENT MODULES OVER ONE YEAR (2016).....	88
FIGURE 15 TAXONOMY COMPOSITION OF THE DIFFERENT MODULES.	89
FIGURE 16 VARIATION OF THE DOMINANT OTUS OF THE MAIN MODULES ACROSS THE ENTIRE SURVEY.	91
FIGURE 17 VARIATION OF OTUS FROM THE SPRING MODULE (MODULE 24).....	93
FIGURE 18 SAMPLING POINTS AND DISTRIBUTION OF MAG ID03_00003 ACROSS THE SAMPLES.....	113
FIGURE 19 PHYLOGENOMIC AND TAXONOMIC ANALYSIS OF MAG ID03_00003.....	118
FIGURE 20 LOCALIZATION OF THE DIFFERENT STATIONS SAMPLED DURING THE OSD CONSORTIUM THE 21 ST OF JUNE 2014.	126
FIGURE 21 NON-METRIC DIMENSIONAL SCALING ORDINATION BASED ON THE BRAY-CURTIS DISSIMILATORY INDEX FOR THE BACTERIAL COMMUNITIES FOUND AT THE DIFFERENT STATIONS SAMPLED DURING OSD 2014.....	128
FIGURE 22 VARIATION IN QUANTITY OF DISSOLVED ORGANIC MATTER.....	132

LIST OF TABLES

TABLE 1 MAIN TIME SERIES OF OCEANIC MICROBIAL (BACTERIA AND ARCHAEA) COMMUNITIES.....	20
TABLE 2 MAIN PATHWAY AND GENES PRESENT IN DIFFERENT GENOMES OF RHODOBACTERACEAE.....	114

LIST OF SUPPLEMENTARY FIGURES

FIGURE S 1 CTD PROFILE OF THE DIFFERENT STATIONS FOR ALL THE CRUISES. VALUES WERE AVERAGED EVERY 5 METERS. FLUO: FLUORESCENCE.....	64
FIGURE S 2 GENERAL INFORMATION ABOUT THE CRUISES (A.) AND SATELLITE SURFACE CHLA MEASUREMENTS (B.). THE SHAPES CORRESPOND TO THE DIFFERENT STATIONS SAMPLED DURING EACH CRUISE.....	68
FIGURE S 3 NMDS OF THE BACTERIAL DIVERSITY BASED ON BRAY-CURTIS DISSIMILARITIES AMONG EACH SAMPLING. STRATIFICATION (STATION 5) CLEARLY SEPARATES DISTINCT COMMUNITIES BETWEEN THE SURFACE AND DEPTH. THIS TREND IS LESS OBSERVED IN MIXED WATERS (STATION 2 AND 3) EXCEPT IN JULY. IN MARCH THERE IS NO DISTINCTION OF THE COMMUNITIES WITH DEPTH BUT A CLEAR COASTAL TO OFFSHORE GRADIENT IS CONSPICUOUS	69
FIGURE S 4 MOST SIGNIFICANT BIOMARKERS (LDA > 4) OF A STATION WITHIN EACH CRUISE FOUND WITH LEFSE ALGORITHM. THEIR RELATIVE ABUNDANCE CONSISTED OF THE MEAN OF THE DIFFERENT REPLICATES FOR EACH SAMPLING, WITH THE STANDARD DEVIATION. A. SEPTEMBER 2014, B. MARCH 2015, C. JULY 2015, D. SEPTEMBER 2015.....	71
FIGURE S 5 DOMINANT OTUS FOUND WITHIN EACH MODULE, WITH THEIR TAXONOMY AND RELATIVE ABUNDANCE (IN % OF THE READ) IN THE MODULE.	72
FIGURE S 6 PLOT OF THE COEFFICIENT OF VARIATION (CV) OF AN OTU WITH ITS OCCURRENCE IN A SAMPLE (I.E. RELATIVE ABUNDANCE >0.05%). THE LINEAR MODEL IS THE BLUE LINE AND THE PREDICTION INTERVAL (0.5) ABOVE AND UNDER THE LINEAR MODEL IS IN GREY. HIGH CV OTUS ARE THOSE ABOVE THE PREDICTION INTERVAL AND LOW CV OTUS ARE THOSE BELOW. EACH OTU IS COLORED DEPENDING ON ITS MODULE MEMBERSHIP.	73
FIGURE S 7 CYTOMETRY COUNTS FOR TOTAL BACTERIA AND <i>SYNECHOCOCCUS</i> AT SOMLIT STATION IN THE BAY OF BREST, BETWEEN 2014 AND 2018.	99
FIGURE S 8 VARIATION IN ALPHA DIVERSITY OF THE BACTERIAL COMMUNITIES MEASURED WITH SHANNON INDEX.	99
FIGURE S 9 RELATIVE ABUNDANCE OF THE MOST ABUNDANT OTUS; I.E. WITH A RELATIVE ABUNDANCE UPPER THAN 1% IN MORE THAN 75% OF ALL THE SAMPLES.....	100
FIGURE S 10 PEARSON CORRELATION OF EACH MODULE EIGENVECTOR WITH THE ENVIRONMENTAL VARIABLES.	101
FIGURE S 11 HIERARCHICAL CLUSTERING WITH BRAY-CURTIS DISTANCES BASED ON K-MER FREQUENCY WITH SIMKA BETWEEN SAMPLES. GROUPS OF CO-ASSEMBLY USED IN THIS STUDY ARE DEFINED BY RED CIRCLE...	121
FIGURE S 12 30 MOST ABUNDANT MANUALLY REFINED, NON-REDUNDANT AND HIGH QUALITY MAGS WITH THEIR CHECKM AFFILIATION. THEY WERE DEFINED ABUNDANT BASED ON THEIR CUMULATIVE PERCENTAGE OF READ RECRUITMENT ACROSS THE SAMPLES.....	124

LIST OF SUPPLEMENTARY TABLES

TABLE S 1 ADDITIONAL DATA FOR EACH SAMPLING OF M2BIPAT.	65
TABLE S 2 STATISTICS OF THE CONTIGS WITHIN EACH CO-ASSEMBLY GROUP.	122
TABLE S 3 CHARACTERISTICS OF THE DIFFERENT REDUNDANT BINS OF MAG ID03_00003.	123
TABLE S 4 COMPARISON OF SOME SIGNIFICANT COG CATEGORIES LINKED WITH TROPHIC LIFESTYLE ACCORDING TO LAURO ET AL., 2009.	123

LIST OF ABBREVIATIONS

DNA	Deoxyribonucleic Acid
OM	Organic Matter
DOM	Dissolved Organic Matter
POM	Particulate Organic Matter
AOB	Ammonia Oxidizing Bacteria
NOB	Nitrite Oxidizing Bacteria
AOA	Ammonia Oxidizing Archaea
DMSP	Dimethylsulfoniopropionate
rRNA	Ribosomal Ribonucleic Acid
NGS	Next Generation Sequencing
PCR	Polymerase Chain Reaction
MAG	Metagenome-Assembled Genome
SCG	Single-copy Core Gene
FL	Free-Living
PA	Particle-Attached
COG	Clusters of Orthologous Groups
POC	Particulate Organic Carbon
PON	Particulate Organic Nitrogen
Chla	Chlorophyll a
OTU	Operational Taxonomic Unit
ME	Module Eigengene
TMA	Trimethylamine
ANI	Average Nucleotide Identity
TMAO	Trimethylamine N-Oxide

GENERAL INTRODUCTION

1. MARINE MICROORGANISMS

1.1. MARINE BACTERIA: A RECENT STORY

The presence of microorganisms in marine environment has been shown since the beginning of microbiology at the end of the 17th century by Antonie van Leeuwenhoek, a Dutch trader. He observed "animacules" using microscopes of his own design capable of enlarging a sample up to 300 times.

Subsequently, the first microbiological research in the oceans took place during the oceanographic missions of the French ships, *Le Travailleur* and *Le Talisman* in 1880 and 1883. These expeditions were truly pioneering in the study of marine bacteria and the samples collected allowed the discovery of new species, isolated in Louis Pasteur's laboratories in Paris. Since then, knowledge of marine bacteria has greatly increased. Thus, in the first half of the 20th century their contribution in the recycling of organic matter and in different biogeochemical cycles (e.g. sulphur and nitrogen cycles) was known (Vernadsky, 1945; ZoBell, 1946). Strangely, at that time, the issue of exclusive marine bacteria was still debated among microbiologists, given their morphological and biochemical similarities with terrestrial bacteria. It was not until the 1960s that the endemic nature of marine bacteria was definitely recognized (MacLeod, 1965).

At the same time, microbiologists realized that culture techniques could only study a very small minority of the bacteria present in a given environment. This was defined as "Great Plate Count Anomaly" and it was found that the bacterial abundances observed under microscopes were much higher than those obtained by culture methods. It is estimated that less than 1% of microorganisms in the environment are potentially "cultivable" with standard plating techniques (Staley and Konopka, 1985). Consequently, the main discoveries in microbiology in recent decades strongly relied on key technological advances capable of overcoming the difficulties inherent in the study of these organisms invisible to the naked eye.

For example, determining precisely how many bacteria are contained in a sample of seawater was only possible in the late 1970s with the emergence of epifluorescence microscopy techniques and DNA staining using fluorescent markers (Hobbie et al., 1977) as well as the invention of flow cytometry and its use on seawater samples (Yentsch et al., 1983). These techniques, still used, have demonstrated that seawater at ocean surface and coastal level contains between 100 000 and 1 000 000 bacteria per milliliter of seawater (Porter et al., 1996). This number is, surprisingly, fairly constant in all the oceans of the world, with variations generally not exceeding a factor of 10 (Fuhrman and Caron, 2016). We would thus have an estimated total of 10^{29} bacteria in the first 100 meters of surface of all the oceans (Whitman et al., 1998). Which is more than the number of stars in the visible universe (estimated at 10^{21}). This dizzying number suggests that these microorganisms, however small they may be, are closely linked to the functioning of marine ecosystems.

1.2. KEY ROLE OF PLANKTONIC BACTERIA IN MARINE ECOSYSTEMS

“Basic to the understanding of any ecosystem is the knowledge of its food web, through which energy and material flow.” (Pomeroy, 1974). In a simplistic view, a food web represent a transfer of organic matter (OM) synthesized by the primary producers to the rest of the living organisms through complex pathways. For a long time, small organisms (viruses, bacteria or protozoa) were neglected in the study of OM processing in marine ecosystems as they were thought to be of minor contribution (Steele, 1974). However, advances in microbiology with the possibility to enumerate bacteria, measure their biomass and estimate their production have shown that bacteria are, on the contrary, key actors of the food web. They are mainly involved in the production and degradation of OM and through this to the fate of major chemical elements (e.g. nitrogen, sulphur) in the ocean. This recognition that the smallest members of the food web (*i.e.* viruses, bacteria, phytoplankton or protists) were probably responsible for a large fraction of important system activities represent one of the major breakthroughs in the field of marine science in the last decades (Azam et al., 1983; Fuhrman and Caron, 2016; Pomeroy, 1974).

1.2.1. ROLE IN PRIMARY PRODUCTION

In marine surface environment, primary producers are composed of photosynthetic microorganisms: microalgae and cyanobacteria. Thanks to the photosynthesis reaction and sunlight as a source of energy, they are able to produce complex organic molecules principally from inorganic nutrients (*i.e.* NO_3^- or PO_4^{2-}) and carbon (CO_2). Despite their small size, they account for nearly 50% of global net primary production in our planet and are able to supply the complex trophic chain in the oceans (Field et al., 1998; Katz et al., 2004). Cyanobacteria represent the most abundant and ancient photosynthetic microorganisms found on earth (Hedges et al., 2001). Contrary to the widely diverse eukaryotic photosynthetic organisms, marine planktonic cyanobacteria are composed of two dominant genus : *Prochlorococcus* and *Synechococcus* that numerically dominate most phytoplankton assemblage in the oceans (Scanlan et al., 2009). These two taxa have a different global repartition (Figure 1). *Prochlorococcus* species represent the dominant cyanobacteria of the oligotrophic and warm tropical waters while *Synechococcus* members are ubiquitous and dominates in higher latitudes and nutrient-rich environments such as temperate coastal waters (Flombaum et al., 2013; Olson et al., 1990).

Marine cyanobacteria contribute from 10 to 25% of the net primary production in the ocean (Flombaum et al., 2013; Gregg and Rousseaux, 2016). But this value will depend on the environment. In oligotrophic waters, these small cells present competitive advantages compared to large eukaryotic cells and they sometimes are found to represent almost half of the phytoplankton biomass. In these regions, they can account for up to 48% of the net primary production (DuRand et al., 2001; Vaulot et al., 1995). In more eutrophic waters *Synechococcus* competes with the larger eukaryotic phytoplankton (as diatoms or dinoflagellates) which usually dominates in biomass, activity and abundance. In these waters, relative contribution to primary production of cyanobacteria is less important but still relevant. For example, it was found to be about 6% in a neritic front of the Atlantic Ocean (Glover et al., 1986).

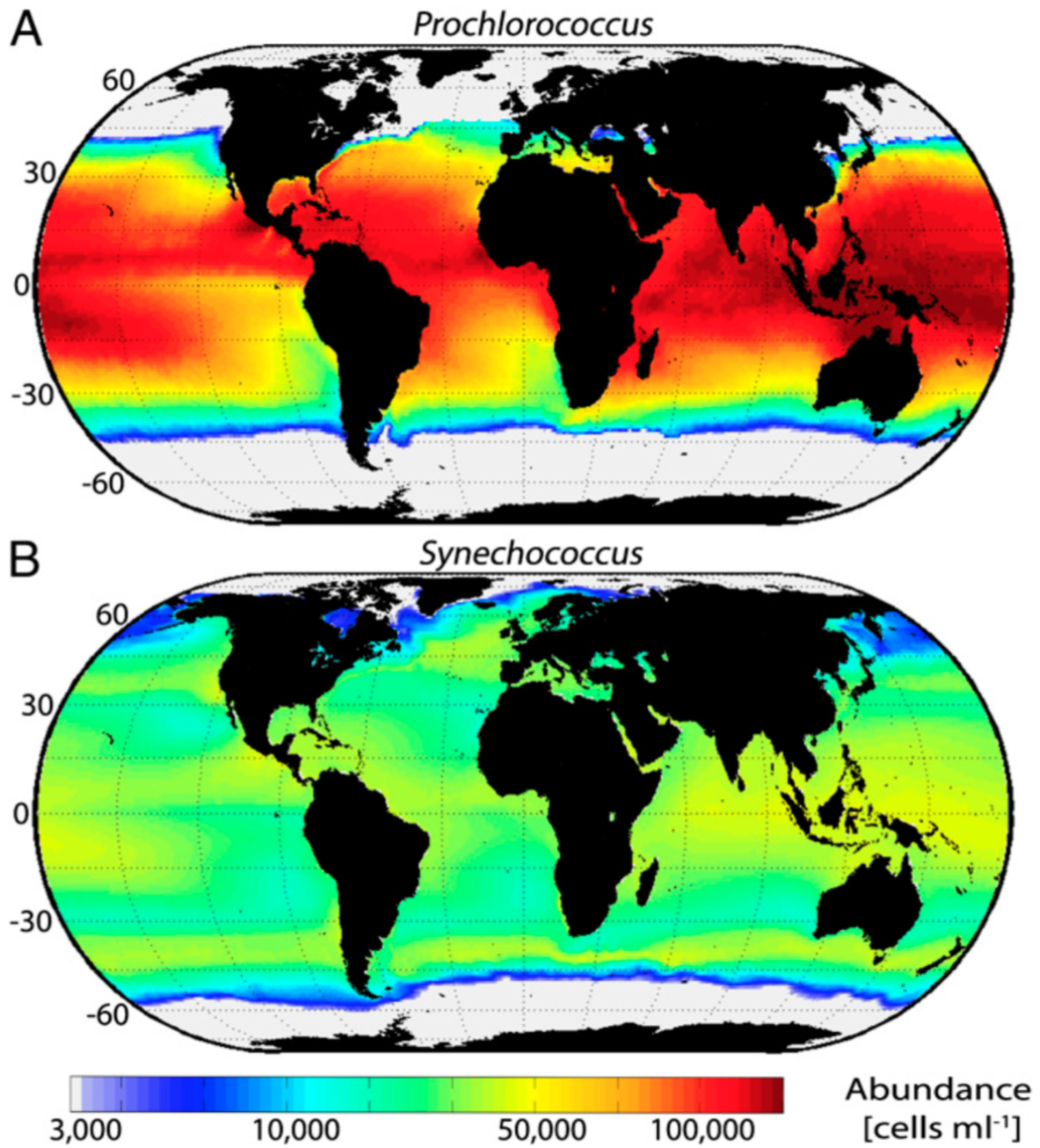


Figure 1 | Global distribution of (A) *Prochlorococcus* and (B) *Synechococcus* mean annual abundance at the sea surface. (Flombaum et al., 2013).

1.2.2. ROLE IN ORGANIC MATTER RECYCLING AND THE FOOD WEB: THE MICROBIAL LOOP

Possibly one of the most recognized roles of bacteria in marine ecosystems is the processing of OM. It is estimated that about a half of the total primary production in the oceans is consumed by microbial heterotrophs, making them one of the major pathways for organic carbon fluxes in the food web (Azam, 1998). These heterotrophic bacteria are osmotrophs and they mainly feed on Dissolved Organic Matter (DOM) which consist of a complex mixture of molecules (*i.e.* lipids, proteins, various metabolites or carbohydrates) that solubilized in the water column (Baines and Pace, 1991). Phytoplankton represent one of the main sources of DOM which account for 10% to 30% of their global production in OM (Fenchel and Jørgensen, 1977). Bacteria are also able to directly attack the Particulate Organic Matter (POM) using hydrolytic exoenzymes in order to liberate DOM (Azam et al., 1994).

The importance of bacteria in OM fluxes have major consequences in marine ecosystems (Azam, 1998). Among other, in the grazing food chain: bacteria represent prime preys for small protozoa, which leads a part of the DOM to re-enter the classical food chain (Figure 2). This phenomenon has been named the microbial loop (Azam et al., 1983). Through this microbial loop, organic matter is remineralized, both by the action of bacteria, and by the protozoa that release large amounts of ammonium and phosphate as waste. They thus accelerate the mineralization of inorganic nutrient making them available again for the primary producers (Caron et al., 1988; Lomas et al., 2002). After its first description, the importance of heterotrophic bacteria as a link between DOM and higher trophic levels has long been debated (Fenchel, 2008). Later studies supposed that the major part of the carbon they assimilate is dissipated as CO₂, and a few will be used to create biomass and then possibly enter the higher trophic levels (Williams and Martinez, 2000).

In fact, the classic microbial loop (DOM → Bacteria → Protozoa → Higher trophic levels) appeared to be much more complex than we thought (Sherr and Sherr, 2002). The primary production consumed within the bacteria follows complex cycles of respiration, mineralization, with even more actors such as viruses and mixotrophic algae (Suttle, 2007).

The relative importance of the microbial loop in the food web will change depending on the environment (Cotner and Biddanda, 2002). Many studies suggested that the microbial loop is predominant in oligotrophic waters, dominated by small phytoplankton (Legendre and Le Fèvre, 1995). In these waters, small cyanobacteria can be responsible of the largest portion of the total primary production and they represent key preys for small protists. In these conditions, the only link between the biomass of cyanobacteria and the other trophic levels is via the microbial loop (Fuhrman and Caron, 2016). On the contrary, in more eutrophic waters, the eukaryotic phytoplankton is dominant and will be directly grazed by herbivorous while DOM will be remineralized and respired into CO_2 via the bacteria. This will have direct consequences on the ecosystem: in oligotrophic waters, bacteria will take up most of the primary production with a little left for the rest of the highest levels of the trophic chain, leading to waters with poor fisheries resources (Sommer et al., 2002; Turley et al., 2000).

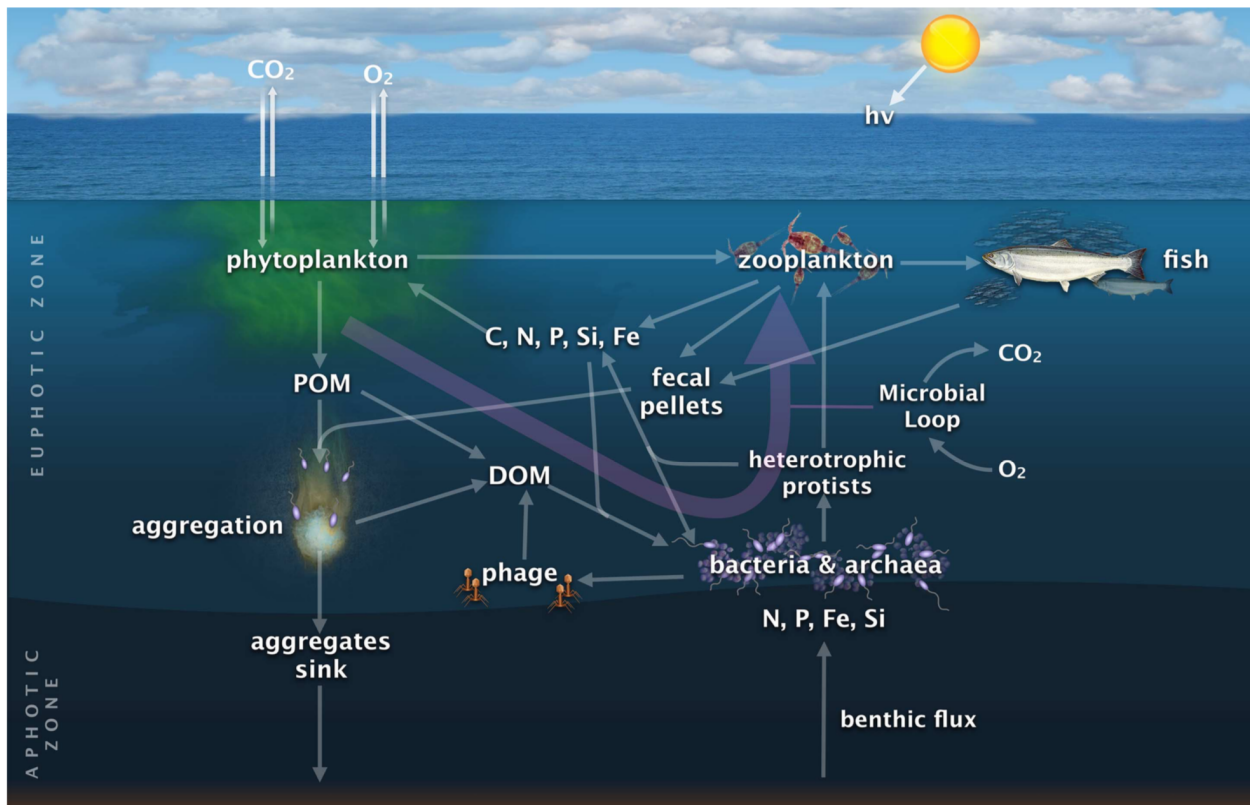


Figure 2 | Global view of the pelagic food web. (Worden et al., 2015).

1.2.3. ROLE IN GLOBAL BIOGEOCHEMICAL CYCLES

Thanks to their abundance, and their large metabolic capacities, microorganisms dominates the flux of energy and biologically important chemical elements (carbon, nitrogen, sulfur and phosphorus) in the oceans. These elements move from biotic to abiotic compartments (and reversibly) in complex pathways called biogeochemical cycles with major implication in the Earth system.

1.2.3.1. CARBON

As seen above in the text, with the microbial loop marine bacteria will structures carbon fluxes in all major pathways : grazing food chain, sinking (CO_2 respiration), carbon storage and carbon fixation (Azam, 1998). The comprehension of carbon biogeochemical cycle is of utmost interest as this element is closely linked to the climate of our planet. Atmospheric CO_2 in particular, regulates the earth's surface temperature and its increasing in concentration is considered to be responsible of the Global warming (Matthews et al., 2009). A part of the carbon (CO_2) fixed by photosynthetic organisms will sink as POM in the bottom of the oceans where it can be sequestered within the sediments for centuries. This process is called the biological carbon pump and permits to remove a consistent part of atmospheric CO_2 from the atmosphere (Ducklow et al., 2001; Riebesell et al., 2009). Moreover, marine DOM is also supposed to be one of the most important Organic Carbon reservoir on earth, with quantities equivalent to the atmospheric reservoir of CO_2 or land plant biomass (Azam and Worden, 2004; Hedges et al., 1992; Jiao et al., 2010). Thus, the way heterotrophic bacteria degrades POM and DOM will actively control the fraction of carbon that will stay in the oceans or return back in the atmosphere through respiration (Mitra et al., 2014).

1.2.3.2. NITROGEN

Nitrogen in the ocean is considered as one of the major limiting factors for primary production (Tyrrell, 1999). It present a variety of chemical forms (*i.e.* N_2 , NH_4 , NO_2 , NO_3 , N_2O , urea) that can be rapidly converted by microorganisms. The main reservoir of Nitrogen is the dinitrogen gas (N_2) present in the atmosphere but it is not usable in this form to most organisms. Some rare diazotrophs among bacteria

and archaea can convert it into Ammonium (NH_4^+) that represent an important source of nitrogen for most primary producers. In marine environment, the cyanobacteria are among the most important diazotrophs, mainly in oligotrophic, N-limited waters (e.g. Western tropical south pacific) (Bonnet et al., 2017). In this oligotrophic water, N_2 fixation can contribute to 50% of the primary production. But our knowledge is still limited on the dynamic and intensity of this diazotrophy (Moisander et al., 2017). Moreover, in coastal waters the main source of nitrogen is from the terrestrial runoff or upwelling of deep-waters where it is present in large quantities.

The regeneration of organic nitrogen to ammonium or nitrates is a fundamental process to maintain the primary production in summer. Organic nitrogen when degraded, release NH_4^+ that will rapidly be nitrified under aerobic conditions to form Nitrates (NO_3). All this process is catalyzed by bacteria or archaea (Zehr and Ward, 2002) and occurs along the water column. It was long admitted that the entire nitrification could be done by two different groups of proteobacteria: ammonium-oxidizing bacteria (AOB : $\text{NH}_4^+ \rightarrow \text{NO}_2$) with the genus *Nitrosomonas/Nitrosospira* and nitrite-oxidizing bacteria (NOB : $\text{NO}_2 \rightarrow \text{NO}_3$) with genus such as *Nitrospina*. But in reality, some abundant pelagic archaea, the Ammonium oxidizing Archaea (AOA) are able to oxidize ammonium in nitrites and probably in a more important way than the AOB (Wuchter et al., 2006). These archaea are members of the phylum Crenarchaota and are 10 to 10000 times more abundant than AOB. One has been cultivated in 2005 : *Nitrosopumilus maritimus*, an autotrophic AOA (Könneke et al., 2005). Even more, a *Nitrospira* has been found to be able to carry out the whole nitrification from ammonium to nitrates (Kits et al., 2017). All these recent evidences pointed out that the nitrogen cycle in water column is carried out by a complex microbial community with large variety of metabolically versatile microorganisms (Kuypers et al., 2018).

1.2.3.3. SULFUR

Sulfur is a fundamental element of the proteins. It is assimilated in the surface ocean by the phytoplankton as dissolved sulfates mainly to form methionine and cysteine. Some microalgae convert methionine into dimethylsulfoniopropionate (DMSP), a stable form of organic sulfur that is used as an osmolyte or an antioxidant. Some phytoplankton (i.e. Dinoflagellates) produce a very large amount of

DMSP. A study in the North Atlantic ocean found that the phytoplankton invested 7% of the net primary production in the DMSP synthesis (Simó et al., 2002). The sulfur cycle was extensively studied over the past decades as it was found that the DMS, a volatile product of the DMSP degradation, affects clouds formation and change their albedo, with a possible role on the climate (Charlson et al., 1987) even if this importance is debated (Quinn and Bates, 2011). Some studies demonstrated that this single compound can support 1 to 13% of the bacterial carbon demand in the surface oceans, as well as their need in sulfur (Kiene et al., 2000). Heterotrophic bacteria are the main actors of DMSP degradation and DMS release in the atmosphere (Kiene et al., 2000).

In conclusion, marine bacteria, as tiny as they can be, are fundamental in all the major ecosystems processes with even implication on the global climate. These great findings and theories were discovered in second part of the 20th century with the knowledge acquired from cultured taxa and the possibility to measure microbial biomass and fluxes that passes through them. However, at this time we had totally no idea of the different type of bacteria that compose such an important biological compartment (e.g. their size, specific activity or specific growth rate) and thus, the different mechanisms that contribute to the biogeochemical cycles. It was only in the end of the 20th century and the development of sequencing technologies that their incredible diversity and potential metabolisms became accessible.

1.3. THE REVOLUTION OF MOLECULAR TOOLS IN MARINE MICROBIAL ECOLOGY

One of the most important technical revolution of the last past years in marine microbiology was the possibility to assess the non-culturable part of microorganisms, using their DNA. For now, this has been the only way to approach the complexity of environmental microbial communities' diversity. Those technologies allowed microbiologists to answer fundamental questions about the ecology of marine microorganisms and get insight the complexity of their natural communities.

1.3.1. ACCESSING DIVERSITY USING 16S SEQUENCING

Traditionally, organisms (prokaryotes, eukaryotes) were classified according to phenotypic features but Carl Woese and others used the small subunit ribosomal RNA gene, also called 16S rRNA, as a phylogenetic marker to compare and reconstruct the tree of life and define the three domains of life that are Bacteria, Archaea and Eukarya (Woese et al., 1990; Woese and Fox, 1977).

1.3.1.1. METHODS AND LIMITS

The first microbial DNA sequences were obtained with the Sanger technique in the end of the 1980s (Pace et al., 1986). The 16S rRNA gene was amplified by PCR then cloned and sequenced. With this fastidious technique, it was possible to obtain around 100 sequences for one sample, thus observing mainly the dominant species. This method has dominated the field of microbiology for 20 years (Figure 3) before the introduction of New Generation Sequencers (NGS). Today, Illumina sequencers can produce 10^7 to 10^9 reads allowing to possibly assess the global microbial diversity of an environmental sample. With NGS, an important period of description of all possible microbial communities started (e.g. diversity, richness, evenness or dynamic), as shown by the inventory of numerous microbiomes : the human microbiome (Turnbaugh et al., 2007), soil microbiome (Fierer and Jackson, 2006) or oceanic microbiome (Sunagawa et al., 2015).

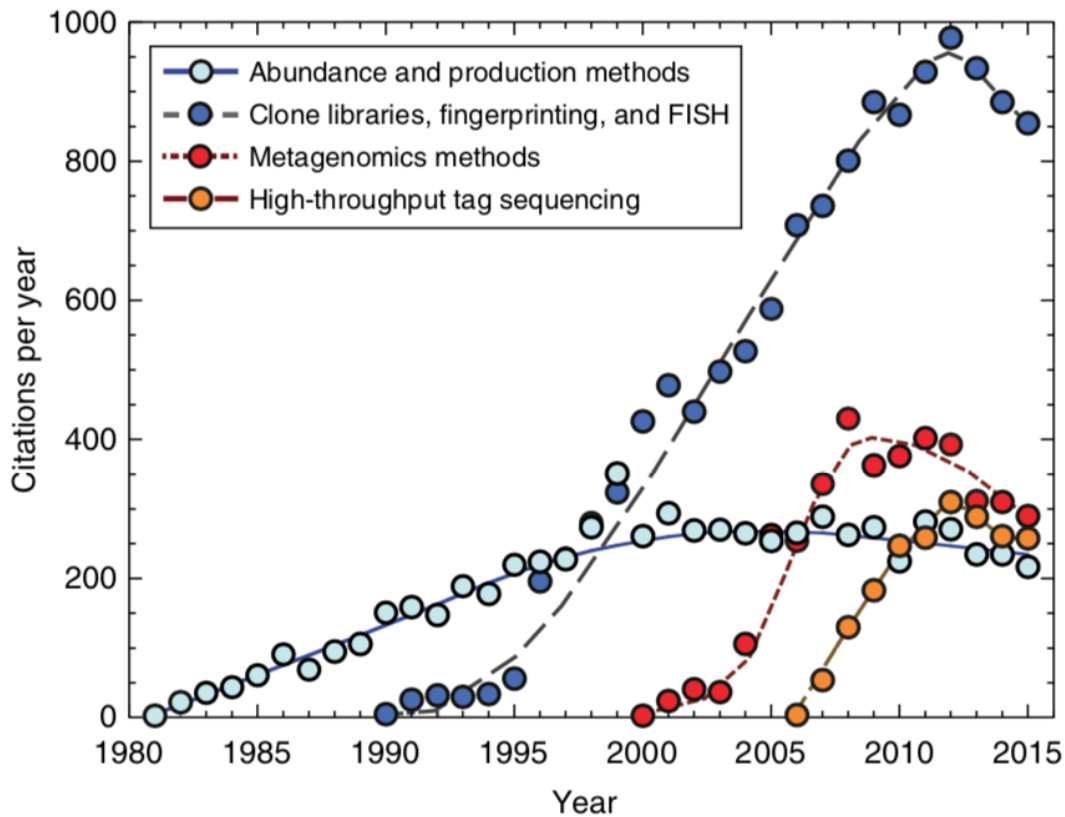


Figure 3 | Graph showing the augmentation of paper citations following sequencing techniques evolution in the field of microbiology. (From *Microbial ecology of the oceans*, 3rd edition, 2018).

The main challenge in 16S analysis is the clustering of sequences to ecologically relevant units. 16S amplicons are clustered into Operational Taxonomic Units (OTUs) to account for intra-species variability, but also the sequencing errors that are inherent to NGS technologies. Traditionally, a minimum similarity of 97% was required to consider two 16S rRNA sequences to belong to the same microbial species (Stackebrandt and Goebel, 1994). But the use of an a priori genetic distance threshold to delineate fundamental units of diversity has however some strong limitations because it is not always applicable or relevant (Brown et al., 2015; Nebel et al., 2011) and limits the resolution of microbial diversity analysis by masking important patterns occurring between closely related organisms (Eren et

al., 2013a). Hence, the recent development of new high resolution methods leveraging sequencing error analysis (DADA2, Callahan et al., 2016), information entropy analysis (oligotyping and MED, Eren et al., 2013, 2014) or error distribution profiles (SWARM, Mahé et al., 2015) constitutes an important advance in microbial diversity analysis. For instance, the Swarm algorithm agglomerated sequences based on pairwise similarity with maximum one nucleotide difference and delineate OTUs based on abundance values – *i.e.* break in abundance « valley » linking two otherwise separate OTUs (Mahé et al., 2015).

The main limits of NGS technologies are the read length (around 2x250 bp) and the amplification step. In one hand, the read length is problematic for the diversity as only 500 bp is often too small to resolve the taxonomy of different species. In the other hand, PCR biases lead to problems of quantification as for instance some taxa are over amplified (Sinclair et al., 2015). Other biases are inherent with the use of 16S rRNA amplicon sequencing, such as the creation of chimeric sequences during the PCR amplification. These artificial sequences are made with the fusion of two (or more) sequences, but are—by definition—lower in abundance than the “parent” sequences and can be identified by software such as VSEARCH (Rognes et al., 2016).

1.3.1.2. AN UNSUSPECTED DIVERSITY IN THE OCEAN

As soon as it was used in marine environment, 16S ribosomal gene sequencing revealed the presence of an important unknown diversity. Its first utilization in 1990 in the Sargasso Sea, detected sequences of *SAR11 Clade*, a completely new clade of the *Proteobacteria* and phylogenetically distant from all cultured organisms known at the time (Giovannoni et al., 1990). Since its discovery, this taxa is now considered to be the most abundant known organism of the planet (Morris et al., 2002). The cultivation of a *SAR11 Clade* member, *Candidatus Pelagibacter ubique* in 2002, showed that it groups small, free-living, heterotrophic bacteria that feed on small organic compounds (Rappé et al., 2002). Added to this, they possess a streamlined genome, that is to say, minimized in size and complexity (Giovannoni et al., 2005b). During this pioneering work in the Sargasso sea and other marine environments, numerous

dominant but unknown lineages were found as the *Clade SAR86*. Unlike *SAR11 Clade*, this taxon has still, almost 30 years later, no representative cultivated.

Another major discovery of 16S ribosomal gene sequencing in marine environment was the presence of marine Archaea, that were thought to be limited to extreme environments at this time (DeLong, 1992; Fuhrman et al., 1992). Marine archaea are in fact highly abundant in the global ocean where they may compose up to 20% of the whole marine picoplankton (DeLong and Pace, 2001). Their phylogeny is far less known and described than bacteria and is composed of two main phyla in marine environment. The most abundant is the *Thaumarchaeota* phylum, which has been revealed to have a great role in the global nitrogen cycle (Brochier-Armanet et al., 2008). The second one is the *Euryarchaeota*, where different groups are found : *Marine Group II* that inhabit surface waters, *Marine Group III* and *Marine Group IV*, which are more abundant in mesopelagic and bathypelagic area (Galand et al., 2009). The ecology of *Euryarchaeota* has only been approached through culture-independent techniques (Haro-Moreno et al., 2017; Ouverney and Fuhrman, 2000).

Thanks to molecular ecology and years of sampling in numerous oceans, we now have a global picture of the microbial diversity in the euphotic and coastal marine environment. There is a total of more than 30,000 species estimated in the ocean microbiome, divided up in 12 bacterial and archaeal phyla (Sunagawa et al., 2015; Yilmaz et al., 2016). In surface waters, most of this diversity is affiliated to three dominant phyla : the *Proteobacteria*, the *Bacteroidetes* and *Cyanobacteria* (Giovannoni and Stingl, 2005; Sunagawa et al., 2015). Overall, the majority of marine bacterial diversity is only assessed by DNA sequencing, making difficult to understand the role of natural assemblages in marine ecosystem (Rinke et al., 2013).

1.3.2. METAGENOMICS

1.3.2.1. METHODS AND LIMITS

The cost-effective high-throughput sequencing allowed for great improvement over metabarcoding: instead of sequencing a marker gene, all genomes present in a sample are fragmented and sequenced. The genetic potential of community can therefore be assessed, and genome populations can be reconstructed into metagenome-assembled genome (MAG) (Nielsen et al., 2014).

New generation sequencers produce short reads of DNA (around 150pb, Illumina HiSeq/NextSeq). As a consequence, the original genomes need to be re-assembled in a process called metagenomic assembly. The product of that assembly is longer fragment of DNA called contigs. The challenge with assembly is that conserved part of genome among different organisms will break the assembly, as well as variations within a species population. To overcome this issue, co-assembly can be performed: DNA sequences from different samples are assembled together to increase the chance of recovering low abundance microorganism's genome.

The binning process is the classification of contigs into discrete clusters in which a bin represent a genome. It can be done automatically using computational tools such as CONCOCT (Alneberg et al., 2014), MetaBAT (Kang et al., 2015), BinSanity (Graham et al., 2017) or Maxbin2 (Wu et al., 2016). The method is based on the sequence composition of contigs (tetramer frequencies, Teeling et al., 2004) and their differential coverage across samples (Albertsen et al., 2013). But automatic methods are not perfect and can result in conflation errors (more than one genome in a bin) or fragmentation errors (one genome in multiple bins), which can be solved by manual curation of the binning results with software like Anvi'o (Eren et al., 2015). The single-copy core genes (SCG), which are expected to be in all microorganisms and only once, are used to assess the completion and redundancy of a bin, and therefore the quality of the binning. Because some microorganisms can lack some of these genes and/or have multiple copies, there is an accepted margin of 90% completion and 5% redundancy to characterize a MAG as a high quality MAG (Bowers et al., 2017).

The first bias of this approach is the same as for the amplicon sequencing technique: the need for DNA amplification, which can result in chimera sequences and overestimation of abundant sequences. Microbial population with high intra-species variability would be poorly assembled because these variations are likely to break the assembly step. A prime example is the *SAR11 Clade* that often presents poorly recovered in metagenomics studies (Venter et al., 2004). Longer DNA sequences would help overcome this issue as with Nanopore technologies (Loman et al., 2015).

1.3.2.2. NEW CAPACITIES DISCOVERED

One of the major contributions of metagenomics and genomics analysis in microbiology is to give clues on the physiology and ecology of uncultivated taxa. Single Cell genome sequencing and metagenomic analysis on *Euryarchaeota*, from whom no representative has been cultivated yet, revealed that they could be heterotrophs, able to uptake and degrade different organic compounds such as lipids (Iverson et al., 2012; Tully, 2019; Zhang et al., 2015a). Similarly, the genome of a member of the *SAR86 Clade* was sequenced, revealing characteristics similar to members of *SAR11 Clade*: very tiny, free-living heterotroph (Dupont et al., 2012). Moreover, large metagenomics surveys revealed the importance of certain genomic characteristics in the global ocean microbiome, giving clue on essential adaptive traits in this environment. For instance, surface free-living bacteria typically present very small genomes compared to most of the cultivated bacteria (Luo et al., 2014). Cells with streamlined genome, typically represented by *SAR11 Clade*, are the most successful in oligotrophic environments (Swan et al., 2013). Those bacteria present small genomes (ca 1-2 Mb) and metabolisms targeting small, very common, low-molecular weight molecules that require minimal functional complexity. For instance, *SAR11 Clade* compete for small, labile organic matter widely produced by marine organisms as osmolytes, external metabolites (methanol, formaldehyde) or amino acids, that is commonly found in the ambient environment (Sun et al., 2011). This leads to the formulation of the streamline genome hypothesis: very large populations that often experiment nutrient limitations tend to simplify their genome complexity. In addition, genome streamlining is often linked to a requirement of some unusual nutrients for the microorganisms, such as reduced sulfur compounds as they lack some fundamental

pathways (Tripp et al., 2008), giving one possible explanation of why we fail cultivating so many pelagic microorganisms. Another interesting and widespread characteristic of surface marine bacteria found thanks to metagenomic surveys is the large presence of proteorhodopsin coding proteins in their genomes (Béjà et al., 2000) which was totally unsuspected. These proteins are localized in the membrane and can harvest light directly to produce energy. There are many types of proteorhodopsin that are highly widespread in many abundant surface bacterioplankton such as *SAR11 Clade*, *SAR86 Clade* and some pelagic archaea (Pinhassi et al., 2016). Despite its ubiquity in numerous marine bacteria, we don't have yet a complete understanding of how proteorhodopsin serves bacteria metabolisms. For now, only very few publications have shown their importance in enhancing growth under light condition (Gómez-Consarnau et al., 2007).

Metagenomic also allowed to directly investigate functional ecology in the oceans. Based on genes functions, numerous studies found that global functions of an ecosystem were relatively stable in contrast with the high variability of taxonomic composition (Louca et al., 2016; Sunagawa et al., 2015). This led to the hypothesis of functional redundancy; *i.e.* different taxa are able to perform the same function in an ecosystem. But this hypothesis is strongly debated (Galand et al., 2018).

Besides, the main fraction of potential genes recovered by metagenomic surveys have still unknown functions. For instance, the pioneer study of metagenomes by Craig Venter in 2005 found an important unknown genes pool of 1.2 million previously unsequenced gene. More recently, the Tara Oceans metagenomic survey encountered around 40% of the orthologous groups of genes of unknown function (Sunagawa et al., 2015). We are hence just starting to understand and assess the potential role of bacteria in marine ecosystems.

1.4. MARINE MICROBIAL ECOLOGY BASED ON DNA DATA

In a few words, marine microbial ecology aims to understand the structure and dynamic of microbial assemblages in their natural environment. Environmental microbial community are composed of a wide range of distinct bacteria that will present different physiological characteristics (growth rates, optimum

growing conditions, trophic strategies) and resource specificity that are largely unknown. It is a great challenge to understand the underlying mechanisms that shape the microbial communities' distribution. Understanding this dynamic is the first step to go inside the role of marine bacteria and archaea in the functioning of marine ecosystems (Karl, 2007). And more, knowing what factors (biological or environmental) drives microbial diversity is essential to predict their evolution in a context of global climate change an increasing anthropogenic pressure (Doney et al., 2012).

1.4.1. BACTERIAL COMMUNITIES CHARACTERISTICS

1.4.1.1. FREE-LIVING AND PARTICLE-ATTACHED BACTERIA

The first studies of marine bacteria ecology separated pelagic bacteria into two different lifestyle: free-living (FL) or particle-attached (PA). These lifestyles were generally distinguished directly under the microscope or by their ability to pass (or not) through a 3µm filter (Garneau et al., 2009). The majority of pelagic bacteria are free-living, while the rest -10% on average- are attached to different kind of particles such as phytoplankton, zooplankton, organic matter debris or marine snow, microplastics among others (Dussud et al., 2018; Turner, 2015). These two compartment harbour different taxonomic composition (Rieck et al., 2015). Small bacteria such as members of *SAR11 Clade*, *SAR86 Clade* or *Cyanobacteria* are almost only present in the FL fraction, while PA bacteria are dominated by members of *Verrucomicrobia*, *Alteromonadales* and *Bacteroidetes* (Crespo et al., 2013; Mestre et al., 2017).

These two fractions likely present different ecological function. PA bacteria tend to present larger cells, with higher individual uptake rates of organic compounds (e.g. glucose) and higher hydrolysis rates suggesting a possible distinct role in the ecosystems and notably in organic matter processing (Iriberry et al., 1987; Middelboe et al., 1995; Unanue et al., 1992). They could be important for the degradation of Particulate Organic Matter and its fate in the deep ocean (Azam and Long, 2001). Adding to this, PA bacteria present a complex role when they are associated with other pelagic organisms. For instance, members of the *Roseobacteraceae* showed mutualistic interactions with some microalgae, providing

essential nutrient for their growth (Croft et al., 2005) while other are strict pathogens producing algicide compounds (Seyedsayamdost et al., 2011).

Despite their different role and composition, the separation of FL and PA bacteria is widely used in most of the marine bacterial communities ecological studies, that focus in their vast majority on the free-living fraction (Chafee et al., 2018; Fuhrman, 2009; Sunagawa et al., 2015; Teeling et al., 2016).

1.4.1.2. RARE BIOSPHERE OF SEA WATER

With NGS technologies, scientists observed that environmental communities present a few abundant species and a great number of rare OTUs (Sogin et al., 2006). At the beginning, these rare OTUs were associated with possible sequencing errors and technological biases (Reeder and Knight, 2009). But now, their existence is widely admitted in microbial ecology and are often identified as OTUs with a relative abundance below 0.01% in a sample (Galand et al., 2009). In addition, they possess a key role in bacterial community ecology and in the environment. Some studies showed that some members of the rare biosphere presented coherent ecological patterns and were actually active in marine environment (Alonso-Sáez et al., 2015; Campbell et al., 2011; Hugoni et al., 2013; Reveillaud et al., 2014) and could play significant role in global biogeochemical cycles (Hausmann et al., 2018; Musat et al., 2008). Moreover, these rare OTUs might be bacteria in a state of dormancy but could answer quickly as soon as the conditions become favorable in the ocean (Jones and Lennon, 2010; Pedrós-Alió, 2006). These characteristics make marine bacterial community composition highly dynamic in space and time (Fuhrman et al., 2015).

1.4.2. BACTERIAL COMMUNITIES DYNAMICS

1.4.2.1. TEMPORAL DYNAMIC

Time-series studies have been crucial features of oceanography to understand the ecological processes occurring in the oceans (Ducklow et al., 2009). Observations of the evolutions of microbial communities through time have been increasingly studied over the past decades in marine environment. Several

monitoring of bacterial communities have been implemented in the main ocean regions. The main ones are listed in Table 1.

Table 1 | Main time series of oceanic microbial (bacteria and archaea) communities. (Adapted from Bunse and Pinhassi, 2017)

Region	Location	References
Polar	Antarctic peninsula	Ghiglione et al., 2012
	Franklin Bay, Western Antarctic	Alonso-Sáez et al., 2008
Temperate	RADIALE time-series project (E2), east-Atlantic	Alonso-Sáez et al., 2015
	Western Channel Observatory (L4), English Channel	Gilbert et al., 2012
	Linnaeus Microbial Observatory (LMO), Baltic Sea	Lindh et al., 2015
	German Bight, North sea	Teeling et al., 2016
	Blanes Bay Microbial Observatory (BBMO), Mediterranean sea	Gasol et al., 2016
	SOLA time-series, Mediterranean sea	Galand et al., 2018
(Sub)- tropical	Hawaiian Ocean time-series (HOT), north Pacific	Bryant et al., 2016
	Bermuda Atlantic time-series study (BATS), west Atlantic	Vergin et al., 2013
	San Pedro Ocean time-series station (SPOTS), east Pacific	Needham et al., 2013

All these time surveys found that marine microbial communities are very dynamic in time, and at different scales, from days to years (Gilbert et al., 2012; Hatosy et al., 2013). At each scale, microbial community composition dynamic was linked to both physicochemical (*i.e.* temperature, salinity, inorganic nutrient concentrations) and biological (*i.e.* phytoplankton dynamic, predation) changes in the environment (Fuhrman et al., 2015). Looking at small scales, marine environment experiments small spatial and temporal patchiness in water masses (*e.g.* eddies) or meteorological events, as well as different dynamic of phytoplankton or viruses (Martin-Platero et al., 2018; Needham and Fuhrman, 2016). These small variations leads to the sharp set-up of ephemeral microbial communities that present a high turnover within few days (Lindh et al., 2015b; Martin-Platero et al., 2018).

At the scale of days to weeks, one of the major drivers of bacterioplankton dynamic is the phytoplanktonic blooms that recurrently occurs during spring and summer in temperate coastal waters (Taylor and Ferrari, 2011). Phytoplankton during its growth or senescence will release a great amount of a variety of DOM that represent an important source of nutrient for heterotrophic bacteria. Given the diversity of these organic molecules, their degradation is based on an high diversity of heterotrophic bacteria mainly affiliated, among others, to the classes *Flavobacteria*, *Gammaproteobacteria* and *Roseobacter clade* (Buchan et al., 2014; Bunse et al., 2016; Lucas et al., 2015, 2; Pinhassi et al., 2004).

Surprisingly, despite substantial variations across days to weeks, and the importance of fast environmental changes, marine bacterial communities present highly robust seasonal patterns that remain stable years after years over more than a decade (Fuhrman et al., 2015; Giovannoni and Vergin, 2012). This trend was also observed for taxa answering to the blooms that were found recurrent over different years in the North sea (Teeling et al., 2016) or even individual members that showed a strong rhythmicity over seven years (Lambert et al., 2018).

The underlying processes that could explain this regularity are still poorly understood. Global seasonal drivers such as temperature, circadian cycles, nutrient concentrations or seasonal primary production are most likely candidates. For instance, seasonal patterns in deep waters are weaker if not absent (Cram et al., 2015; Hatosy et al., 2013). But these seasonal patterns cannot explain alone such recurrence. An increasing number of studies suggested that the interactions (positive or negative) between pelagic organisms could represent a key parameter of the communities structures and dynamic (Lima-Mendez et al., 2015; Needham and Fuhrman, 2016; Worden et al., 2015). The Black Queen hypothesis formulated recently (Morris et al., 2012a) supports these observations and suggests that marine bacteria could form interdependent cooperatives interactions (Sachs and Hollowell, 2012).

1.4.2.2. SPATIAL DYNAMIC

Along with the set-up of time series, the study of global biogeographical patterns has also been increasingly investigated over the past decades. Thanks to the democratization of sequencing

technologies, numerous expedition sampled the ocean in vast geographical scales such as the Global Ocean Sampling in 2007 (Rusch et al., 2007), the Malaspina expedition in 2010 (Duarte, 2015), the Ocean Sampling Day since 2014 (Kopf et al., 2015) or Tara Ocean expeditions (**Figure 4**) from 2009 to 2013 (Sunagawa et al., 2015), giving great insights on microbial assemblages in the oceans, both in horizontal and vertical gradients.

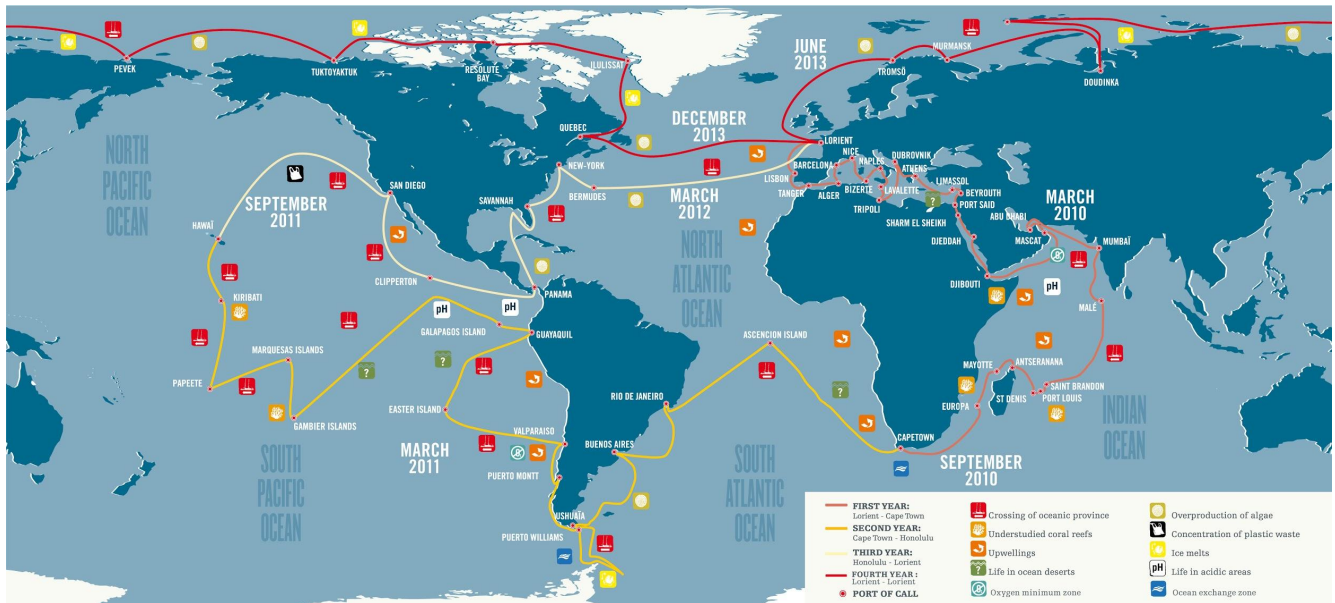


Figure 4 | Map indicating the different sampling stations of the expedition Tara ocean. (<https://oceans.taraexpeditions.org/>)

As for temporal surveys, marine microbial communities present different biogeographical patterns, at different scales (Baltar et al., 2015; Sunagawa et al., 2015). Within horizontal distances, important biogeographic patterns have been found at the basin scale (Sunagawa et al., 2015). Microbial communities would be different between polar, tropical and temperate waters (Swan et al., 2013). Some highly ubiquitous taxa such as *SAR11 Clade*, *SAR86 Clade* or *Cyanobacteria* are present at the ocean scale, but they will present different ecotypes adapted to different temperature and latitudes (Brown et al., 2012; Delmont et al., 2017). At smaller distances, studies observed no or moderate changes in bacterioplankton community composition (Li et al., 2018) leading to the idea that they are

homogeneous in distances under 50 km (Hewson et al., 2006). But in transition zone, such as coastal area, strong environmental gradients (e.g. salinity) can occur and that structure the bacterial communities (Cottrell and Kirchman, 2003; Herlemann et al., 2011; Wu et al., 2006; Zhang et al., 2006). It has even been suggested that along with marine and freshwater microbiomes exists a brackish microbiome that is composed of bacteria adapted to intermediate salinities (Hugerth et al., 2015). Moreover, sharp physicochemical changes linked to physical structures such as eddies or upwellings can lead to strong spatial heterogeneity of microbial communities (Allen et al., 2012; Baltar et al., 2010).

Important changes in microbial communities are found with the depth and are widely observed at large scale. Different communities are found across the global bathyal layers of the oceans : euphotic, mesopelagic and bathypelagic (Jing et al., 2013). This evolution in the water column is observed at the domain level, as Archaea are more abundant in mesopelagic and bathyal environments (Karner et al., 2001; Varela et al., 2008). At a smaller scale, within the euphotic layer, vertical thermal stratification determine most of the variance in community composition of bacterial communities (Cram et al., 2015; Sunagawa et al., 2015) that highlighted the importance of local factors (such as environmental variables and biology) on microbial communities.

1.4.2.3. EVERYTHING IS EVERYWHERE

The most widespread ecological theory in explaining marine microbial distribution in the ocean is resumed in the famous sentence of Baas-Becking : « everything is everywhere but the environment selects » (Baas-Becking, 1934). This concept posits that microorganism's assemblage results from a selection by local environmental conditions rather than dispersal limitation in the environment, intrinsically considered perfect in this assumption. This statement is strongly debated among scientists, particularly for the first part of the tenet « everything is everywhere ». Marine environments are thought to encounter no or few dispersal barriers compared to other environment (e.g. terrestrial) thanks to large currents occurring at a global scale (Broecker, 1991). Moreover particular physiological capacities of marine bacteria such as their ability to enter a stage of dormancy could allow them to travel through long distances and thus reduce dispersal limitations (De Meester, 2011; Lennon and Jones, 2011). But

studies that were interested in dispersal limitations of marine bacteria or archaea at basin scale found contradictory results (Sintes et al., 2015; Sul et al., 2013).

Nevertheless, environmental selection (biotic and abiotic) on bacterial communities, has been widely observed through spatial and temporal surveys as detailed above and very few studies observed the opposite (Kan et al., 2007). It can be the dominant process that structures microbial communities in some marine environment (Barberán and Casamayor, 2010). However, it has to be noted that understanding the importance of environmental selection rely on the sampling of a wide quantity of variables, which is complex and rarely done. Particularly for the biotic interactions such as competition, symbiosis or predation that represent key parameters on the structure of marine bacteria (McFall-Ngai et al., 2013; Needham et al., 2018; Thingstad, 2000; Worden et al., 2015).

1.4.3. MARINE BACTERIA TROPHIC STRATEGIES

Marine surface waters represent an extreme environment for marine heterotrophic bacteria as they often experiment strong nutrient limitations due to the sparsity of their resource (Stocker, 2012). Nutrients are often found in patchiness (e.g. phytoplankton, marine snow) in a bulk of low-nutrient interstitial water (Luo and Moran, 2014). Moreover, nutrients availability is highly dynamic in time and space linked to mesoscale heterogeneity (e.g. upwellings) or seasonal spring blooms (Lauro et al., 2009). Marine bacteria have evolved with a large array of strategies to maximize reproductive success under these heterogeneous conditions. These strategies have major consequences in the processing of organic matter and our comprehension of bacterial community's composition dynamic.

Marine heterotrophic bacteria usually fall into two general trophic strategies based on their capacity to grow at different substrate concentrations. Oligotrophs are adapted to thrive with low ambient concentrations of organic matter, while copiotrophs rely on higher levels (Giovannoni et al., 2014; Hunt et al., 2013; Koch, 2001). Typical and successful oligotrophs are members of *SAR11 Clade* that can represent up to 50% of the cells in the oligotrophic ocean (Morris et al., 2002). On the contrary, copiotrophs typically represented by *Alteromonadales* or some *Rhodobacteraceae* will be found in

abundance associated to particles and planktonic organisms or under highly productive conditions such as phytoplankton blooms in temperate coastal waters (Polz et al., 2006).

While this notion is a fundamental ecological trait of marine bacteria, it has no strict consensus definition (See table 1 in Schut et al., 1997) and terms oligotroph and copiotroph are sometime used interchangeably with r- and K-strategy despite their different meanings (Ho et al., 2017; Newton and Shade, 2016; Semenov, 1991). The former, r-strategists, are also called opportunists and they are characterized by a capacity of replicate intensively when the conditions are favorable. They will thus present highly variable growth rates, harboring a feast and famine lifestyle. K-strategist bacteria are on the contrary adapted to environments where the resource supply rates are lower, and won't be able to take advantage of increasing nutrients loadings. Consequently, r and K- strategists integrates key notions such as susceptibility to predation or growth rates capacities (Giovannoni et al., 2014).

Another way to differentiate bacteria is their capacity to utilize a variety of different organic compounds. Thought this notion is less investigated in marine environments (Kirchman, 2016). Some heterotrophic bacteria are generalist, being able to grow on a wide variety of different compounds (Mou et al., 2008). This is notably the case for copiotrophic *Rhodobacteraceae* that present large and versatile genomes and can inhabit diverse niches (Moran et al., 2004). Specialist bacteria will be highly adapted to a specific niche by targeting few compounds. Members of the *Bacteroidetes* are often considered as specialist in degrading complex polysaccharides compounds (Fernández-Gómez et al., 2013; Unfried et al., 2018). At low resource concentrations, substrate specialists have been shown to be the most active (Sarmiento et al., 2016). Sometimes generalist or specialist lifestyle are also described as synonyms to copiotrophs and oligotrophs respectively (Christie-Oleza et al., 2012). Again, there is a lack of consensus of these definitions, as the typical oligotrophic members of the *SAR11 Clade* « [...] cannot plausibly be characterized as specialists, being one of the most successful and widely distributed chemoheterotrophic groups in the ocean » as written by Grote et al., 2012.

Different studies have been trying to characterize marine bacteria trophic strategies, employing a wide variety of technics. For instance, the importance of nutrient concentrations and their specific incorporation by some heterotrophic taxa have been investigated using stable isotopes probing on different substrates (Alonso and Pernthaler, 2006; Bryson et al., 2017; Mayali et al., 2014). They revealed the complexity of these different strategies that fall on a continuum between copiotrophy and oligotrophy and high and low activity (Mayali et al., 2014). For instance, Mayali et al., 2012, uncovered at least 7 trophic strategies for marine bacteria regarding their capacity of exploiting substrates at increasing nutrient concentrations. However, these complex trophic lifestyles seems to be conserved within different heterotrophic taxa. Members of *SAR11 Clade* typically grow slowly and are outcompeted at high nutrient concentrations. While members of *Alteromonadales* grow fast at increasing concentrations of nutrients (Alonso and Pernthaler, 2006; Bryson et al., 2017).

The signature of oligotrophy and copiotrophy was investigated in the genomes (Lauro et al., 2009). Based on two organisms typically defined as oligotroph or copiotroph, they found genomic characteristics or functions (represented by cluster of orthologous genes) specific of each trophic strategy. Copiotrophs organisms present a higher genetic potential to respond faster to changing environmental conditions such as nutrients input or depletion and are more adapted to a particle-attached lifestyle. Oligotrophs, on the contrary present a smaller genome and are adapted to a free-living lifestyle and use a broader range of transported substances. Following the same logic, Song et al., 2017, investigated the metagenomic response of bacteria from the soil to copiotrophic or oligotrophic conditions. They found an enhancement of particular cluster of genes that would match with the genes defined by Lauro et al., 2009. This confirms that different trophic conditions are important and select adapted organisms with particular genomic characteristics.

2. THE IROISE SEA AND THE BAY OF BREST

2.1. COASTAL OCEANS: A CRUCIAL BUT FRAGILE ECOSYSTEM

Within the oceans, coastal areas are of utmost interest and have recognized ecological values (Costanza et al., 1997). Thanks to terrestrial input, they are among the most fertilized systems in the world, promoting high primary and secondary productivity (Barnes and Hughes, 1999; Nixon and Buckley, 2002). Thus, while coastal areas represent 8% of the global ocean surface, they account for 10 to 25% of the global marine primary production (Cloern et al., 2014). In temperate coastal waters, the inflow of inorganic nutrients leads to recurrent phytoplankton blooms in spring that are able to sustain a complex trophic chain. This diversity and productivity are influenced by a strong hydrodynamism, with currents, wind, freshwater inflow that are important controls for the onset of spring blooms (Ragueneau et al., 1996; Sverdrup, 1953). This hydrodynamic also leads to the formation of mesoscale physical features (*i.e.* upwellings, eddies or frontal zones) that strongly shape the biology of the ecosystem, starting with the microbial compartment (Casotti et al., 2000). This large hydrodynamism is supposed to be important actors of the diversity and the richness of the ecosystem (Legendre and Demers, 1984). Such dynamic features make temperate coastal waters challenging to study and require multiple approaches as well as long term observations and model simulations (Cadier et al., 2017; Weisberg et al., 2015).

In addition, a monitoring of this environment is important since it represents a sensible area. We have evidence that temperate coastal zones are highly degraded (Waycott et al., 2009). Mainly linked to anthropic pressures as around 70% of the world megacities are in the coastal zone (Ramesh et al., 2015). Eutrophication is one of the major anthropic impacts on coastal systems (Rabalais et al., 2002; Smith et al., 1999). It causes a disequilibrium in the major biogeochemical cycles (*i.e.* Carbon, Nitrogen and Phosphorus). Increase in nitrogen and phosphorus are due to human or animal waste and the use of land fertilizers that at the end, arrive into coastal waters. Excess nitrogen can lead to increasing toxic phytoplanktonic blooms that may have a critical impact on the global ecosystem, with significant

mortality events for a large number of species (Rosenberg and Loo, 1988). Other anthropic pressures lead to an increasing pollution of the coastal environment, with oil spills, chemical pollution (e.g. pharmaceutical products) or plastic release (Dussud et al., 2018; Kennish, 1991).

The influence of such contaminants on natural microbial communities have been observed, but their consequences are still poorly understood. On the one hand, microbial communities could participate to the degradation of pollutant molecules (Barra Caracciolo et al., 2015) and lead to their removal from the environment. On the other hand, they could be directly impacted in their richness or composition (Sauret et al., 2016).

In France, the Iroise sea and Brest Bay represent particular coastal environments with respects to their rich biodiversity and different societal challenges. These areas are intensively studied through the presence of the « Zone Atelier Brest Iroise (Zabri) » which aim at characterize the interactions between anthropic pressures and ecosystem functioning in this pilot site. This objective is possible through the presence of numerous major laboratories (e.g. UMR, Ifremer, CNRS, UBO) that carry different scientific surveys such as microbial survey of shellfish production area (REMI), phytoplankton and phycotoxin survey (REPHY), the survey of chemical contaminations (ROOCH), the Service d'Observation du Milieu Littoral (SOMLIT).

2.2. THE IROISE SEA AND THE USHANT FRONT

The Iroise sea is located in the north-western part of France (Figure 5). It is part of Celtic seas, close to the English channel in the North-East Atlantic ocean. It is separated in the north by the Ushant Island (48°5N) and in the south by the Sein Island (47°4N). The Iroise Sea at the west is under the Atlantic Ocean circulation regime while at the east it is influenced by land and freshwater inputs from different rivers, the main ones being the Aulne and the Elorn.

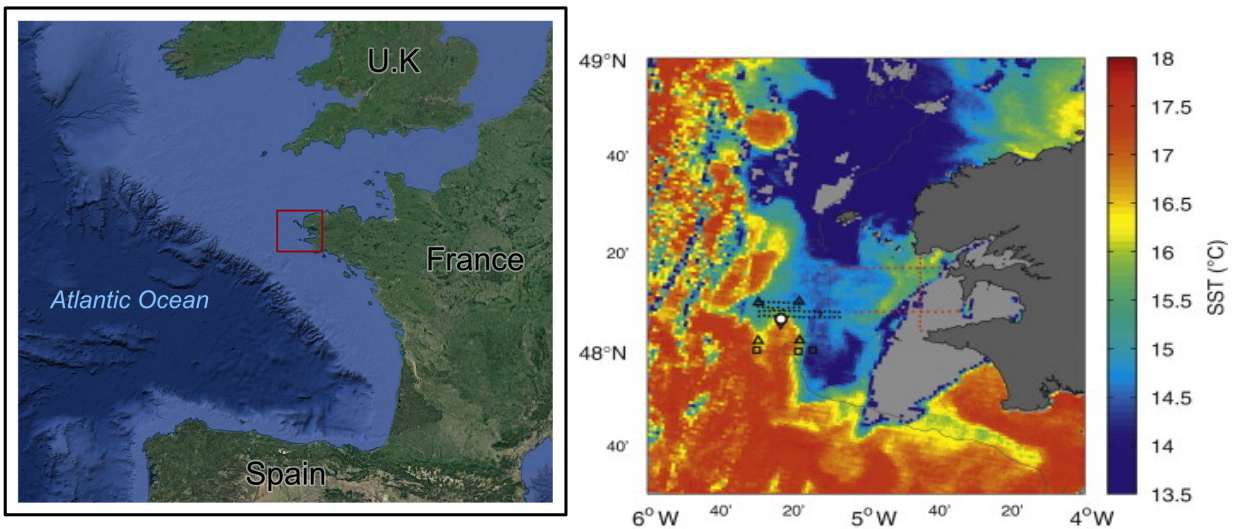


Figure 5 | Localization of the Iroise sea and Satellite sea surface temperature highlighting the presence of the Ushant Front.

This region is characterized by high macrotidal regimes (Muller et al., 2007). Along with wind and atmospheric fluxes, tide represent the main physical forces that governs the physical dynamic of this sea. All these physical forcing are responsible of the formation of a remarkable front, that represent the predominant feature of the Iroise sea between May to October (Le Boyer et al., 2009). In summer, the wind are less important and solar radiations heat the sea surface leading to a stratification of the water column. This stratification is broken down due to strong tidal currents that can reach $4\text{m}\cdot\text{s}^{-1}$ east of the Ushant Island, leading to a complete mixing of the water column. The Ushant front represent the area between stratified and mixed waters and it is characterized by a sharp horizontal gradient with changes in sea surface temperature of 4°C in less than a mile (Figure 5).

Frontal area are of high interest as they represent interface with enhanced planktonic biomass and productivity (Sournia et al., 1987). The Ushant front has been recognized as an exemplar of frontal area in a coastal system (Le Fèvre, 1987). It was discovered quite early in the 1950s and extensively studied in the following years, with numerous notable cruises such as SATIR-DYNATLANT in the 80s and FROMVAR in 2009 that characterized its physicochemical and biological characteristics (GREPMA, 1988; Le Boyer et al., 2009). In the Ushant front, the western water masses are typical of stratified

waters with a marked thermocline and phytoplanktonic growth limited by nutrient concentrations at the surface. On the contrary western mixed waters are not limited in nutrients but the homogenization limits the access to light for phytoplanktonic cells. At the frontal area, input of nutrients from the cold deep waters or from the mixed area are possible, stimulating phytoplanktonic growth and accumulation of zooplankton, leading to a remarkable productivity. This particular concentration of biomass at the frontal area likely attract numerous birds, fish larvae and species of particular interest for the conservation (Schultes et al., 2013).

Overall, the presence of a remarkable biodiversity in the Iroise Sea with numerous species of algae, birds and marine mammals as well as its rich cultural heritage lead to the creation of the first marine natural park in France (Parc Marin Naturel de l'Iroise, **Erreur ! Source du renvoi introuvable.**).

2.3. THE BAY OF BREST AND THE SOMLIT STATION

The Bay of Brest represent a macrotidal semi enclosed ecosystem located in the North-Western coast of France. In the one hand, this bay is influenced by terrestrial and freshwater inflow through the presence of 6 rivers, where the two dominants are the *Aulne* and the *Elorn* (Figure 6). The river discharge is maximal on November and February linked to the winter storms, and minimal in summer (Ragueneau et al., 1996).

On the other hand, the Bay of Brest is influenced by oceanic waters from the Iroise sea through the tide. These two areas communicates through a narrow channel called *le Goulet* (1.8 km width, 50 m depth, 6 km long). The renewal of marine water through tide mixing is so important – about a third of its entire volume – that this semi-enclosed bay is considered as a typical marine environment (Delmas and Treguer, 1983).

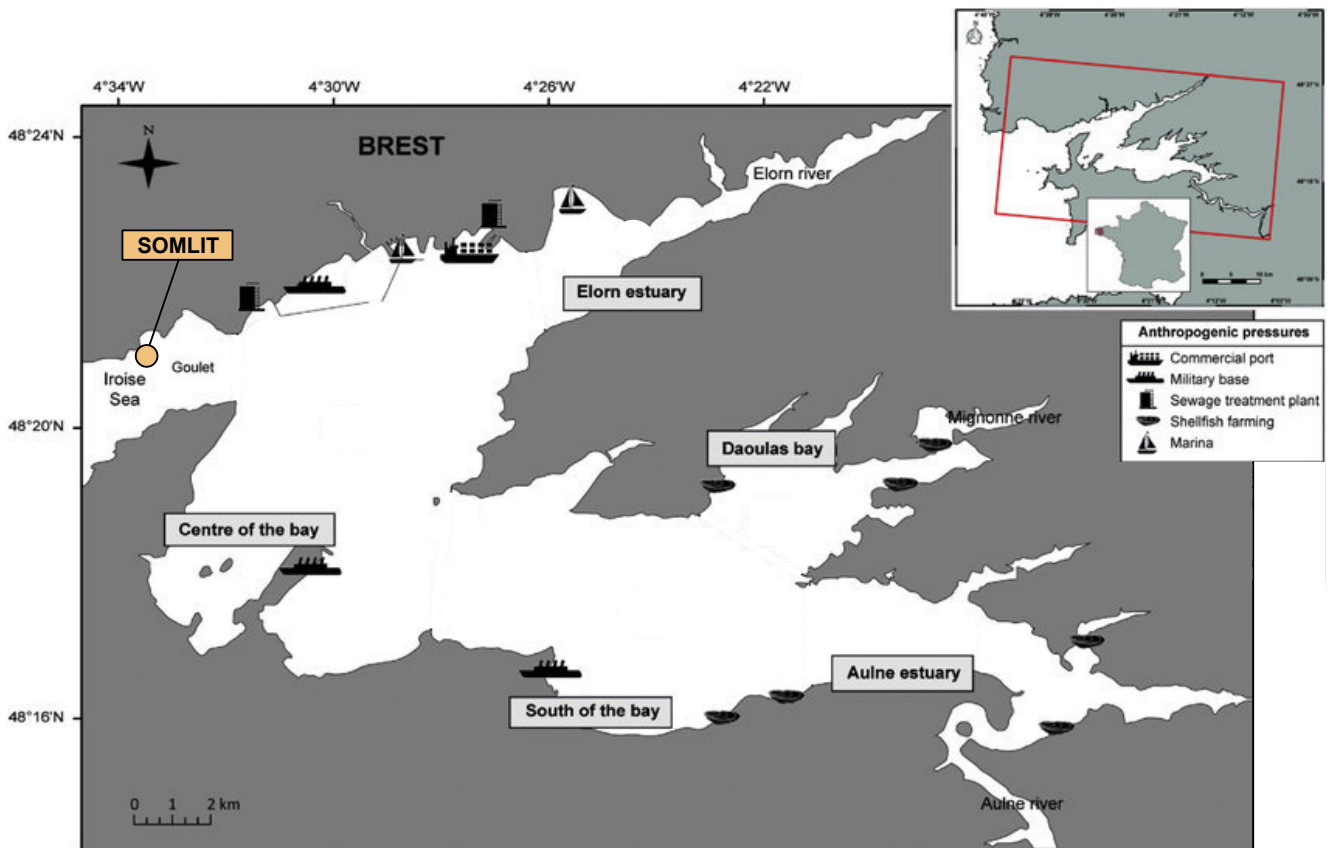


Figure 6 | Map of the Bay of Brest. Adapted from Frère et al., 2017.

The Bay of Brest represents a rich area, typical of temperate zones of the Northern Hemisphere, with an annual pelagic primary production estimated to reach $280 \text{ g C}\cdot\text{m}^{-2}\cdot\text{a}^{-1}$ (Quéguiner and Tréguer, 1984). Its productive period is situated from late March to early October with spring blooms occurring in April. The phytoplankton communities are dominated by diatoms and its development is strongly influenced by tide mixing that will change nutrient and light availability (Figure 7) (Ragueneau et al., 1996).

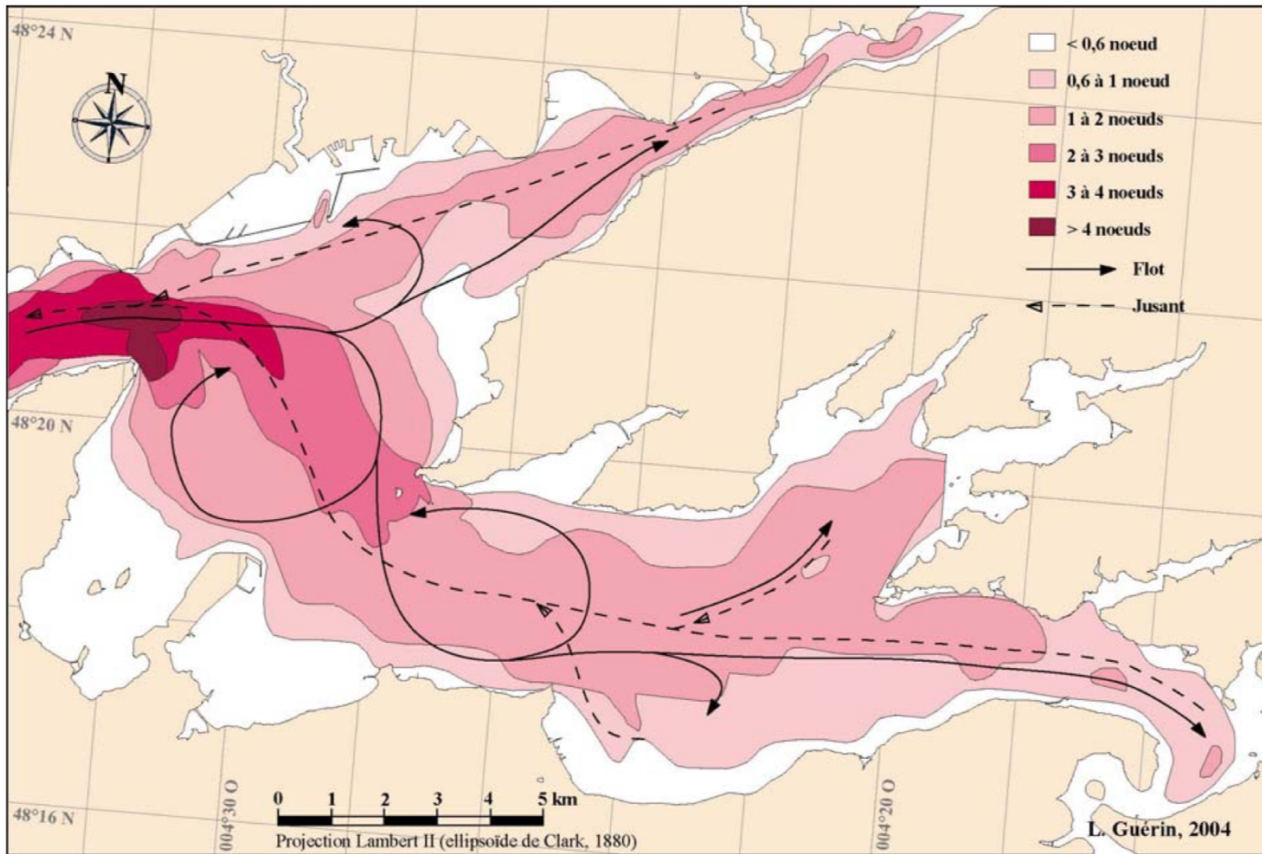


Figure 7 | Tide currents in the Bay of Brest.

This area is impacted by different anthropogenic pressure with numerous commercial, fishing and recreational activities. A total of 27 cities are present in the coastline of the Bay of Brest, the most important one being Brest with 139 163 inhabitants (INSEE data, 2015). This city possess two treatment plants in the commercial harbour and near the Goulet (Maison Blanche) that release treated waters in the Bay of Brest. Brest possess the largest harbour of Brittany with an active commercial activity that cumulated a total tonnage of 2 800 000 in 2018. The main two basins, represented by the Aulne and the Elorn harbor different agricultural activities and of hog production (Troadec and Le Goff, 1997). Since the beginning of the century, the anthropogenic phosphorus and nitrogen carried by the rivers presented a drastic increase of 10 fold (Le Pape et al., 1996; Treguer and Queguiner, 1989). This increase of nitrogen presents an influence on Diatoms populations that tend to diminish towards more Dinoflagellates and possibly harmful phytoplankton (Beucher et al., 2004).

A recent study of microplastic pollution in the Bay of Brest showed a contamination of around 0.24 ± 0.35 microplastics/m³ which is very low compare to other highly contaminated estuaries as in China (Frère et al., 2017). However, these microplastics seems to enrich a particular diversity of bacteria such as members of the genus *Vibrio* compare to the surrounding water (Frère et al., 2017).

A station of the national Service d'Observation Marine du Littoral (SOMLIT) survey is also present in the Bay of Brest since 1998 (**Figure 6**). Every week more than 15 physicochemical (e.g. temperature, salinity, oxygen, nitrates, phosphates) and biological parameters (e.g. chlorophyll a, cytometric cells counts) are assessed at this station. This survey is completed by the presence of an instrumented buoy, MAREL-Iroise, from the marine observatory of the IUEM, that record the main physicochemical parameters (*i.e.* temperature, conductivity, dissolved oxygen, fluorescence and turbidity) every 20 minutes since 2000.

Overview and Objectives

Despite the central role of microorganisms in marine ecosystems and the presence of a dense observation network in the Bay of Brest and the Iroise Sea, pelagic bacterial community's ecology has not been studied in these two environments for more than 30 years with the work of Evelyne Jacq in the 1980s. At that time, microbial ecology mostly relied on microscopic cell counts and cultures, providing a relatively low resolution to study bacterial dynamics. This thesis has for objectives the set-up of a microbial observatory in the Bay of Brest and the Iroise Sea, based on technological advances in metabarcoding and metagenomics to better study these complex microbial communities. This new subject in this area was made possible thanks to the collaboration between two laboratories, LM2E and LEMAR, combining expertise in microbial ecology and knowledge of the coastal marine environment.

A large part of this thesis, anticipated from a Master II internship, consisted in acquiring the samples, during 4 three-day campaigns in the Iroise Sea (2 of which I participated in) and the regular, weekly and bi-monthly samples, at the SOMLIT station, as well as setting up the molecular biology protocols for DNA extraction, libraries preparation for metabarcoding and metagenomics. For all these steps I was helped by Morgan Perennou, engineer at LEMAR. Bioinformatics and statistical pipelines for the analysis of metabarcoding data were also produced during this thesis. The network analyses were carried out on the advice of Damien Eveillard (LS2N, Nantes) and Antonio Fernandez-Guerra (Max Planck Institute, Bremen). Finally, Johanne Aube, engineer at LM2E, carried out the metagenomic analysis presented in the last chapter.

The first objective of this work, was to investigate the dynamics of bacterial communities in these two environments through an amplicon sequencing approach. In Chapter 1, the dynamics of bacterial communities associated with the presence of the Ushant tidal front in the Iroise Sea were studied. During 4 campaigns, the vertical and horizontal structure of the water bodies associated with the front was characterized, representing an exceptional and original setting to understand bacterial communities dynamics. In a second chapter of the thesis, I analyzed the first 3 years of the microbial temporal survey at the SOMLIT station in the Bay of Brest. This survey offers a high temporal resolution, making possible to observe the seasonal patterns of bacteria in relation to the variations in the marine environment characterized by the SOMLIT survey.

The second objective of this work was to investigate the functional content of these communities. Using a metagenomic approach we were able to reconstruct the genomes of microorganisms present in the Bay of Brest and the Iroise Sea, that help to understand the physiology of uncultivated members of marine bacteria. The genome of the most abundant bacterium of these two environments was assembled and is presented in a third chapter.

Chapter I

Spatio-temporal dynamic of bacterial communities in the Iroise Sea

1. RÉSUMÉ EN FRANÇAIS

Comprendre la dynamique spatiale et temporelle des communautés de microorganismes marins est une étape clé pour mieux appréhender leur rôle dans les écosystèmes. Les structures méso-échelles représentent des zones clés pour étudier ces dynamiques car elles influent directement sur la distribution du plancton et leur activité. Dans ce chapitre, nous avons étudié la dynamique des communautés de bactéries libres dans l'eau, associées au front de Ouessant en mer d'Iroise. Cette structure frontale est caractérisée par des eaux stratifiées typiques au large de Ouessant, avec en surface des eaux chaudes et épuisées en nutriments. Cette stratification n'est pas permise dans les eaux plus côtières, notamment de par l'action des courants de marées, qui peuvent homogénéiser la colonne d'eau sur toute sa hauteur. Dans ces masses d'eau instables, le développement du phytoplancton va être limité par ce mélange constant. A l'interface de ces masses d'eau, le front permet le maintien d'une forte production, par un apport local d'eau enrichie en nutriments. Afin d'étudier comment la présence de ce front pouvait structurer les communautés bactériennes, cinq stations réparties autour du front ont été échantillonnées en Septembre 2014, Mars, Juillet et Septembre 2015, et les communautés ont été caractérisées par le séquençage de leur ADNr 16S. La composition des communautés microbiennes va changer en fonction des différentes masses d'eau présentes, mais pas au niveau de la zone frontale. Par une approche en réseau, nous avons ensuite regroupé les OTUs qui présentaient des dynamiques similaires au niveau de ces masses d'eau en posant l'hypothèse qu'ils partageaient des niches écologiques similaires. Cette approche a mis en évidence l'importance des producteurs primaires qui, de par leur répartition hétérogène autour du front et dans le temps, vont définir des environnements plus ou moins riches en matière organique, et ainsi sélectionner des bactéries hétérotrophes avec des stratégies trophiques différentes. Des bactéries oligotrophes vont être présentes toute l'année et dominer en hivers et dans les eaux profondes au large, lorsqu'il y a peu de développement phytoplanctonique. A l'inverse, différentes populations copiotrophes vont apparaître plus abondantes en été et dans les eaux plus côtières et juste après un bloom intense en Juillet, certaines répondant par des variations en abondance fortes. Ces résultats laissent supposer l'importance de ces communautés bactériennes dans la reminéralisation de la matière organique de la boucle microbienne en mer d'Iroise.

Linking Spatial and temporal dynamic of bacterioplankton communities with ecological strategies across a coastal frontal area

C. Lemonnier^{1,2}, M. Perennou², D. Eveillard³, A. Fernandez-Guerra⁴, A. Leynaert², L. Marié⁵, H. Morrison⁶, L. Memery², C. Paillard² and L. Maignien^{1,6}

¹ Univ Brest, CNRS, IFREMER, Laboratoire de Microbiologie des Environnements Extrêmes, F-29280 Plouzané, France

² Univ Brest, CNRS, IRD, IFREMER, Laboratoire des Sciences de l'Environnement Marin, F-29280 Plouzané, France

³ Univ Nantes, CNRS, Centrale Nantes, IMTA, Laboratoire des Sciences Numériques de Nantes, F-44322 Nantes, France

⁴ MPI for Marine Microbiology, Microbial Genomics and Bioinformatics Research Group, D-28359 Bremen, Germany

⁵ Univ Brest, CNRS, IRD, IFREMER, Laboratoire d'Océanographie Physique et Spatiale, F-29280 Plouzané, France

⁶ Josephine Bay Paul Center for Comparative Molecular Biology and Evolution, Marine Biological Laboratory, Woods Hole, USA

Paper prepared for submission in ISME Journal

2. ABSTRACT

Ocean Frontal systems are widespread hydrological features defining the transition zone between distinct water masses. They are generally of high biological importance as they are often associated with locally enhanced primary production by phytoplankton. However, the composition of bacterial communities in the frontal zone remains poorly understood. In this study, we investigate how a coastal tidal front in Brittany (France) structures the free-living bacterioplankton communities in a spatio-temporal survey across four cruises, five stations and three depths. We used 16S rRNA surveys to compare bacterial community structures across 134 seawater samples and defined groups of co-varying taxa (modules) exhibiting coherent ecological patterns across space and time. Seasonal variations in primary producers, and distribution in the water column appeared as the most salient parameters controlling heterotrophic bacteria. Different dynamics of modules observed in this environment were strongly consistent with a partitioning of heterotrophic bacterioplankton in oligotroph and copiotroph ecological strategies. Overall, this study shows a strong coupling between bacterioplankton communities dynamic, trophic strategies, and seasonal cycles in a complex coastal environment.

3. INTRODUCTION

Bacteria dominate marine environment in abundance, diversity and activity where they support critical roles in the functioning of marine ecosystems and oceanic biogeochemical cycles (Cotner and Biddanda, 2002; Falkowski et al., 2008; Madsen, 2011). In the coastal environment, they are closely linked to other planktonic organisms (e.g. viruses, phytoplankton and zooplankton) during the recycling of organic matter and inorganic nutrients through the so-called microbial loop (Azam and Malfatti, 2007; Pomeroy et al., 2007). They form complex and highly dynamic assemblage (Giovannoni and Vergin, 2012), with bacterioplankton diversity variations in space and time linked to changes in functional diversity (Galand et al., 2018). Therefore, understanding how the bacterioplankton composition varies in the environment remains one of the central question to understand coastal ecosystem functioning better (Fuhrman et al., 2015).

Organic matter processing implies diverse heterotrophic bacterioplankton among which one could pinpoint members of *Bacterioidetes*, *Roseobacter group* or *Gammaproteobacteria* (9). These taxa contribute to the complexity of the marine ecosystem via different adaptive strategies, owing to the unequal access to their respective resource (10) For instance, heterotrophs are generally distinguished between oligotrophs and copiotrophs that compete at low and high nutrient concentrations respectively (Giovannoni et al., 2014; Koch, 2001). They also present different degrees of ecological specialization, with generalist bacteria able to assimilate a broad variety of substrates, while specialists will compete for a narrow range of nutrients (Mou et al., 2008). Analysis of these ecological traits offer a simplified view of complex microbial communities and has gained interest to understand the dynamic of natural microbial communities better and get an insight of their role in the ecosystem (Haggerty and Dinsdale, 2017; Krause et al., 2014; Raes et al., 2011).

Marine fronts are very common mesoscale features in the ocean and lie at the transition between water masses of different physicochemical characteristics that actively shape the distribution of microbial organisms (phytoplankton, zooplankton and bacteria). Driven by currents and mixing, local nutrients input in the vicinity of the front generally enhance primary and secondary production making frontal

zone area of high biological importance (Olson and Backus, 1985) and microbial processing of organic matter (Baltar et al., 2015; Heinänen et al., 1995). However, the bacterial communities composition involved in such dynamic systems remains to investigate (Baltar et al., 2016).

The Ushant Front in the Iroise Sea (Brittany, France) is considered as a model of a coastal tidal front (Le Fèvre, 1986). Its position and characteristics are highly dynamic and influenced by atmospheric forcing and tidal coefficient (Le et al., 2009). It occurs from May to October and leads to contrasted physicochemical environments with higher biomass at the frontal area (Le Fèvre et al., 1983). West of the front, stratification results in warmer oligotrophic surface waters and colder nutrient-rich deeper waters separated by a marked thermocline. East of the front, associated with highly variable conditions, permanently mixed coastal waters are characterized by an unlimited quantity of inorganic nutrients but with highly fluctuating conditions. These contrasted water bodies structure primary producers distribution with the dominance of small phytoplankton and dinoflagellates in surface stratified waters and diatoms in mixed waters (Birrien et al., 1991; GREPMA, 1988; Videau, 1987).

In this study, we tested the hypothesis that such contrasted water mass in physicochemical and biological parameters will strongly drive bacterioplankton communities. Using a network analysis, we defined groups of co-varying bacterial OTUs that present the same dynamic across the samples, suggesting that they possibly share the same ecological niches.

4. RESULTS

4.1.1. ENVIRONMENTAL SETTINGS

This study was carried out between September 2014 and September 2015 in the Iroise Sea, off Brittany (N.-E. Atlantic), in the vicinity of the Ushant island. In general, Ushant front position and characteristics were estimated using Satellite Surface Temperature (SST) maps (Figure 8) and CTD data collected at the dates of sampling (Figure S 1). The March 2014 sampling took place before the onset of the Ushant front and presented a homogeneous temperature around 10°C across all the stations. High nutrient concentrations at all stations (Si(OH)_4 : 1.82 to 4.38 μM , Nitrates: 5.88 to 12.13 μM ,

Table S 1), low surface chlorophyll a ($< 1 \mu\text{g.L}^{-1}$ except for Station 1, Figure S 2) and overall low phytoplankton cells counts observed during this cruise (

Table S 1) indicated that sampling occurred prior to the development of the phytoplankton spring blooms. In summer (September 2014, July 2015 and September 2015), SST maps showed a sharp transition between coastal and offshore temperatures confirming the presence of a frontal area. The different observed stratification regimes (Figure S 1) coincided with distinct physicochemical patterns and phytoplanktonic patterns across seasons (Figure 8.B,

Table S 1). In late summer (September samples), offshore deep waters shared close characteristics with winter waters (March sample, Temperature: 11.7 to 12.3°C), with few phytoplankton cells, similar concentration of Si(OH)_4 (1.61 to 3 μM) and Nitrates (3.99 to 6.41 μM). Conversely, surface waters presented high temperatures (14.7 to 18.2°C), a nutrient depletion (Si(OH)_4 : 0.02 to 1.65 μM , Nitrates: 0.00 to 0.09 μM) and high phytoplankton cells counts. In early summer (July), the nutrients

concentrations were overall lower than in September, probably consumed by the spring bloom coinciding with the onset of the front around May-June. A significant bloom occurred at stations 2 and 3, on the 27th of June, five days before the sampling period as seen in the satellite surface Chl_a observations (Figure S 2.B).

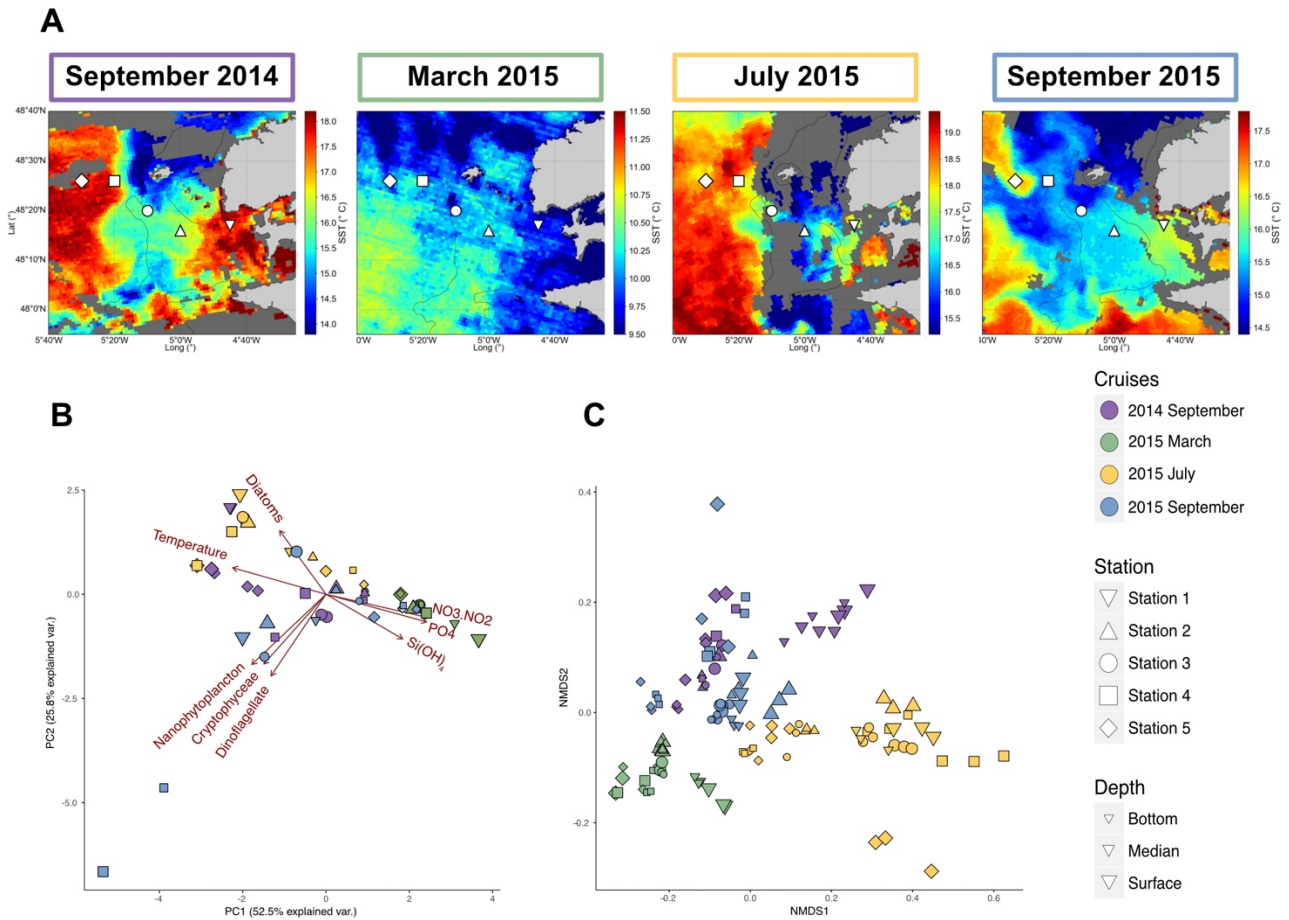


Figure 8 | Environmental and bacterial communities characteristics of the Iroise Sea for the different cruises. **A.** Map of the Satellite Sea Surface Temperatures (SST) in the Iroise Sea for each cruise (September 2014 the 8th, March 2015 the 6th, July 2015 the 2nd and September 2015 the 8th), highlighting in the summer cruises the presence of the Ushant front that separates offshore warm stratified waters and coastal cooler mixed waters. The shape correspond to the different stations sampled, with at each time 2 or 3 depths. Surface waters are highly dynamic and some eddies are conspicuous during September 2015, bringing cold water in surface stratified waters. Thus, SST map are not enough to define the frontal area and vertical profiles are needed to characterize the different water masses (Figure S 1). **B.** Principal Component Analysis of the environmental characteristics for the different samples based on temperature, inorganic nutrients (Si(OH)₄, NO₃+NO₂, PO₄) concentrations and microscopic count of the different phytoplanktonic groups (*Diatom*, *Nanophytoplankton*, *Cryptophyceae* and *Dinoflagellate*) **C.** NMDS of the bacterial diversity based on Bray-Curtis dissimilarities among the different cruises, stations and depth (stress = 0.091).

4.1.2. BACTERIAL COMMUNITY DYNAMICS IN THE IROISE SEA

Bacterial diversity structure presented unusual seasonal and spatial patterns, mirroring environmental physicochemical variations represented in the PCA plot of environmental variables (Figure 8.B), as shown in the NMDS ordination plot (Figure 8.C, for each cruise separately see Figure S 3).

In winter, communities were highly similar throughout the water column (Permanova test on Bray-Curtis dissimilarities ~ Depth was non-significant, $Pr(>f) = 0.225$), but presented a coastal to offshore gradient (Permanova test on Bray-Curtis dissimilarities ~ Station was significant, $Pr(>f) = 0.001$). Conversely, with the onset of stratification in summer, communities were much more heterogeneous and their structure followed the different water masses: deep water communities remained similar to those present in winter (Figure 8.B). In contrast, surface bacterial communities in stratified regimes mostly departed from this typical winter structure, with depth becoming a significant structuring driver (Permanova test Bray-Curtis dissimilarities ~ Depth $Pr(>f) = 0.004$). This trend was not significant in mixed coastal samples (Permanova test Bray-Curtis dissimilarities ~ Station was not significant, $Pr(>f) = 0.222$ and 0.434). Besides, similar communities were found in stratified waters in September 2014 and 2015 samples (Figure 8.C) indicating a clear recurring seasonal pattern.

Using LEfSe algorithm (Segata et al., 2011), we found that some stations favored specific biomarker OTUs within each sampling time (Figure S 4). Biomarkers such as OTU1 (*Amylibacter*) and OTU4 (*Planktomarina*) were found at the most near-shore station (Station 1) in September 2014 and March 2015 and other biomarkers were found in surface stratified waters in early summer (OTU29 *NS4 marine group*) and late summer (OTU10 *Synechococcus*) cruises, but no biomarker was identified for the mixed and frontal stations except in July 2015 at Station 2 with OTU35 (*Aliivibrio*) and OTU74 (*Pseudoalteromonas*).

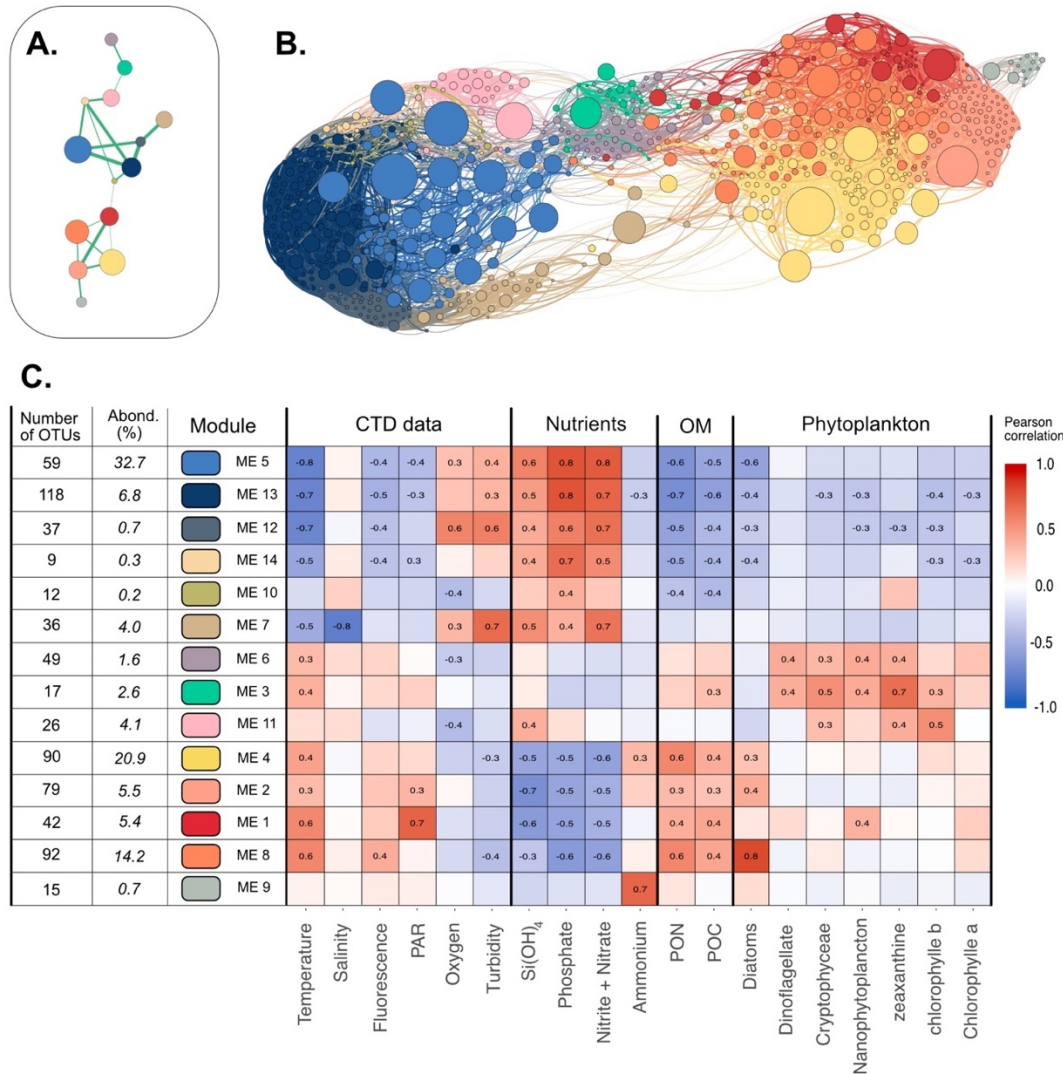


Figure 9 | Visualization of the different modules and their main characteristics. A. Representation of the main connectivity between the different modules detected with the Louvain algorithm. Each node represent a module, each edge represent the median of the positive correlations between two modules. Only the strongest connection are shown. B. Network visualization of the dataset. Each node represents an OTU, while each edge represent a positive correlation (>0.3) obtained from the covariance matrix calculated with SPIEC-EASI. The node size depends on the abundance of one OTU (in number of sequences) in the entire dataset. The edge size depends on the value of the correlation. Nodes colors represent the OTU affiliation to the different Louvain communities. Network was represented using Gephi and Force Atlas layout algorithm. C. Presentation of the 14 modules detected. Pearson correlations between each module eigengene and the different environmental data are presented in the heatmap. OM: Organic Matter, PAR: Photosynthetically Available Radiation, PON: Particulate Organic Nitrogen, POC: Particulate Organic Carbon. Only the significative correlations (p value > 10^{-4}) are shown. The number of OTUs in each module and the relative abundance of each module (in % of all the reads) are detailed.

4.1.3. DESCRIPTION OF MODULES WITH A SPECIFIC DYNAMIC, DIVERSITY AND CORRELATION WITH ENVIRONMENTAL PARAMETERS.

We further investigated groups of co-varying OTUs that could potentially share the same ecological niches in this contrasted marine environment via a co-occurrence network analysis. In the inferred network, we were able to identify 14 sub-networks (Figure 9) defining groups of OTUs (named modules hereafter) with similar distribution patterns across the entire study. Two modules (5 and 4) were dominant in our dataset and respectively accounted for 32.7% and 20.9% of all sequences. Eight modules represented between 1.6% and 14.2%, and four were rarer with less than 1% in abundance. The OTU taxonomy in each module is summarized in Figure S 5 and Figure 11 presents the distribution of the dominant families among each module.

Since the eigengene can characterize each module, we investigated to which extent module distribution could be correlated (*i.e.* Pearson correlation) with environmental parameters (Figure 9.C). Correlation patterns partitioned modules into three significant subnetworks. The first one mostly comprises modules 5 and 13 representing 32.7% and 6.8% of the dataset respectively, that correlated positively to inorganic nutrients and negatively to temperature and Particulate Organic Matter (POC, PON) values. Their relative abundance in the different samples showed that they were dominant in oligotrophic waters: together they ranged from 39% at Station 1 to 72% at Station 5 of the late winter communities, and 54% to 59% of the deep stratified waters in September (Figure 10). *SAR11 Surface 1* (46.7% of the module sequences), *ZD0405 marine group* (12%) and *SAR86 Clade* (6.4%) dominated the main module (*i.e.* module 5). The module 13 showed a different diversity, dominated by members of *Marinimicrobia* (15.8%), *SAR11 Deep 1* (12%), *SAR11 Surface 1* (8.9%) and *Salinisphaeraceae* (7.9%).

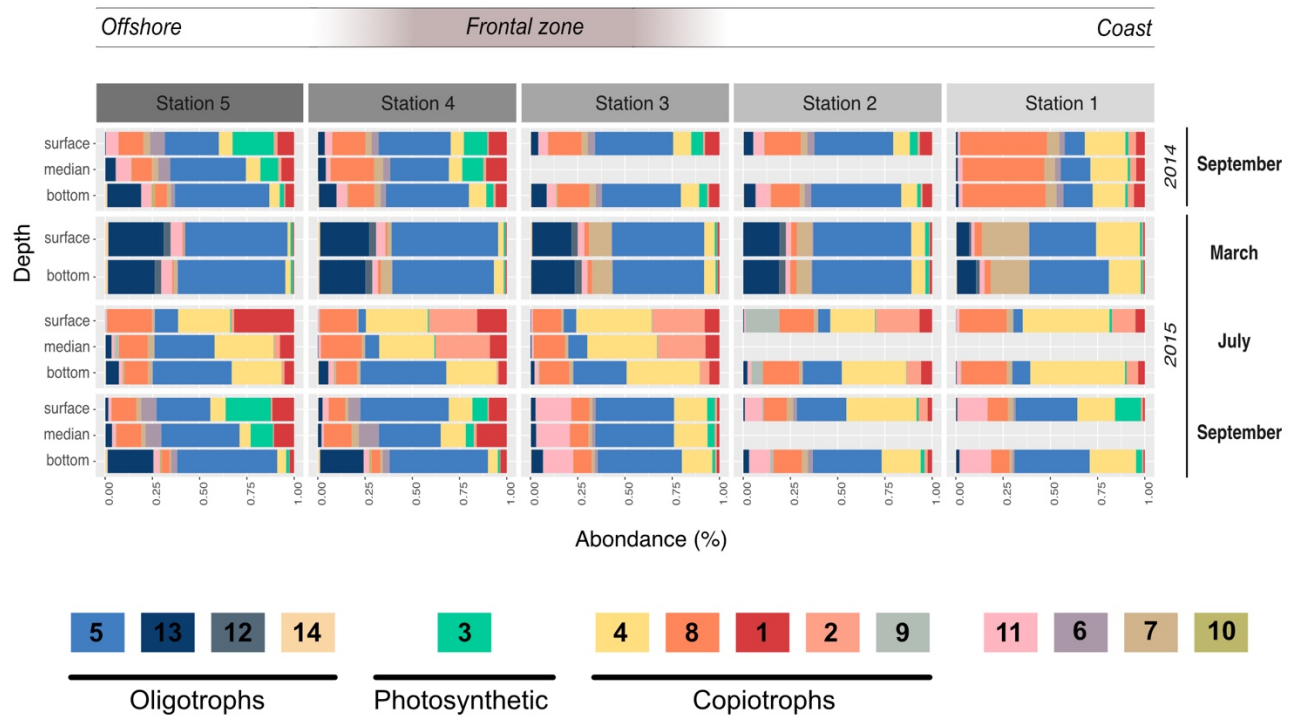


Figure 10 | Modules dynamic across the different campaigns, stations and depth. Relative abundance of the different modules in each campaign, station and depth. The value shown resulted of the mean of the triplicates for each sampling. The legend includes the different hypothesis of the modules trophic strategies based on their correlations to the environmental parameters, their taxonomy and their dynamic. Oligotrophic modules are likely to present heterotrophic bacteria adapted to low organic matter levels, while copiotrophic bacteria are likely more competitive under higher organic matter levels. Finally, the photosynthetic module is almost only composed of Cyanobacteria.

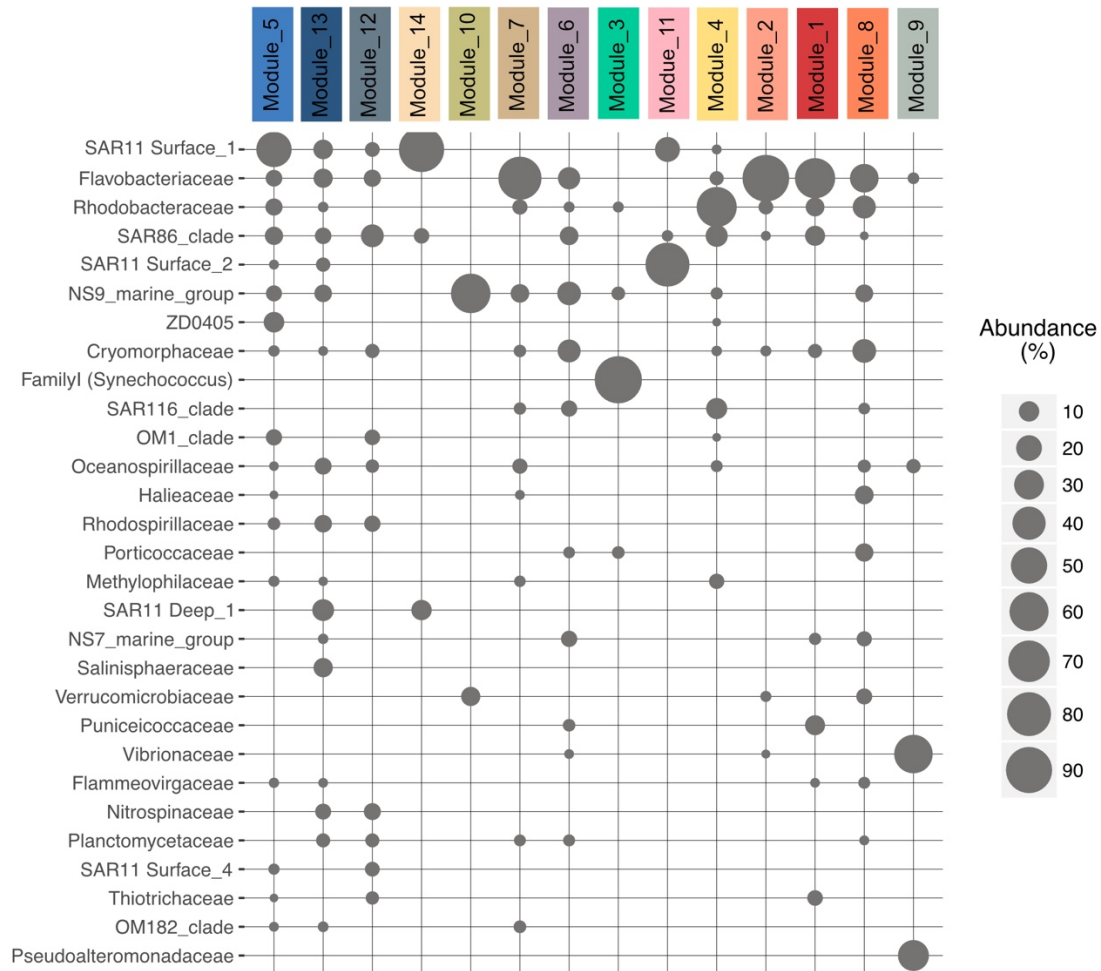
Modules 4, 8, 2 and 1 contributed to the second group, representing 20.9%, 14.2%, 5.5% and 5.4% of the dataset respectively. They presented clear inverse correlations compared with modules of the first group, *i.e.* they correlated positively to temperature, POM and Diatoms but negatively to inorganic nutrient concentrations. In relative abundance they dominated the samples of July and in the surface and coastal stations. Their taxonomy was distinct from the first group, *i.e.* together they gathered the majority of *Flavobacteriaceae* (70%) and *Rhodobacteraceae* (88%). Also, module 9 was enhanced in July 2015 but was almost exclusively present in Station 2 and strongly correlated with Ammonium (0.72, p

< 0001). Its taxonomy was also particular as dominated by *Vibrionaceae* (58%) and *Pseudoalteromonadaceae* (30%).

The third group (modules 6 and 3) was less abundant and presented different correlations with environmental parameters. They respectively account for 1.6 and 2.6% of the dataset, were correlated with specific phytoplankton groups (Dinoflagellates, Nanophytoplankton) and were dominant in the surface and DCM of stratified waters in September where they made up to 34% of the community. Among them, module 3 was constituted almost of Cyanobacteria (94%) affiliated to *Synechococcus* genus only.

To examine whether these modules exhibited different ecological behavior (*i.e.* feast and famine or rather slow steady growth), we computed the coefficient of variation (CV) for all OTU used in the network analysis and plotted this metric against OTU occurrence as proposed in Newton and Shade, 2016, with some modifications in the interpretation (see material and methods). We then examined their module membership (Figure S 6 and Figure 11.B): modules showed a clear gradient ranging from relatively low CV (LCV) OTUs (modules 4, 11, 14, 5 and 13) to high CV (HCV) OTUs (modules 9, 10, 7 and 2).

A.



B.

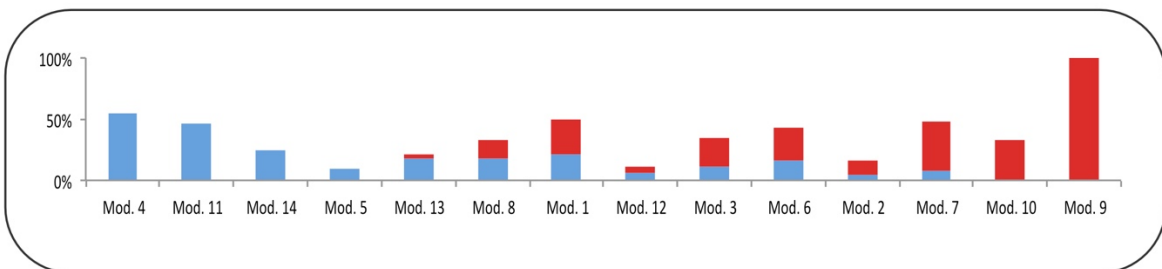


Figure 11 | Modules taxonomies and Coefficient of Variation of their OTUs **A.** Relative abundance of the 30 main families in each module. **B.** Proportion of OTUs (in % of all the OTUs in a module) that present a High Coefficient of Variation (HCV) in red, or a Low Coefficient of Variation (LCV) in blue defined based on Newton method (Newton and Shade, 2016).

5. DISCUSSION

The spatial and temporal dynamics of bacterial communities associated with a coastal tidal front

In this study, we investigated the spatiotemporal variations of free-living bacterial community composition over one year across contrasted coastal water masses characterized by a seasonal tidal front structure. For each cruise (season) we observed a coastal to open ocean gradient in the bacterioplankton community structure (Figure S 3). However, winter communities are comparatively much more homogeneous than summer ones, highlighting the importance of the onset of a tidal front and open ocean stratification for bacterioplankton community structure. This result is mainly due to the sharp divergence of summer surface samples, as deep summer samples are firmly related to winter ones (Figure 8.C). The existence of two contrasting types of bacterioplankton dynamics (winter and deep vs. summer surfaces) reflects similar patterns in sample ordination based on biogeochemical parameters (Figure 8.B). These observations of a frontal zone as a sharp ecological transition for bacterioplankton among different water masses are coherent with reports from another frontal area (Baltar et al., 2016) and in stratified waters (Cram et al., 2015; Ghiglione et al., 2008).

Interestingly, the frontal area itself around station 3-4, did not exhibit a specific bacterial composition (Figure 8.C and Figure 10) but rather seems to result from mixing between adjacent communities, as already suggested (Sournia, 1994). We indeed could not identify any biomarker OTUs associated with frontal stations. Hence, bacterioplankton communities associated with enhanced productivity at the Ushant front (Videau, 1987) is resulting from an increase in resident plankton density (Franks, 1992) rather than the emergence of distinct communities.

Besides, we observed a remarkable seasonal recurrence between the two September cruises, within each station and depth. This pattern was comparatively more accentuated toward the open ocean (station 4 and 5) than near shore (stations 1 to 3). This temporal dynamic is coherent with previous reports of seasonal patterns in single stations as shown in the North Sea (Chafee et al., 2018) or the English Channel (Gilbert et al., 2012) time series. Our results suggest that despite a substantial

heterogeneity in our system, such seasonal recurrence patterns also extend to stations along geographical and bathyal gradients.

Distinct bacterioplankton trophic strategies associated with tidal front temporal and spatial dynamics

Using network analysis on samples exhibiting strong temporal and spatial variations, we could define modules of ecologically coherent OTUs. As suggested elsewhere, such modules can be considered as OTUs sharing the same environmental niches with coherent ecological strategies (Eiler et al., 2012). In association with previously characterized dominant taxa in each module, we aim at identifying higher order bacterioplankton communities organization and revealing ecological drivers of community dynamics in the frontal zone.

Our study shows an evident module partitioning into two dominant subnetworks dominated by heterotrophs, exhibiting distinct inverse covariance patterns (Figure 9.A.B). Their correlations with environmental variables also clearly separate conditions between water masses with active primary producers (low nutrients and high POM values) that were present in summer surface samples, and conditions with overall low development of phytoplankton typical of winter and deep samples. Development and decay of phytoplankton lead to the release of dissolved organic matter, which is almost only available for heterotrophic bacteria (Azam and Malfatti, 2007). Thus, these two different conditions likely select for copiotrophic or oligotrophic heterotrophic bacteria adapted to different concentrations of organic matter in the environment (Giovannoni et al., 2014). Taxonomy affiliation was highly coherent with these observations. Modules 5 and 13 are dominated by *SAR11 clades*, which are typically thriving in basal concentrations of organic matter (Morris et al., 2012b) or members of taxa typical of the sub-euphotic zones such as *Marinimicrobia*, or *SAR11 Deep 1* (Agogu  et al., 2011). Conversely, modules 4, 8, 2, 1 and 9 were dominated by members of the *Rhodobacteraceae* and *Flavobacteriaceae*, typically associated with phytoplankton-derived organic matter (Buchan et al., 2014; Teeling et al., 2016). Thus, availability of phytoplankton-derived organic matter will drive heterotrophs dynamic.

The last subnetwork with module 3, dominated by *Synechococcus*, highlights the contribution of phototrophic bacteria. These cyanobacteria became strongly dominant in the surface of well stratified-waters in late summer, which is already known to favor these small phytoplanktonic cells (Lomas and Lipschultz, 2006; Smith et al., 2014).

Interestingly, chemoautotrophs involved in the nitrification such as Nitrospinaea did not form a separate module and were included in the oligotrophic module 13. Nitrification in the water column is typically found in the sub-euphotic zone as nitrite-oxidizing bacteria could be light-sensitive (Lomas and Lipschultz, 2006) and can be outcompeted by phytoplankton for the uptake of ammonium (Smith et al., 2014; Wan et al., 2018). This fact could explain why they emphasize a dynamic similar to oligotrophic bacteria.

Differential bacterioplankton responses to organic matter availability

Using OTUs coefficient of variations, we examined whether heterotrophic bacteria had different responses to organic matter availability in summer samples. Their dynamic and composition was highly coherent with a partitioning of the microbial loop between different taxa (Bryson et al., 2017). For instance, several modules presented a majority of HCV OTUs (Figure 11.B): *Tenacibaculum* sp. (a Flavobacteriaceae dominating Module 2) was highly abundant reaching up to 20.7% of the sequences in the 2015 July surface waters shortly after significant phytoplankton bloom. One single *Tenacibaculum* OTU dominated the surface of 4 stations over a 40 km distance just a few days after a phytoplankton bloom. This dynamic of *Tenacibaculum* genus is coherent with previous observations of a recurrent population increase after seasonal phytoplankton blooms in coastal water (Teeling et al., 2016) while being part of the rare biosphere rest of the time (Alonso-Sáez et al., 2015). At the same period, module 9 dominated by Vibrionales (*Aliivibrio* sp) and Alteromonadales (*Pseudoalteromonas* sp.) class was especially abundant at Station 2, reaching up to 27% of the sequences (Figure 10). Those taxa are known copiotroph bacteria that can sharply increase in abundance in response to high substrates loadings (Tada et al., 2012). These results strongly support an opportunist lifestyle of bacteria in

modules 2 and 9. When they are dominant, bacteria with such an opportunist strategy have a critical role as they can contribute to a substantial fraction of organic matter recycling (Pedler et al., 2014).

In contrast, module 4 mostly comprised LCV OTUs with a cosmopolitan distribution across time, depth and geography. The dominant family in module 4, the *Rhodobacteraceae*, are known as generalist bacteria with a versatile substrate utilization (Moran et al., 2004; Newton et al., 2010). Some of the free-living *Rhodobacteraceae* grow on low molecular weight organic compounds and numerous studies pointed out that they compete for the same substrate as *SAR11 Clade*, but are more competitive at higher organic matter concentrations. They can be very successful in coastal waters, such as a representative of the genus *Planktomarina* in the North Sea (Voget et al., 2015). Other dominant taxa of this module are widespread lineage of marine environments that also target small organic matter molecules such as *SAR86 Clade* (Dupont et al., 2012), *Methylophilaceae* (Huggett et al., 2012) or members of *SAR116 Clade* (Oh et al., 2010)

Module 8 presented mixed CV value among its OTUs. Many members of this module can use complex algal-derived substrates as *Flavobacteriaceae* (Buchan et al., 2014), *Verrucomicrobiacea* (Martinez-Garcia et al., 2012) or *Cryomorphaceae* (Bowman, 2014), and present a typical increase in abundance after phytoplankton blooms (Teeling et al., 2016). We interpret these observations by the relatively low sampling time resolution that probably fail to resolve rapid OTU variations of other opportunistic bacteria.

Overall, this study permitted to delineate several groups of bacterioplankton trophic strategies and their variations in response to biogeochemical cycles in a highly dynamic coastal environment. Oligotrophs, typically represented by *SAR11* OTUs, were present in all water masses. They outcompete other heterotrophic organisms in shallow organic matter environments such as winter and deep offshore water masses, while copiotrophs communities develop during the high productivity period. Within copiotrophs, several taxa show patterns of adaptation to rapidly changing conditions in their substrate availability: opportunist taxa can punctually become highly dominant (represented by *Tenacibaculum* and *Alteromonadales*) after local events such as phytoplankton bloom, while more

generalist taxa such as *Rhodobacteraceae* exhibit a ubiquitous distribution. This complex heterotrophic community dynamics highlight the central role of free-living bacteria in organic matter cycling. Indeed, previous studies of the Iroise Sea nutrient cycling demonstrated that after the initial depletion of winter nutrients pool, phytoplankton growth strongly rely on recycled nutrients through the microbial loop (LHelguen et al., 2005), through which they are remineralized several times in a seasonal cycle (Birrien et al., 1991). In this system, bacteria could be directly responsible for up to 25% of urea with the remaining originating from ciliates mainly, which also mostly feed on bacteria (LHelguen et al., 2005).

6. CONCLUSION

Here, using a complex 3D (vertical, horizontal, and temporal) survey and network analysis, we were able to get insight on ecologically coherent modules of free-living bacteria in correlation with the seasonal dynamics of a coastal tidal front. This study suggests that heterotrophs constitute the main modules and abstract the main complexity of the marine ecosystem as they are driven by the availability of organic matter produced by the phytoplankton. This hypothesis should be confirmed with measurements and characterization of dissolved organic matter quantity and quality in the water column. Moreover, to better understand these systems, future studies should include the characterization of other key members of the planktonic communities such as viruses, ammonium oxidizing archaea, eukaryotic phytoplankton and grazers, as well as particle-attached bacteria that are central for organic matter and biogeochemical cycles.

7. MATERIAL AND METHODS

7.1. STUDY SITE AND SAMPLING DESIGN

For this study, we realized four East-west transects of about 60 km across the Iroise sea in September 2014 (from the 9th to the 11th), March 2015 (from the 10th to the 12th), July 2015 (from the 1st to the 3rd) and September 2015 (from the 8th to the 10th) aboard the R/V Albert Lucas. Station positions remained identical across campaigns and were meant to cross the front (Figure 8.A). Station 5 (48°25 N, 5°30 W) is the most offshore station and is characterized by a strong stratification in summer. Station 4 (48°25 N, 5°20 W) was closer the front. Station 3 (48°20 N, 5°10 W) and 2 (48°15 N, 5°00 W) are present in the mixed area and Station 1 (48°16 N, 4°45 W) is the most coastal station, near the outflow of the Bay of Brest. At each station, we obtained CTD profiles to assess the physical characteristics of the water column and set the depth of the different biological and chemical sampling: surface, near the bottom and at the Deep Chlorophyll Maximum (DCM).

7.2. NUTRIENTS, PHYTOPLANKTON COUNTS AND PIGMENTS ANALYSIS

Seawater was sampled at each depth for nutrients, biogenic silica (BSi), chlorophyll a, particulate organic carbon and nitrogen (POC/PON) concentrations and microscopic phytoplankton cell counts and identification. Dissolved inorganic phosphate (DIP) and silicate (DSi) were determined from filtered seawater on Nuclepore membrane filters (47 mm, and DIN on Whatman GF/F filters (25 mm). Samples for dissolved inorganic nitrogen (DIN) and DIP were then frozen, whereas samples for DSi were kept at 4°C in the dark. Concentrations were later determined in the lab by colorimetric methods using segmented flow analysis (Auto-analyzer AA3HR Seal-Analytical) (Aminot and K erouel, 2007), BSi was determined from particulate matter collected by filtration of 1 L of seawater through 0.6 µm polycarbonate membrane filter. The analysis was performed using the alkaline digestion method (Ragueneau and Tr guer, 1994). Total Chl a and phaeopigments (Pheo) were determined in particulate organic matter collected on 25 mm Whatman GF/F filters. The filters were frozen (-20°C) and analyzed

later by fluorometric acidification procedure in 90% acetone extracts (Holm-Hansen et al., 1965). Particulate material for POC/PON measurements was recovered on pre-combusted (450 °C, 4 h) Whatman GF/F filters. Samples were then analyzed by combustion method (Strickland and Parsons, 1972), using a CHN elemental analyzer (Thermo Fischer Flash EA 1112). Phytoplankton samples were fixed with Lugol's solution and cell counts were carried out using the Lund et al., 1958, method.

7.3. BACTERIOPLANKTON COMMUNITIES SAMPLING

Seawater for bacterial diversity analysis was collected at different depth using Niskin Bottle and directly poured in 5L sterile carboys previously rinsed three times with the sample. For each depth, three biological replicates sampling were done, consisting of three different Niskin deployments, except Station 4 in September 2014 Filtrations started in the on-shore laboratory 3 to 4 hours after sampling. Water samples were size-fractionated using three in-line filters of different porosity: 10 µm, 3 µm (PC membrane filters, Millipore®) and 0.22 µm (Sterivex® filters). In this study, we focused only on the 0.22 - 3 µm free-living fraction. All filters were frozen in liquid nitrogen before storage at -80°C. In total, 2 to 5 liters of seawater were filtered each time, depending on filters saturation.

7.4. DNA EXTRACTION AND SEQUENCING

Half-filters were directly placed in a Tube matrix B® (MP Biomedicals) with lysis buffer (Tris, HCl, EDTA), SDS 10% and Sarkosyl 10% and subjected to physical lysis for 5 minutes on a vortex plate. The liquid fraction was then collected in order to perform a phenol-chloroform extraction using PCI (25:24:1) and a precipitation step with Isopropanol and 5 M sodium acetate. DNA was resuspended in 100µL of sterile water. Bacterial diversity was assessed targeting the v4-v5 hypervariable regions of 16S rDNA with the primers 518F (CCAGCAGCYGCGGTAAN) / 926R (CCGTCAATTCNTTTRAGT-CCGTCAATTTCTTTGAGT - CCGTCTATTCCTTTGANT) (Nelson et al., 2014). PCR products were purified using AMPure XP® kit and DNA quantity was measured using PicoGreen® staining and a plate fluorescence reader. Each sample was diluted to the same concentration and pooled before sequencing in an Illumina MiSeq sequencer at the Marine Biological Laboratory (Woods Hole, USA).

7.5. BIOINFORMATICS ANALYSIS

We obtained 22 523 398 raw reads, with a range from 31 583 to 2 001 687 reads per sample. Reads were merged and quality-filtered according to recommendations in Minoche et al., 2011, using the Illumina-Utils scripts (Eren et al., 2013b). Those steps removed 17 % of all the sequences. We used the Swarm algorithm (Mahé et al., 2014) to cluster the 18 876 655 remaining sequences into 1 611 447 operational taxonomic unit OTUs. Chimera detection was done using Vsearch *de novo* (Rognes et al., 2016) resulting in 281 825 filtered OTUs. Taxonomic annotation for each OTU was done with the SILVA database v123 (Quast et al., 2012) using Mothur classify.seqs command (Schloss et al., 2009). 19 516 OTUs (1 247 339 sequences) were affiliated to unwanted taxa (e.g. Archaea, Chloroplasts, Mitochondria, Eukaryota) and removed from the dataset. As a result, we obtained 262 308 OTUs for 134 samples, with a large number of singletons (205 292 OTUs).

7.6. STATISTICAL ANALYSIS

Libraries were normalized for read number using DESeq package in R (Anders and Huber, 2010). A visualization of the bacterial community structure similarity was assessed with an NMDS based on Bray-Curtis dissimilarities using vegan R package. Significant influence of the depth, the station and the sampling time on bacterial communities was investigated using a permutational multivariate analysis of variance (PERMANOVA) based on Bray-Curtis dissimilarities and 999 permutations. We used principal coordinate analysis (PCA) ordination with the R vegan package to characterize water samples based environmental parameters. Finally, we examined the presence of biomarker OTUs of the different stations within each cruise using the LEfSe software (Segata et al., 2011).

7.7. NETWORK ANALYSIS

Network analysis was done based on a truncated matrix only containing the most abundant OTUs with more than 50 sequences in at least three samples. The filtrated matrix contained 681 OTUs that accounted for 79% of all the sequences. This filtration step was necessary to avoid false positive

correlations in the network analysis. Our module detection analysis followed part of the pipeline assembled by Fernandez-Guerra and available online (<https://github.com/genomewalker>). First, the application of SPIEC-EASI computes pairwise co-variance based correlation between OTUs in order to address the sparsity and compositionality issues inherent of microbial abundance data (Kurtz et al., 2015). The possible interaction was then inferred using the glasso probabilistic inference model with a lambda.min.ratio of 0.01. Based on this model, only the significant covariance values were extracted and transformed into a correlation matrix using the `cov2cor` function.

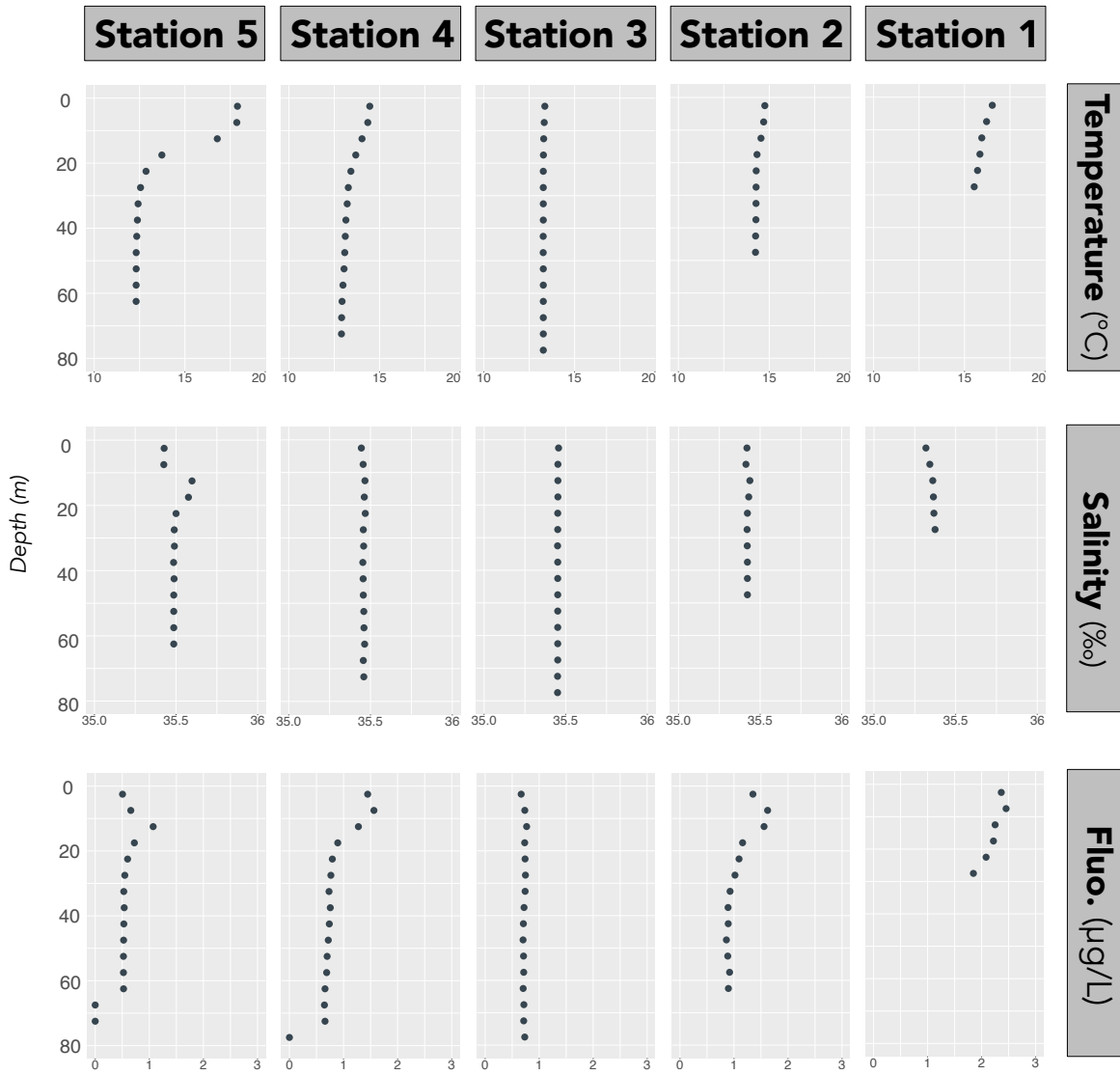
We then delineated network modules as groups of highly interconnected OTUs that presented very closely related dynamic among the studied samples. Those modules were defined using the Louvain algorithm (Blondel et al., 2008). We used Gephi (Bastian et al., 2009) and Force Atlas 2 layout algorithm to generate a visualization of main correlations (>0.3). Module eigengenes (ME, a Principal Component Analysis of the abundance of the different OTUs of a module) was calculated based on the relative abundance matrix using WGCNA function `moduleEigengenes` (Langfelder and Horvath, 2008). Those ME were used to calculate Pearson correlations between the different modules and the environmental variables, using the WGCNA commands `moduleTraitCor`, `moduleTraitPvalue`.

We used the methods proposed by Newton and Shade, 2016, to infer the trophic strategy of OTUs within each module. In this study, they differentiated opportunist taxa with high abundance variability in space and time, and marathoners with low abundance variability. For measuring this they used the Coefficient of Variation of OTU dynamic combined with their abundance and prevalence in the samples (Newton and Shade, 2016).

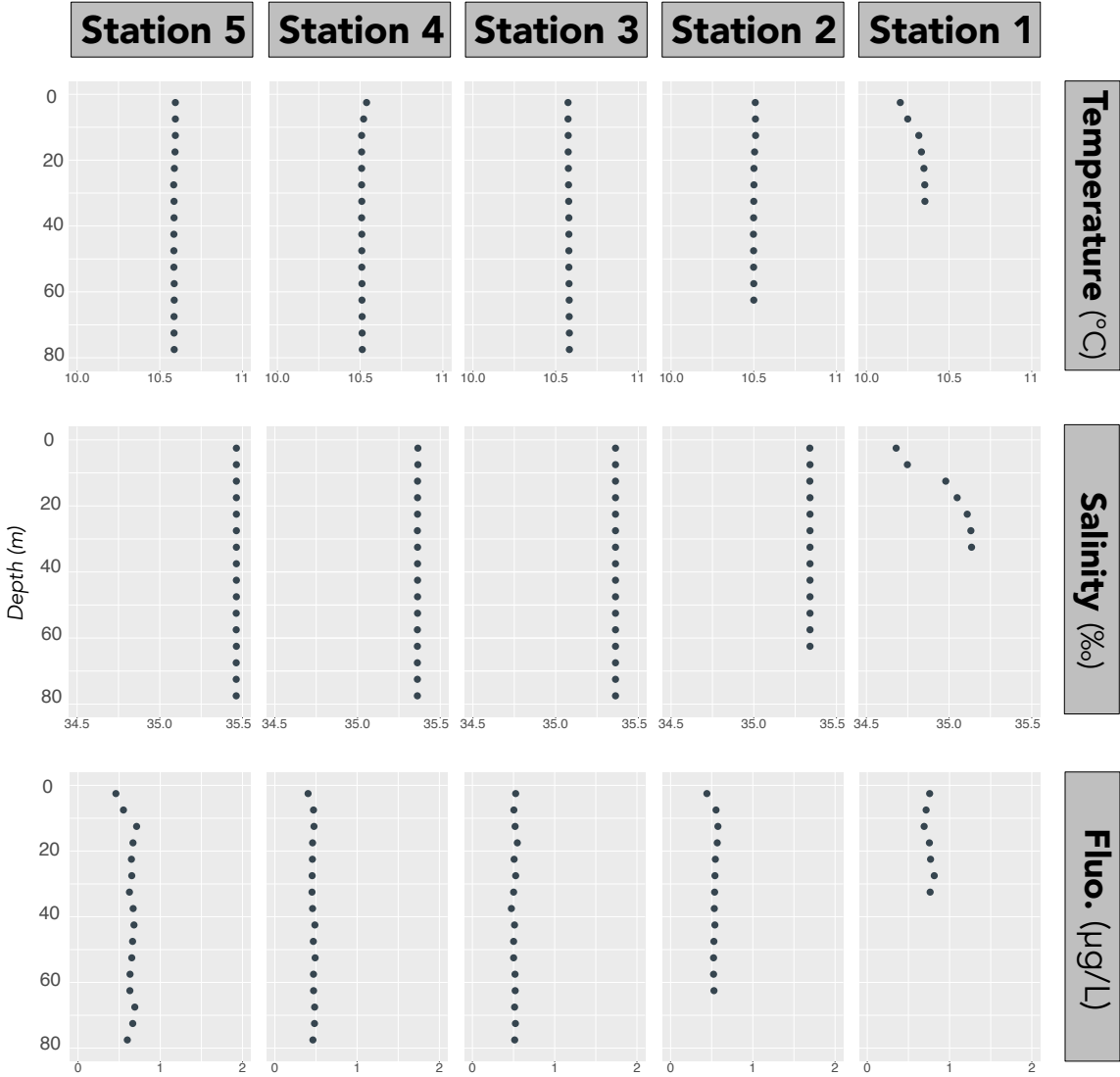
Based on their method, we defined opportunists OTUs or HCV (resp. marathoners, or LCV) as OTU's above (resp. below) the 5% upper (resp. lower) boundary of the linear modeling 95% confidence interval in a CV plot against OTU occurrence. In our analysis, OTU occurrence was defined by a relative abundance $> 0.05\%$.

8. SUPPLEMENTARY DATA

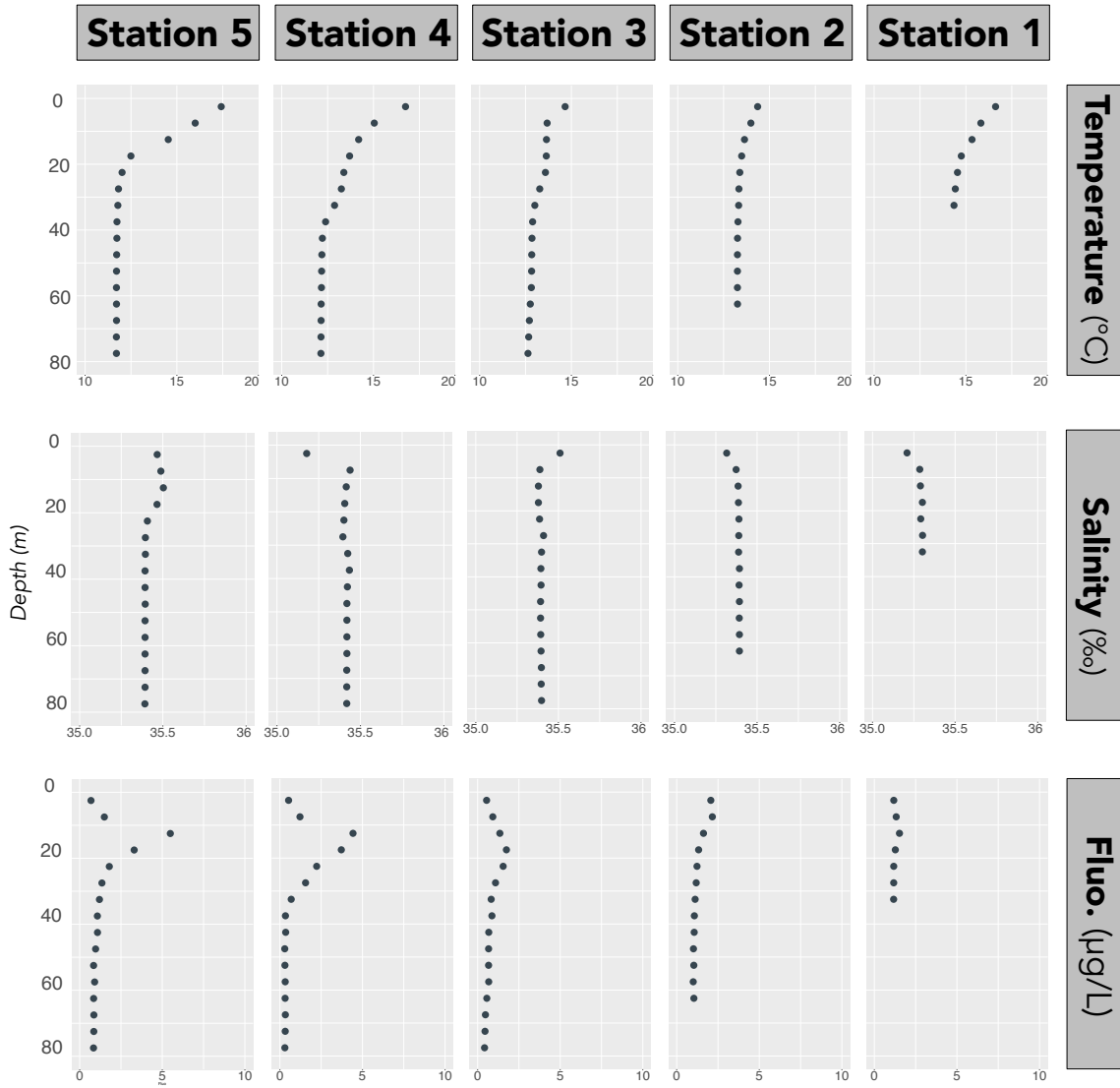
A. September 2014



B. March 2015



C. July 2015



D. September 2015

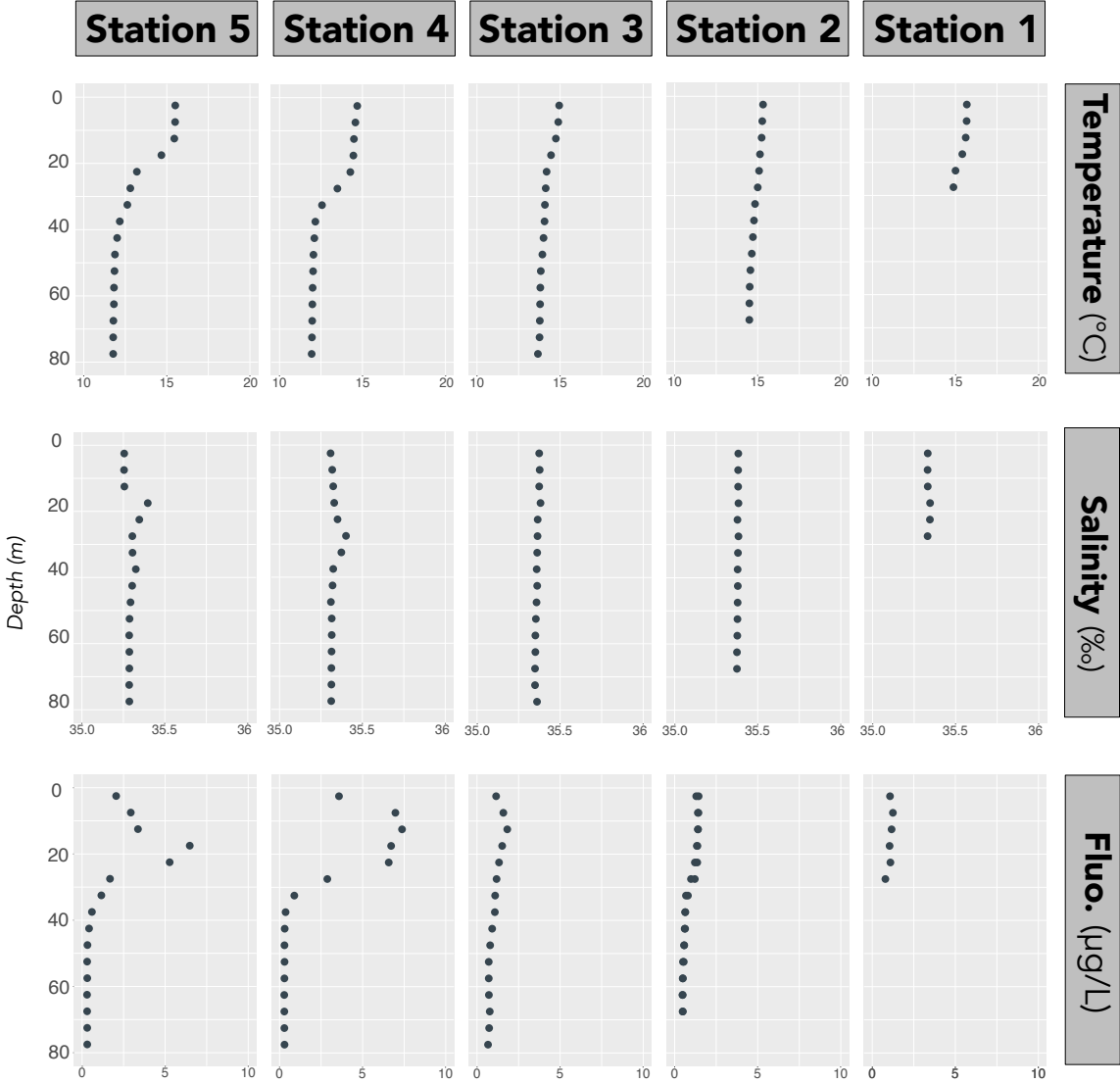


Figure S 1 | CTD profile of the different stations for all the cruises. Values were averaged every 5 meters. Fluo: Fluorescence.

CHAPTER I: SPATIO-TEMPORAL DYNAMIC OF BACTERIAL COMMUNITIES IN THE IROISE SEA

Table S 1 | Additional data for each sampling of M2BiPAT.

ID	Cruise	Date	Station	depth (m)	Depth	Temperature (°C)	Salinity (psu)	Fluorescence (µg/L)	PAR
1	2014_September	11/09/2014	Station 1	3	surface	NA	NA	NA	NA
2	2014_September	11/09/2014	Station 1	6	DCM	16,4	35,337	2,298	128,25
3	2014_September	11/09/2014	Station 1	25	bottom	15,7	35,366	2,071	7,56
4	2014_September	09/09/2014	Station 2	12	surface	14,5	35,413	1,398	NA
5	2014_September	09/09/2014	Station 2	60	bottom	14,2	35,42	0,847	0,13
6	2014_September	10/09/2014	Station 3	10	surface	14,2	35,42	1,247	124,14
7	2014_September	09/09/2014	Station 3	60	bottom	NA	NA	NA	NA
8	2014_September	10/09/2014	Station 4	3	surface	14,5	35,431	1,311	230,11
9	2014_September	10/09/2014	Station 4	8	DCM	14,2	35,447	1,538	184,56
10	2014_September	10/09/2014	Station 4	40	bottom	13,1	35,452	0,751	4,96
11	2014_September	10/09/2014	Station 5	3	surface	18,3	35,419	0,641	695,24
12	2014_September	10/09/2014	Station 5	12	DCM	12,3	35,484	0,531	0,35
13	2014_September	10/09/2014	Station 5	60	bottom	12,4	35,484	0,487	0,28
14	2015_March	10/03/2015	Station 1	5	surface	10,2	34,686	0,746	67,91
15	2015_March	10/03/2015	Station 1	20	bottom	10,3	35,071	0,777	0,58
16	2015_March	12/03/2015	Station 2	10	surface	10,5	35,34	0,509	150,19
17	2015_March	12/03/2015	Station 2	50	bottom	10,5	35,34	0,538	0,44
18	2015_March	10/03/2015	Station 3	10	surface	10,6	35,361	0,501	40,94
19	2015_March	10/03/2015	Station 3	60	bottom	10,6	35,362	0,493	0
20	2015_March	11/03/2015	Station 4	10	surface	10,5	35,362	0,46	112,91
21	2015_March	11/03/2015	Station 4	60	bottom	10,5	35,362	0,492	0
22	2015_March	11/03/2015	Station 5	10	surface	10,6	35,463	0,587	168,97
23	2015_March	11/03/2015	Station 5	60	bottom	10,6	35,463	0,631	0,4
24	2015_July	01/07/2015	Station 1	10	surface	15,7	35,255	1,41	67,56
25	2015_July	01/07/2015	Station 1	30	bottom	14,4	35,298	1,159	3,67
26	2015_July	03/07/2015	Station 2	5	surface	14,2	35,334	1,906	372,84
27	2015_July	03/07/2015	Station 2	50	bottom	13,3	35,392	1,026	0,79
28	2015_July	01/07/2015	Station 3	5	surface	14,8	35,452	0,622	657,67
29	2015_July	01/07/2015	Station 3	20	DCM	13,6	35,381	1,697	105,38
30	2015_July	01/07/2015	Station 3	60	bottom	12,8	35,394	0,575	2,1
31	2015_July	02/07/2015	Station 4	5	surface	16,9	35,164	1,027	851,28
32	2015_July	02/07/2015	Station 4	20	DCM	13,7	35,393	3,678	77,41
33	2015_July	02/07/2015	Station 4	50	bottom	12,2	35,42	0,307	3,75
34	2015_July	02/07/2015	Station 5	5	surface	18,2	35,349	0,514	1192,21
35	2015_July	02/07/2015	Station 5	20	DCM	13,5	35,395	2,955	72,37
36	2015_July	02/07/2015	Station 5	50	bottom	11,7	35,394	0,911	1,26
37	2015_September	08/09/2015	Station 1	3	surface	15,7	35,332	1,002	378,93
38	2015_September	08/09/2015	Station 1	27	bottom	15,0	35,33	0,76	5,87
39	2015_September	10/09/2015	Station 2	3	surface	15,3	35,379	1,319	81,1
40	2015_September	10/09/2015	Station 2	40	bottom	14,8	35,378	0,637	0,27
41	2015_September	08/09/2015	Station 3	3	surface	14,9	35,373	1,094	408,88
42	2015_September	08/09/2015	Station 3	15	DCM	14,6	35,369	1,701	151,11
43	2015_September	08/09/2015	Station 3	50	bottom	14,0	35,357	0,945	2,09
44	2015_September	09/09/2015	Station 4	3	surface	14,8	35,332	2,394	661,63
45	2015_September	09/09/2015	Station 4	25	DCM	14,0	35,318	5,36	3,78
46	2015_September	09/09/2015	Station 4	50	bottom	12,1	35,317	0,308	0,02
47	2015_September	09/09/2015	Station 5	3	surface	15,5	35,26	1,754	450,98
48	2015_September	09/09/2015	Station 5	25	DCM	13,8	35,233	4,512	6,18
49	2015_September	09/09/2015	Station 5	50	bottom	12,2	35,288	0,294	0,75

CHAPTER I: SPATIO-TEMPORAL DYNAMIC OF BACTERIAL COMMUNITIES IN THE IROISE SEA

ID	Turbidity (ntu)	Si(OH) ₄ (μM)	PO ₄ (μM)	NO ₃ +NO ₂ (μM)	NO ₂ (μM)	NH ₄ (μM)	PON (μg.L)	POC (μg.L)
1	NA	1,44	0,02	0,15	0,02	0,15	NA	NA
2	NA	1,52	0,03	0,12	0,01	0,2	NA	NA
3	NA	1,82	0,06	0,43	0,06	0,35	NA	NA
4	NA	3	0,2	3,25	0,27	0,34	NA	NA
5	NA	3,35	0,24	4,85	0,31	0,29	NA	NA
6	NA	3	0,24	3,73	0,28	0,26	NA	NA
7	NA	2,77	0,28	5,14	0,18	0,24	NA	NA
8	NA	2,57	0,13	2,08	0,13	0,13	NA	NA
9	NA	2,61	0,16	2,08	0,13	0,2	NA	NA
10	NA	2,81	0,26	4,81	0,17	0,16	NA	NA
11	NA	0,57	0	0,09	0	NA	NA	NA
12	NA	0,62	0,01	0,13	0	0,46	NA	NA
13	NA	3	0,37	6,41	0,08	0,31	NA	NA
14	1,374	4,38	0,39	12,13	0,21	0,24	38,80	220,20
15	3,866	3,63	0,41	9,46	0,2	0,21	29,51	215,29
16	0,732	2,65	0,34	7,2	0,21	0,1	16,29	102,25
17	1,01	2,67	0,35	7,3	0,22	0,09	19,35	140,50
18	1,73	2,46	0,41	6,91	0,18	0,23	18,65	146,60
19	1,836	2,46	0,42	6,79	0,18	0,21	16,10	75,50
20	0,855	2,82	0,42	7,17	0,2	0,14	15,21	187,74
21	2,224	2,8	0,35	6,48	0,18	0,18	16,21	134,19
22	0,639	1,82	0,39	5,88	0,18	0,15	15,34	80,75
23	0,85	1,86	0,38	5,88	0,25	0,27	13,83	94,54
24	0,445	0,97	0,05	0	0	0,36	58,41	314,62
25	1,809	1,72	0,11	0,52	0,08	0,63	49,51	247,02
26	0,275	0,53	0,05	0,64	0,08	1,93	38,32	143,50
27	0,373	1,1	0,17	2,3	0,16	0,64	77,37	357,56
28	0,247	0,24	0,06	0,24	0,04	0,23	51,86	315,36
29	0,274	0,48	0,11	0,83	0,09	0,5	58,40	304,65
30	0,263	1,02	0,25	2,33	0,2	1,11	27,08	116,52
31	0,258	0,01	0	0	0,01	0,08	39,01	262,03
32	0,26	0,13	0,03	NA	NA	0,08	44,52	287,72
33	0,27	1,27	0,28	3,4	0,25	0,84	21,35	87,44
34	0,277	0,02	0,01	0	0,03	0,08	51,59	331,77
35	0,322	1,2	0,23	3,14	0,2	0,25	38,07	176,52
36	0,334	1,61	0,3	3,99	0,24	0,38	30,14	142,94
37	0,298	2,43	0,07	0,01	0,01	0,13	49,82	283,83
38	0,297	2,92	0,25	2,23	0,23	0,51	31,70	145,53
39	0,317	2,18	0,14	1,02	0,14	0,13	59,58	284,38
40	0,278	2,45	0,25	3,33	0,3	0,46	22,61	106,18
41	0,272	2,4	0,19	2,06	0,22	0,29	43,04	230,66
42	0,282	2,29	0,16	1,96	0,21	0,3	54,65	371,32
43	0,261	2,42	0,37	4,13	0,29	0,32	26,85	130,84
44	0,48	1,65	0,09	0,05	0,01	0,22	74,38	535,40
45	0,448	1,75	0,09	0,08	0,01	0,13	115,03	678,37
46	0,29	2,54	0,44	5,92	0,09	0,06	10,95	49,10
47	0,378	1,07	0,07	NA	NA	0,45	37,49	279,78
48	0,361	2,63	0,42	5,12	0,13	0,13	14,82	73,39
49	0,296	2,87	0,47	6,41	0,12	0,14	15,16	77,04

CHAPTER I: SPATIO-TEMPORAL DYNAMIC OF BACTERIAL COMMUNITIES IN THE IROISE SEA

ID	Diatoms (cells.L)	Dinoflagelleae (cells.L)	Cryptophyceae (cells.L)	Nanophyto-plancton (cells.L)	Total phyto-plancton (cells.L)	Zeaxanthine (ng.L)	Chlorophylle b (ng.L)	Chlorophylle a (ng.L)
1	654280	38720	80600	44300	817900	5	29	1138
2	511770	41000	169180	46390	768340	7	49	1067
3	522060	69600	211470	39650	842780	6	31	816
4	61900	34240	122850	20500	239490	13	100	936
5	37440	13000	54380	18200	123020	12	108	1136
6	66270	16830	44750	754540	882390	8	60	550
7	20990	24190	42290	106750	194220	14	113	1015
8	52880	34370	205430	135200	427880	52	216	1721
9	124890	187790	346410	736110	1395200	43	217	1566
10	19700	37010	66460	125140	248310	14	76	666
11	200	328560	56390	599510	984660	8	31	245
12	160	285610	62100	661260	1009130	NA	NA	NA
13	2640	22260	36250	54540	115690	57	28	0
14	4780	18000	8816	26260	57856	0	49	307
15	3380	0	4000	2000	9380	0	43	271
16	1800	10000	36250	33790	81840	NA	NA	NA
17	1720	6120	31220	6000	45060	0	29	269
18	3260	80	14100	10000	27440	0	26	354
19	3264	40	6000	8040	17344	0	30	1291
20	3408	10150	32200	12040	57798	0	35	266
21	1600	160	2000	8000	11760	0	30	291
22	10645	4040	20140	8000	42825	0	30	365
23	10390	8140	16110	0	34640	0	27	364
24	482040	41953	26180	99580	649753	13	108	680
25	134790	14520	50350	102740	302400	6	101	695
26	349094	175652	58406	207667	790819	NA	NA	NA
27	76696	9100	10100	135410	231306	NA	NA	NA
28	307965	73340	30250	253330	664885	6	53	872
29	12400	NA	NA	NA	NA	8	108	1144
30	NA	12770	12100	52440	89710	0	13	264
31	38840	125080	58410	95400	317730	9	64	549
32	428050	85830	46370	106040	666290	8	100	1480
33	10960	240	2000	4140	17340	0	8	159
34	25615	103450	14130	988320	1131515	8	53	778
35	26170	71888	11190	156804	266052	0	18	722
36	6730	76566	6040	70880	160216	0	13	499
37	15284	137960	517598	437036	1107878	68	138	465
38	13280	95200	223778	337887	670145	15	83	314
39	15890	118250	360510	451540	946190	22	165	1272
40	2920	24250	26210	243340	296720	8	65	260
41	334550	64570	35450	298736	733306	20	227	1395
42	12090	135880	462304	769405	1379679	31	310	2181
43	13746	10120	82570	178590	285026	7	87	491
44	2670	3909040	624590	2403874	6940454	26	115	1704
45	2950	2562782	772518	1210887	4548857	34	158	5448
46	1300	4580	28200	72560	106640	5	37	217
47	8380	1685680	292850	255760	2242670	51	84	1241
48	680	198820	104730	129870	434100	14	59	486
49	640	6520	34200	2000	43360	8	32	194

A.

	Sept. 2014	March 2015	July 2015	Sept. 2015
Date	9, 10 and 11	10, 11 and 12	1, 2 and 3	8, 9 and 10
Tide coefficient	100-115	60-78	79-93	44-68

B.

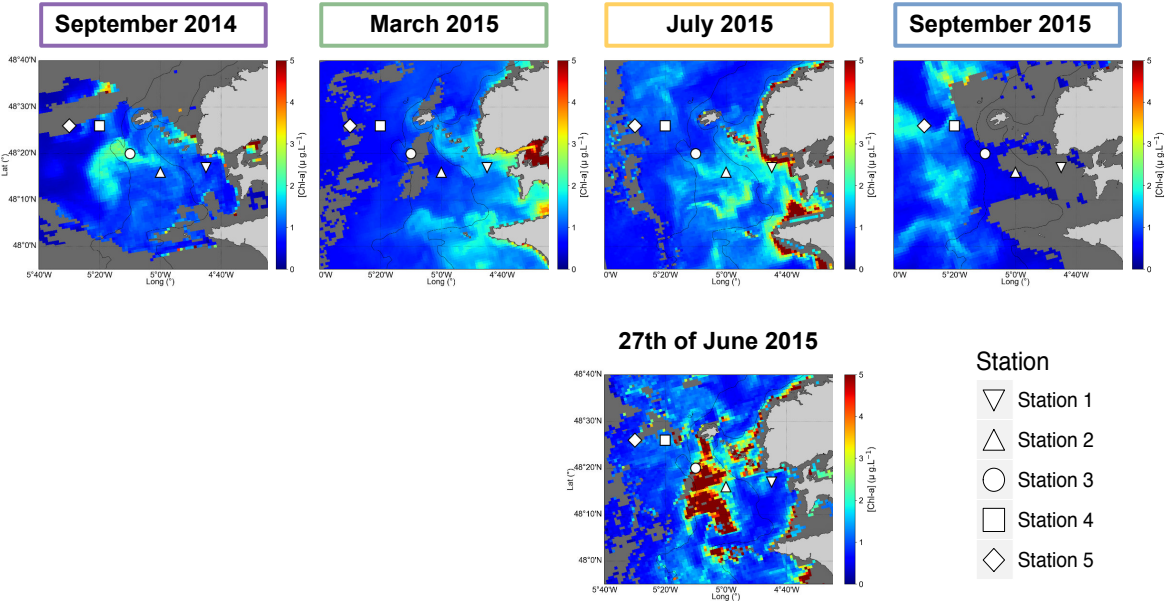


Figure S 2 | General information about the cruises (A.) and satellite surface Chla measurements (B.). The shapes correspond to the different stations sampled during each cruise.

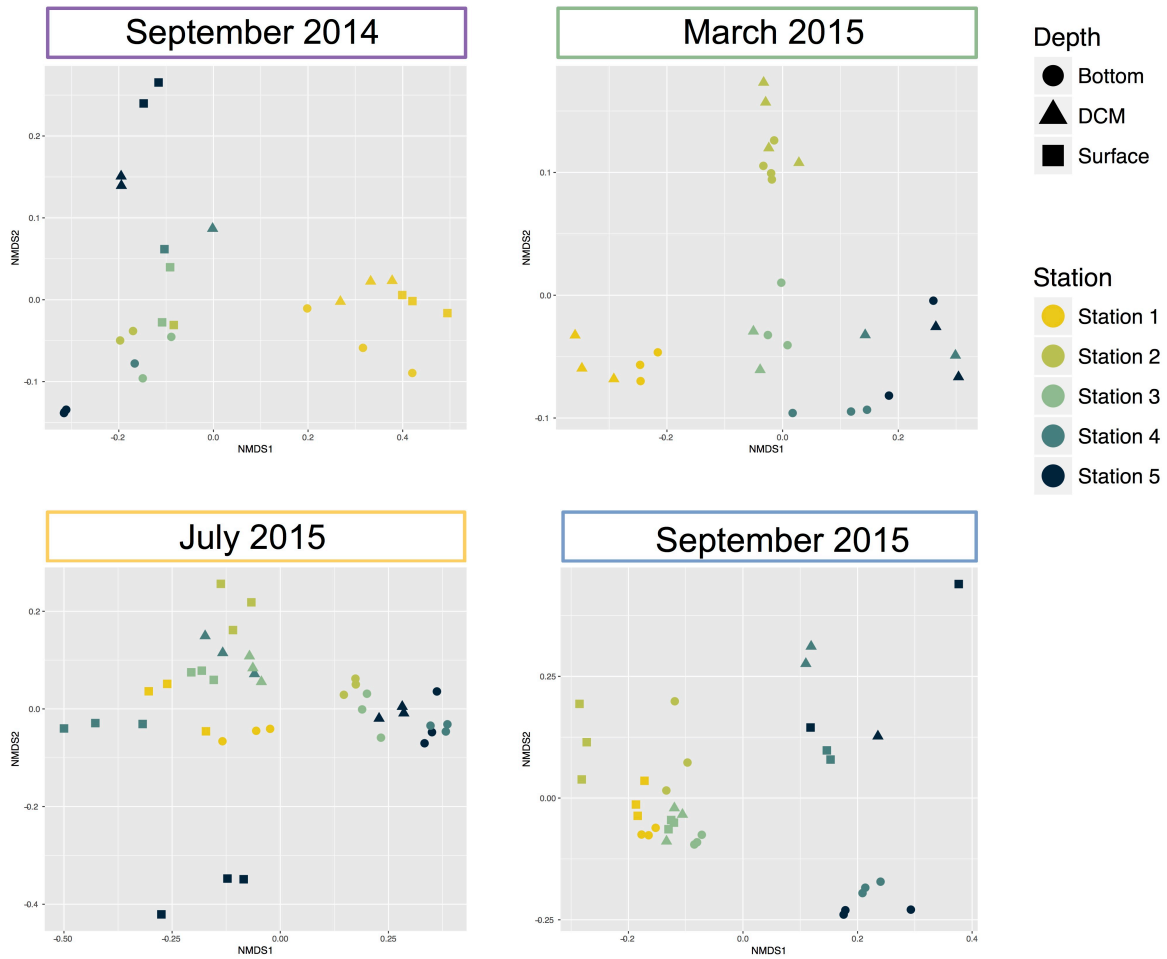


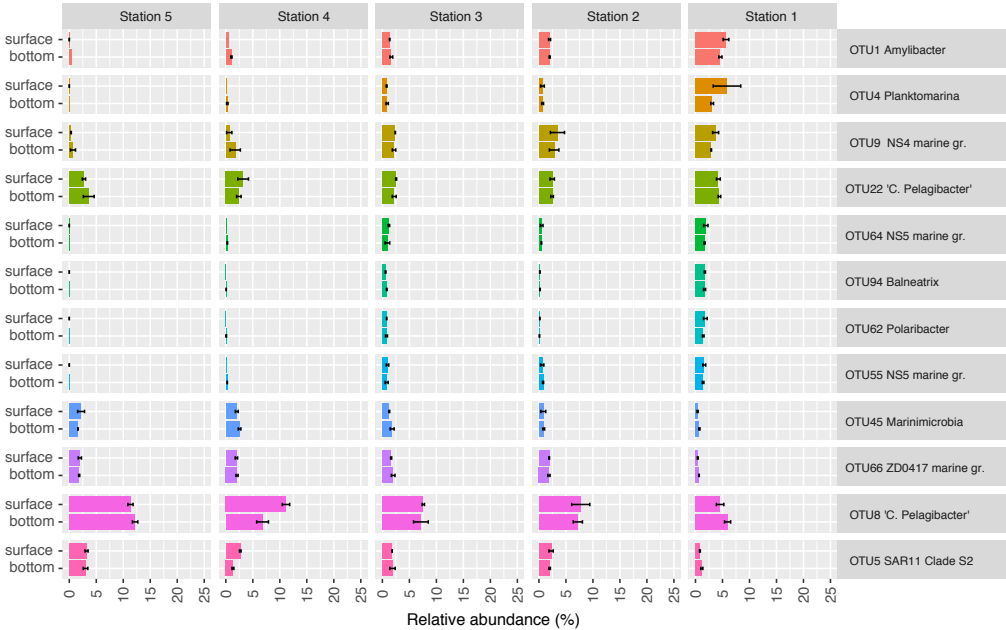
Figure S 3 | NMDS of the bacterial diversity based on Bray-Curtis dissimilarities among each sampling. Stratification (Station 5) clearly separates distinct communities between the surface and depth. This trend is less observed in mixed waters (Station 2 and 3) except in July. In March there is no distinction of the communities with depth but a clear coastal to offshore gradient is conspicuous

CHAPTER I: SPATIO-TEMPORAL DYNAMIC OF BACTERIAL COMMUNITIES IN THE IROISE SEA

A.



B.



C.



D.

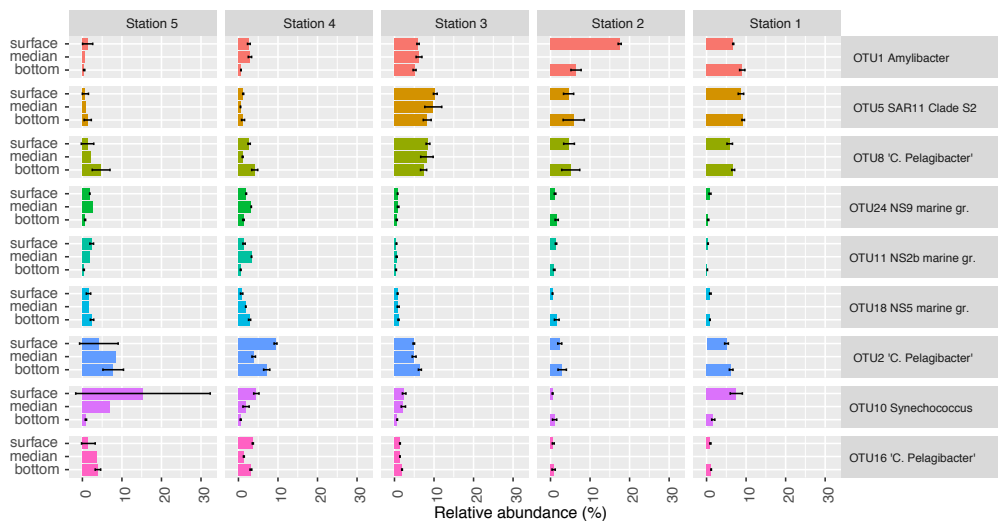


Figure S 4 | Most significant biomarkers (LDA > 4) of a station within each cruise found with LEfSe algorithm. Their relative abundance consisted of the mean of the different replicates for each sampling, with the standard deviation. **A.** September 2014, **B.** March 2015, **C.** July 2015, **D.** September 2015.

CHAPTER I: SPATIO-TEMPORAL DYNAMIC OF BACTERIAL COMMUNITIES IN THE IROISE SEA

com_5			com_13			com_7		
OTU	Taxonomy	abond. (%)	OTU	Taxonomy	abond. (%)	OTU	Taxonomy	abond. (%)
OTU2	<i>Candidatus_Pelagibacter(100)</i>	18,91	OTU37	<i>NS5_marine_group(100)</i>	7,09	OTU9	<i>NS4_marine_group(100)</i>	47,86
OTU8	<i>Candidatus_Pelagibacter(96)</i>	15,70	OTU66	<i>ZD0417_marine_group(100)</i>	6,34	OTU40	<i>Flavobacteriaceae(100)</i>	6,17
OTU22	<i>Candidatus_Pelagibacter(100)</i>	6,94	OTU45	<i>Marinimicrobia(100)</i>	6,11	OTU55	<i>NS5_marine_group(100)</i>	5,97
OTU17	<i>ZD0405(100)</i>	6,61	OTU93	<i>SAR86_clade(98)</i>	4,71	OTU62	<i>Polaribacter(100)</i>	5,27
OTU15	<i>ZD0405(98)</i>	5,46	OTU176	<i>SAR11_Surface_1(100)</i>	4,21	OTU64	<i>NS5_marine_group(100)</i>	5,05
OTU16	<i>Candidatus_Pelagibacter(100)</i>	5,23	OTU106	<i>Marinimicrobia(100)</i>	3,82	OTU81	<i>NS9_marine_group(100)</i>	4,92
OTU20	<i>Candidatus_Actinomorina(100)</i>	4,46	OTU97	<i>Pseudospirillum(100)</i>	2,90	OTU94	<i>Balneatrix(100)</i>	3,65
OTU18	<i>NS5_marine_group(100)</i>	3,48	OTU115	<i>Pseudospirillum(100)</i>	2,51	OTU152	<i>Sulfotibacter(100)</i>	2,21
OTU26	<i>OCS116_clade(100)</i>	3,43	OTU123	<i>Defluviococcus(100)</i>	2,44	OTU92	<i>Owenweeksia(99)</i>	1,70
OTU24	<i>NS9_marine_group(100)</i>	3,22	OTU112	<i>NS9_marine_group(100)</i>	2,32	OTU191	<i>OM182_clade(100)</i>	1,58

com_6			com_3			com_11		
OTU	Taxonomy	abond. (%)	OTU	Taxonomy	abond. (%)	OTU	Taxonomy	abond. (%)
OTU138	<i>KI89A_clade(100)</i>	8,36	OTU10	<i>Synechococcus(99)</i>	68,45	OTU5	<i>SAR11_Surface_2(100)</i>	66,82
OTU137	<i>NS9_marine_group(99)</i>	7,34	OTU57	<i>Synechococcus(100)</i>	15,40	OTU196	<i>SAR11_Clade(100)</i>	3,55
OTU136	<i>Balneatrix(100)</i>	6,88	OTU180	<i>Synechococcus(95)</i>	2,56	OTU238	<i>SAR11_Surface_2(98)</i>	3,41
OTU184	<i>Formosa(100)</i>	6,53	OTU237	<i>Synechococcus(100)</i>	2,53	OTU302	<i>SAR11_Surface_1(100)</i>	3,35
OTU222	<i>Owenweeksia(91)</i>	5,34	OTU296	<i>Synechococcus(100)</i>	1,96	OTU206	<i>SAR11_Surface_1(100)</i>	3,12
OTU162	<i>Chitinophagaceae(100)</i>	5,25	OTU201	<i>Porticoccus(100)</i>	1,75	OTU300	<i>SAR11_Clade(100)</i>	1,80
OTU233	<i>Owenweeksia(100)</i>	4,62	OTU359	<i>NS9_marine_group(100)</i>	1,49	OTU271	<i>SAR11_Clade(100)</i>	1,76
OTU181	<i>SAR86_clade(100)</i>	4,34	OTU608	<i>Asciadiaceihabitan(83)</i>	0,88	OTU231	<i>SAR11_Clade(100)</i>	1,75
OTU235	<i>NS7_marine_group(100)</i>	3,41	OTU1664	<i>Synechococcus(95)</i>	0,77	OTU373	<i>SAR11_Surface_2(92)</i>	1,74
OTU266	<i>NS9_marine_group(100)</i>	2,80	OTU495	<i>Flavobacteriales(100)</i>	0,70	OTU214	<i>SAR11_Clade(100)</i>	1,61

com_4			com_2			com_1		
OTU	Taxonomy	abond. (%)	OTU	Taxonomy	abond. (%)	OTU	Taxonomy	abond. (%)
OTU1	<i>Amylibacter(100)</i>	41,35	OTU3	<i>Tenacibaculum(100)</i>	71,18	OTU11	<i>NS2b_marine_group(99)</i>	34,14
OTU7	<i>SAR86_clade(100)</i>	11,01	OTU147	<i>Tenacibaculum(98)</i>	2,21	OTU33	<i>SAR86_clade(100)</i>	10,23
OTU4	<i>Planktomarina(100)</i>	8,97	OTU148	<i>Formosa(99)</i>	1,43	OTU29	<i>NS4_marine_group(100)</i>	9,76
OTU25	<i>SAR116_clade(100)</i>	5,90	OTU260	<i>Rhodobacteraceae(97)</i>	1,20	OTU44	<i>Asciadiaceihabitan(94)</i>	7,20
OTU32	<i>Candidatus_Puniceispirillum(100)</i>	3,69	OTU295	<i>Flavobacteriaceae(100)</i>	1,14	OTU69	<i>Formosa(100)</i>	4,90
OTU27	<i>OM43_clade(100)</i>	3,24	OTU338	<i>Flavobacteriaceae(100)</i>	0,93	OTU51	<i>Polaribacter(100)</i>	4,77
OTU78	<i>NS9_marine_group(100)</i>	2,03	OTU336	<i>Flavobacteriaceae(100)</i>	0,79	OTU68	<i>NS4_marine_group(100)</i>	4,08
OTU73	<i>Rhodobacteraceae(100)</i>	1,75	OTU276	<i>Tenacibaculum(99)</i>	0,78	OTU85	<i>Puniceococcaceae(100)</i>	3,92
OTU75	<i>Rhodobacteraceae(100)</i>	1,60	OTU294	<i>NS4_marine_group(90)</i>	0,70	OTU111	<i>Lentimonas(100)</i>	3,78
OTU43	<i>NS5_marine_group(98)</i>	1,56	OTU364	<i>Tenacibaculum(100)</i>	0,69	OTU91	<i>Thiothrix(99)</i>	2,88

com_8			com_9		
OTU	Taxonomy	abond. (%)	OTU	Taxonomy	abond. (%)
OTU6	<i>Asciadiaceihabitan(97)</i>	13,06	OTU35	<i>Allivibrio(94)</i>	53,99
OTU14	<i>Owenweeksia(100)</i>	6,06	OTU74	<i>Pseudoalteromonas(100)</i>	21,65
OTU23	<i>Formosa(100)</i>	5,86	OTU116	<i>Pseudoalteromonas(100)</i>	4,96
OTU21	<i>Fluviicola(100)</i>	5,07	OTU242	<i>Pseudoalteromonas(100)</i>	4,06
OTU41	<i>SAR92_clade(100)</i>	4,56	OTU289	<i>Marinomonas(100)</i>	2,87
OTU42	<i>NS5_marine_group(100)</i>	3,34	OTU376	<i>Vibrionaceae(100)</i>	2,10
OTU31	<i>Roseibacillus(100)</i>	2,86	OTU525	<i>Alteromonas(100)</i>	1,77
OTU38	<i>OM60(NORS)_clade(100)</i>	2,73	OTU895	<i>Vibrionaceae(81)</i>	1,44
OTU63	<i>SAR92_clade(100)</i>	2,72	OTU868	<i>Alteromonas(96)</i>	1,43
OTU65	<i>Formosa(100)</i>	2,70	OTU1044	<i>Flavobacteriaceae(100)</i>	1,13

Figure S 5 | Dominant OTUs found within each module, with their taxonomy and relative abundance (in % of the read) in the module.

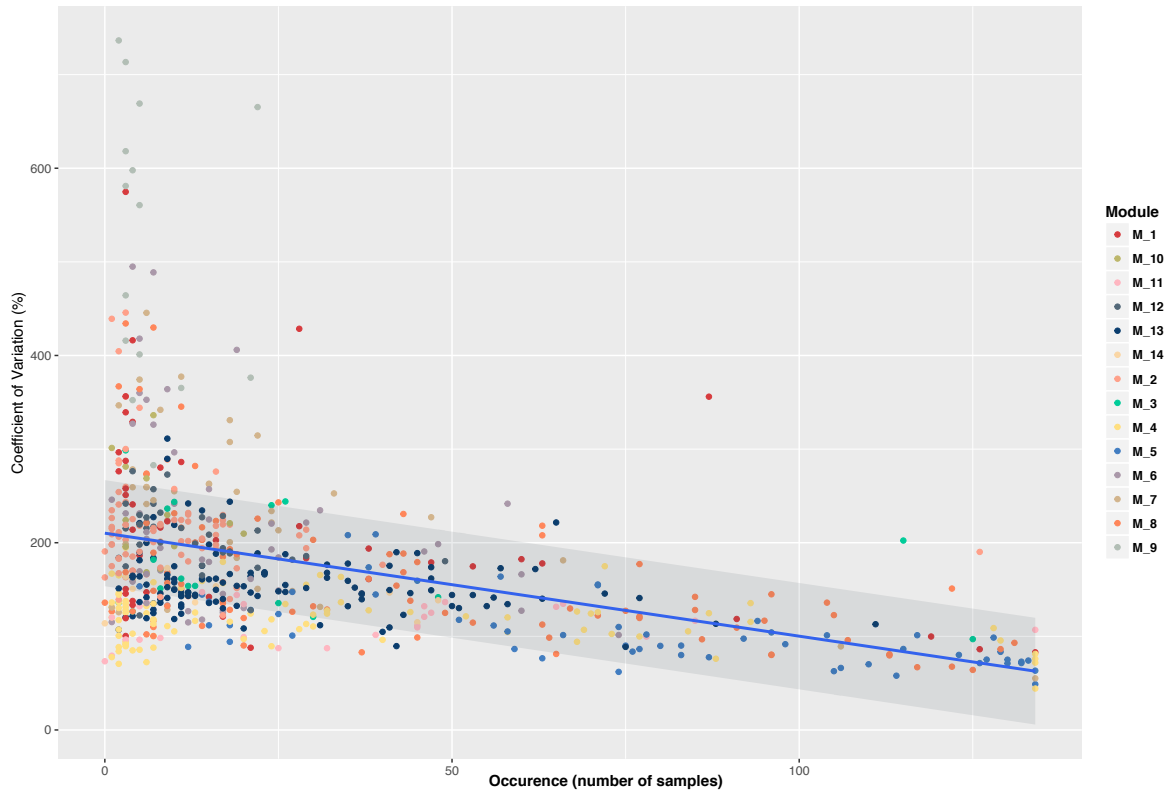


Figure S 6 | Plot of the coefficient of variation (CV) of an OTU with its occurrence in a sample (*i.e.* relative abundance >0.05%). The linear model is the blue line and the prediction interval (0.5) above and under the linear model is in grey. High CV OTUs are those above the prediction interval and Low CV OTUs are those below. Each OTU is colored depending on its module membership.

Chapter II.

Temporal dynamic of bacterial communities in the Bay of Brest

1. RÉSUMÉ EN FRANÇAIS

Nos connaissances actuelles des communautés microbiennes reposent en large partie sur des suivis temporels qui offrent la possibilité d'observer de façon étroite le lien entre la dynamique des communautés et leur environnement. Ceci afin de comprendre les facteurs importants dans la dynamique des communautés bactériennes et d'anticiper l'effet de potentiels changements globaux et anthropiques. La Rade de Brest (Bretagne, France) représente un environnement intéressant pour suivre la dynamique des communautés bactériennes. D'un côté c'est un environnement macrotidal complexe, typique des eaux tempérées qui subit l'influence de deux rivières importantes, ainsi que de la mer d'Iroise, au rythme des marées. La Station SOMLIT en Rade de Brest étudie depuis 1998 les évolutions de l'environnement en mesurant toutes les semaines plus de 15 paramètres physiques, chimiques et biologiques. Ce chapitre de thèse présente les résultats des 3 premières années d'un suivi microbien mis en place en Juillet 2014 au niveau de cette station. Malgré certaines périodes manquantes dans le suivi, la dynamique temporelle des bactéries montre une saisonnalité remarquable. C'est d'autant plus intéressant dans un système dynamique côtier comme la rade de Brest, avec une forte influence de deux rivières et des variations sur l'intensité des blooms printaniers observés lors de notre échantillonnage. Une analyse en réseau a été utilisée pour observer les groupes de bactéries qui se succèdent sur une année. Son extrapolation aux autres années du suivi montre que ces groupes ont une dynamique saisonnière cohérente d'une année sur l'autre. Cette dynamique saisonnière présente un changement rapide lors de la mise en place des blooms printaniers. Cela s'observe par des pics de différentes bactéries copiotrophes, pouvant atteindre jusqu'à 30% de la diversité en 1 semaine. Ce qui influence la dynamique en fin d'été vers l'hiver reste encore à être élucidée, mais peut s'expliquer par plusieurs facteurs, comme les variations générales de température, l'influence des rivières et de la salinité ou encore une augmentation en nutriments inorganiques. Ces résultats pourront être confirmés et mieux compris avec l'intégration d'années d'échantillonnage supplémentaires.

Temporal dynamic of microbial communities in the Bay of Brest

C. Lemonnier^{1,2}, M. Perennou², H. Morrison³, C. Paillard² and L. Maignien^{1,3}

¹ Univ Brest, CNRS, IFREMER, Laboratoire de Microbiologie des Environnements Extrêmes, F-29280 Plouzané, France

² Univ Brest, CNRS, IRD, IFREMER, Laboratoire des Sciences de l'Environnement Marin, F-29280 Plouzané, France

³ Josephine Bay Paul Center for Comparative Molecular Biology and Evolution, Marine Biological Laboratory, Woods Hole, USA

Draft paper

2. ABSTRACT

Temporal surveys are essential to decipher the main environmental drivers of microbial communities dynamic and better predict the effect of potential global and anthropogenic changes. The Bay of Brest (Brittany, France) represents a model environment for monitoring bacterial communities dynamic. It is a complex macrotidal water body, typical of temperate coastal waters with both global and local anthropic influences. Since 1998, the SOMLIT Station in the Bay of Brest has been studying temporal changes in the environment by measuring more than 15 physical, chemical and biological parameters every week. This chapter presents the results of the first 3 years of a microbial monitoring implemented in July 2014 at this station. Despite some missing periods in the survey, temporal dynamics of the bacteria showed a remarkable seasonality. A network analysis was used to observe groups of bacteria, those that covariates within the same year. Its extrapolation to other years of monitoring shows that these groups have a consistent seasonal dynamic from one year to the next. This seasonal dynamic presents a rapid change induced by the seasonal spring blooms. This was observed by peaks of different copiotrophic bacteria, which can reach up to 30% of diversity within a single week period. What influences the dynamics in late summer to winter has yet to be elucidated, but could be explained by several factors, such as general temperature variations, the influence of rivers and salinity or an increase in inorganic nutrients. These conclusions should be confirmed and better understood by integrating additional years of sampling to be carried out in a long-term observatory of microorganisms.

3. INTRODUCTION

Coastal marine environments are important areas of recognized ecological values (Costanza et al., 2014). Thanks to terrestrial input, they are among the most fertilized systems worldwide promoting high primary and secondary productivity (Nixon and Buckley, 2002). In temperate coastal waters, these nutrient loading by rivers support large phytoplankton blooms during the spring period (Del Amo et al., 1997). Marine microbial communities are major component of marine ecosystem functioning where they are involved in key processes such as organic matter degradation and global biogeochemical cycling (Azam and Malfatti, 2007; Falkowski et al., 2008; Madsen, 2011). Time series have improved our understanding of bacterial community composition dynamic related to numerous key environmental factors such as temperature, day length, phytoplankton blooms, or inorganic nutrients (Gilbert et al., 2012; Lucas et al., 2015). In temperate waters, they led to the discovery that free-living bacterial communities present strong recurrent seasonal patterns (Chow et al., 2013; Fuhrman et al., 2015; Lambert et al., 2018) and that are linked to changes in global functions of bacterial communities (Galand et al., 2018). In the last decades, anthropogenic pressure on these ecosystems has dramatically increased due to agricultural and industrial practices (Rabalais et al., 2002). Hence, there is a pressing need to better understand the ecological functioning of these systems to better monitor their evolution and understand how these respond to local perturbations and global changes of the marine environment.

The Bay of Brest is a macrotidal, well-mixed system typical of Western Europe (Delmas, 1981). This shallow Bay (around 8 m depth on average) is a productive coastal area, thanks to nutrients inputs from the Elorn and Aulne rivers. It is influenced by global climate changes with effect of the North Atlantic Oscillation, as well as local anthropogenic changes with for instance the dramatic increase of nitrates loads of terrestrial origin over the past decades (Goberville et al., 2010; Tréguer et al., 2014). Since 1998, a station of a French coastal waters survey (SOMLIT) weekly sampled seawater for numerous physical, chemical and biological parameters measurements in order to deeply characterize the environment and its evolution throughout the seasons and years. Such established long term

environmental monitoring represent a good opportunity to link major geochemical cycles with pelagic microbial community dynamics.

In this study we analyzed the first 2-3 years of a microbial observatory at the SOMLIT station in the Bay of Brest. We aimed at describing the dynamics of free-living bacterial communities during the survey and better understand the factors driving this dynamic.

4. MATERIAL AND METHODS

4.1. 1. STUDY SITE AND SAMPLING STRATEGY

The sampling station for microbiology was located at the SOMLIT station in the Bay of Brest in the temperate North-East Atlantic Ocean. The SOMLIT long term observatory station used in this study was located in le *Goulet* (48°21'32.17" N, 4°33'07.21" O), a very narrow channel linking the Bay of Brest to the Atlantic Ocean. The water of those two areas tightly communicate through tide currents that either bring offshore waters into the bay of Brest at rising tide, or bring brackish waters of the Bay of Brest in the Iroise sea at ebbing tide. To get a repeatable temporal survey, the sampling was done at high tide at a coefficient of around 70 at 2 m depth. Within the SOMLIT survey, numerous physical, chemical and biological parameters are weekly sampled and measured since 1998.

For bacterioplankton sampling, five liters of seawater were collected in triplicate with a Niskin bottle and directly transferred in sterile autoclaved and HCl 1% rinsed carboys. Filtrations started in laboratory, 30 minutes after the sampling. Water samples were size-fractionated using three in line filters of different porosity: 10 μm , 3 μm (PC membrane filters, Millipore®) and 0.22 μm (Sterivex® filters), in order to separate the different class of phytoplankton. For bacteria it allows to distinguish free-living (0.22 - 3 μm) and attached (> 3 μm) lifestyle. In this study, we focused only on the free-living fraction. All the 3 filters were stored at -80°C. In total, 2 to 5 liters of seawater were filtered each time, depending on filters' clogging.

The microbial survey started on July 24, 2014 and ran until May 31 2017, covering almost 3 years. However, severable notable gaps are present: there were no sampling between September the 6th and November the 12th in 2014 as well as between June the 12th and October the 9th in 2015. Only the year 2016 was sampled continuously. In addition, the frequency of sampling changed from weekly in the beginning of the survey to a bimonthly sampling starting from October 9, 2015.

4.2. 2. DNA EXTRACTION AND SEQUENCING FOR BACTERIAL DIVERSITY

For DNA extractions, filters were directly placed in a Tube matrix B® (MP Biomedicals) with lysis buffer (Tris, HCl, EDTA), SDS 10% and Sarkosyl 10%. Bead-beating was performed with an optimized time of 5 minutes on a plate vortex. The liquid fraction was then collected in order to perform a phenol-chloroform extraction using PCI (25:24:1) and a precipitation step with Isopropanol and 5 M sodium acetate. DNA was suspended in 100 µL of sterile water. Bacterial diversity was assessed targeting the v4-v5 hypervariable regions of 16S rDNA with the primers 518F (CCAGCAGCYGCGGTAAN) / 926R (CCGTCAATTCNTTTRAGT - CCGTCAATTTCTTTGAGT - CCGTCTATTCCTTTGANT). All PCR were performed in triplicates before being pooled together. PCR products were purified using AMPure XP® kit and then DNA quantity was measured using PicoGreen®. Each sample was diluted to the same concentration before being pooled together in equimolecular quantities. Sequencing was performed in an Illumina MiSeq sequencer at the Marine Biological Laboratory (Woods Hole, USA).

4.3. 3. BIOINFORMATICS ANALYSIS

Raw reads obtained after the sequencing were merged and quality-filtered according to recommendations in Minoche et al., 2011, using the Illumina-Utills scripts (Eren et al., 2013c). We then used the Swarm algorithm (Mahé et al., 2015) to cluster the remaining sequences into operational taxonomic unit (OTU). Chimera detection was done using Vsearch *de novo* (Rognes et al., 2016), directly on the swarm OTU representative sequences as suggested by Mahé et al., 2015. Taxonomic assignation for each OTU was done with the SILVA database v132 (Quast et al., 2013) using Mothur classify.seqs command (Schloss et al., 2009). Based on this affiliation we then removed sequences affiliated to unwanted taxa (e.g. Archaea, Chloroplasts, Mitochondria, Eukaryota).

4.4. 4. STATISTICAL ANALYSIS

Measures of community characteristics of each sample (OTU richness and Shannon index) were calculated on the raw read counts OTU observation table. For comparison of bacterial community composition across samples, a normalization was done using DESeq package in R (Anders and Huber, 2010). In brief, a size factor of each sample is calculated, bringing all count values to a common scale in order to adjust for differences in sequencing depth. We used the swarm algorithm to partition the dataset in OTUs, and performed a frequency-based filtration of the rarest OTUs removing those present in less than 10% of the samples. Based on this filtered matrix, bacterial community dynamics across the seasons and the years was explored using NMDS based on Bray-Curtis dissimilarities between samples. We carried out Canonical Correspondence Analysis (CCA) to identify environmental variables having the most significant influence on the bacterial community composition. Environmental variables that presented collinearity with other variables were removed using Variation Inflation Factor (VIF, only variables with $VIF < 10$ were kept). Significance of the different remaining factors and robustness of the CCA model were tested using an ANOVA.

4.5. 5. LOOKING FOR GENERAL PATTERNS OF SEASONALITY

Network analysis was used in order to define groups of OTUs that covariates together across the different seasons. However, network analysis are sensible to gaps in the dataset or to changes in frequency. We thus performed the analysis in the year 2016 which was the only one fully complete with a bimonthly sampling. The matrix was filtered to keep only the most abundant 381 OTUs, with at least 50 sequences in 3 samples. Our module detection analysis followed in part the pipeline assembled by Fernandez-Guerra and available online (<https://github.com/genomewalker/>). First, pairwise co-variance based correlation between OTUs was calculated with the SPIEC-EASI software that addresses the sparsity and compositionality issues inherent of microbial abundance data (Kurtz et al., 2015). Correlation groups were then inferred using the glasso probabilistic inference model with a lambda.min.ratio of 0.01. Based on this model, only the significant covariance values were extracted

and transformed in a correlation matrix using the `cov2cor` function. We then delineated network modules as groups of highly interconnected OTUs that presented very closely related correlated changes in the studied samples. Those modules were defined using the Louvain algorithm (Blondel et al., 2008). Module eigengenes (ME, a Principal Component Analysis of the different OTUs abundance in a module) was calculated based on the relative abundance matrix using WGCNA function `moduleEigengenes` (Langfelder and Horvath, 2008). Those ME were used to calculate Pearson correlations between the different modules and the environmental variables, using the WGCNA commands `moduleTraitCor`, `moduleTraitPvalue`.

5. RESULTS

5.1. ENVIRONMENTAL SETTINGS

Seasonal changes of the different environmental parameters are presented in Figure 12. They show typical seasonal variations for coastal temperate systems. In winter, the temperatures were low, around 10°C and with higher inorganic nutrients concentrations. Despite a sampling at rising tide, influence of more brackish water was conspicuous at this station, as salinity was lower in winter. In February 2016, exceptional lowering of salinity to 31.7‰ co-occurred with high nutrients concentrations (NO_3^- : 37.1 μM). Measure of chlorophyll a was used as a proxy of phytoplankton biomass. It showed that the spring phytoplankton bloom occurred in the end of winter between late March and April, depleting the winter stock of inorganic nutrients. The intensity of spring blooms was variable between the 3 years of sampling: it was particularly high in 2015 (6.25 $\mu\text{g}\cdot\text{L}^{-1}$), and low in 2017 (1.73 $\mu\text{g}\cdot\text{L}^{-1}$), which was lower than the late summer bloom in 2014 (2.18 $\mu\text{g}\cdot\text{L}^{-1}$). The temperatures were warmest between July and September, reaching 18°C. At this time of the year, inorganic nutrients concentrations remained very low, at limiting concentrations for primary producers (lower than 1 μM for nitrates). They started to increase again in late summer at the end of the production period in early October.

5.2. BACTERIOPLANKTON COMMUNITY'S CHARACTERISTICS

Abundance of bacteria in the Bay of Brest indicated by flow cytometry counts were typical of coastal environment as it varied between 1.5×10^5 and 2.5×10^6 cells.mL⁻¹. This abundance was variable across the years but was overall lowest during winter (Figure S 7).

A total number of 983 723 swarm OTUs for 14 660 645 sequences were retrieved from the 215 samples. After the filtration of the OTUs present in more than 10% of the samples, we obtained 5828 OTUs representing 93% of the initial sequence number. Diversity of OTUs within each sample measured by Shannon index was the highest in winter and particularly in late January and February (Figure S 8) and was lowest in April, simultaneously with the set-up of spring phytoplankton blooms. Taxonomic

affiliation of the OTUs showed that bacterial communities were dominated by two main phyla: *Proteobacteria* (58% of the sequences) and *Bacteroidetes* (33%). The *Proteobacteria* mostly consisted of member of *Alphaproteobacteria* (75% of *Proteobacteria* sequences) with a great abundance of *SAR11 Clade* (34% of *Alphaproteobacteria* sequences) and *Rhodobacterales* (30%) order. *Bacteroidetes* were dominated by *Flavobacteriales* (95%) order. The most dominant OTUs, abundant (*i.e.* relative abundance > 1%) in more than 75% of all the samples were affiliated to *SAR11 Clade* (OTU2, OTU4 and OTU11), *Rhodobacteraceae* with OTU1 (*Amylibacter sp.*) and OTU3 (*Planktomarina sp.*) and to *Flavobacteriaceae* with OTU14 (*NS4 marine group*). Together, they represented between 13% in April 2015 and up to 42% of the sequences in March 2015 (Figure S 9).

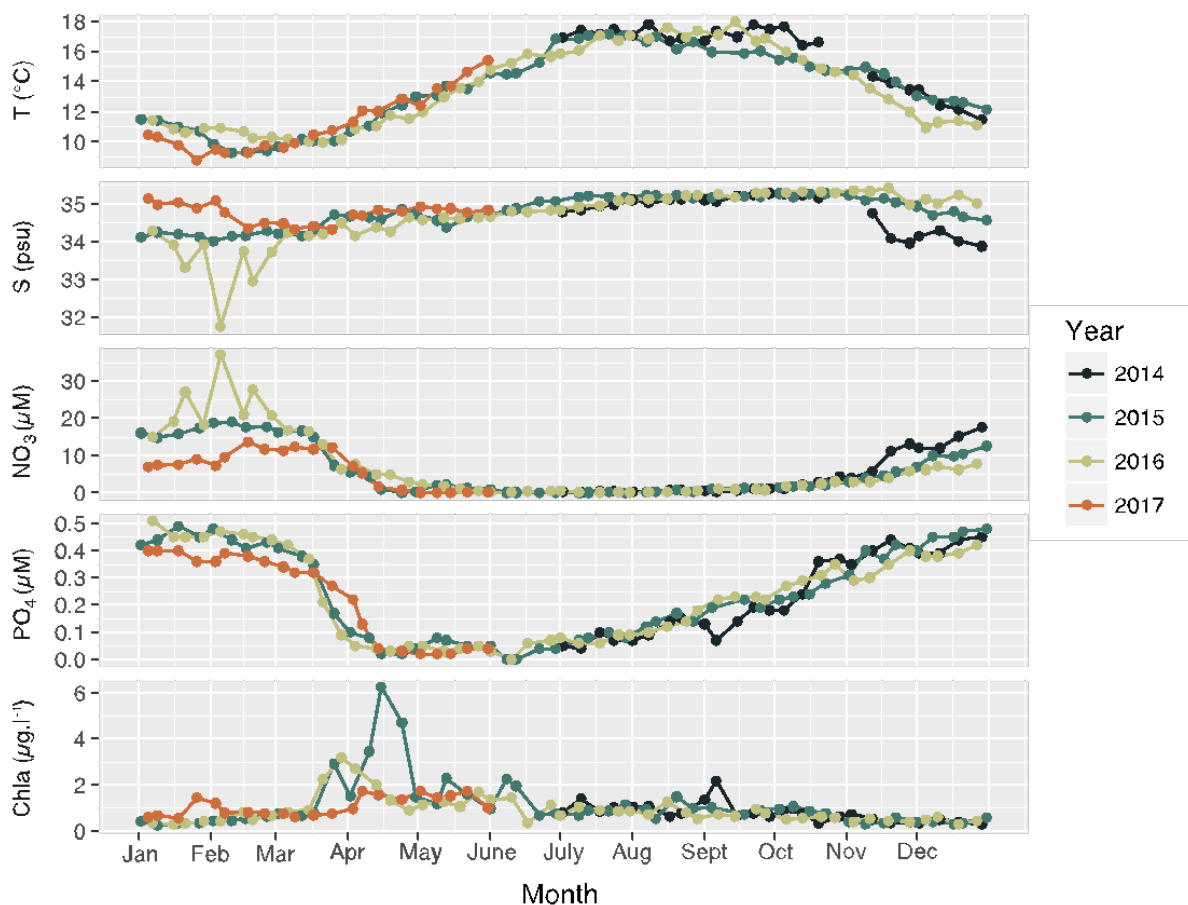


Figure 12 | Seasonal variations of the environment in the Bay of Brest. Variation of Temperature (T), Salinity (S), Nitrates (NO₃) Phosphates (PO₄) and Chlorophyll a (Chla) between 2014 and 2017 at the SOMLIT station in the Bay of Brest.

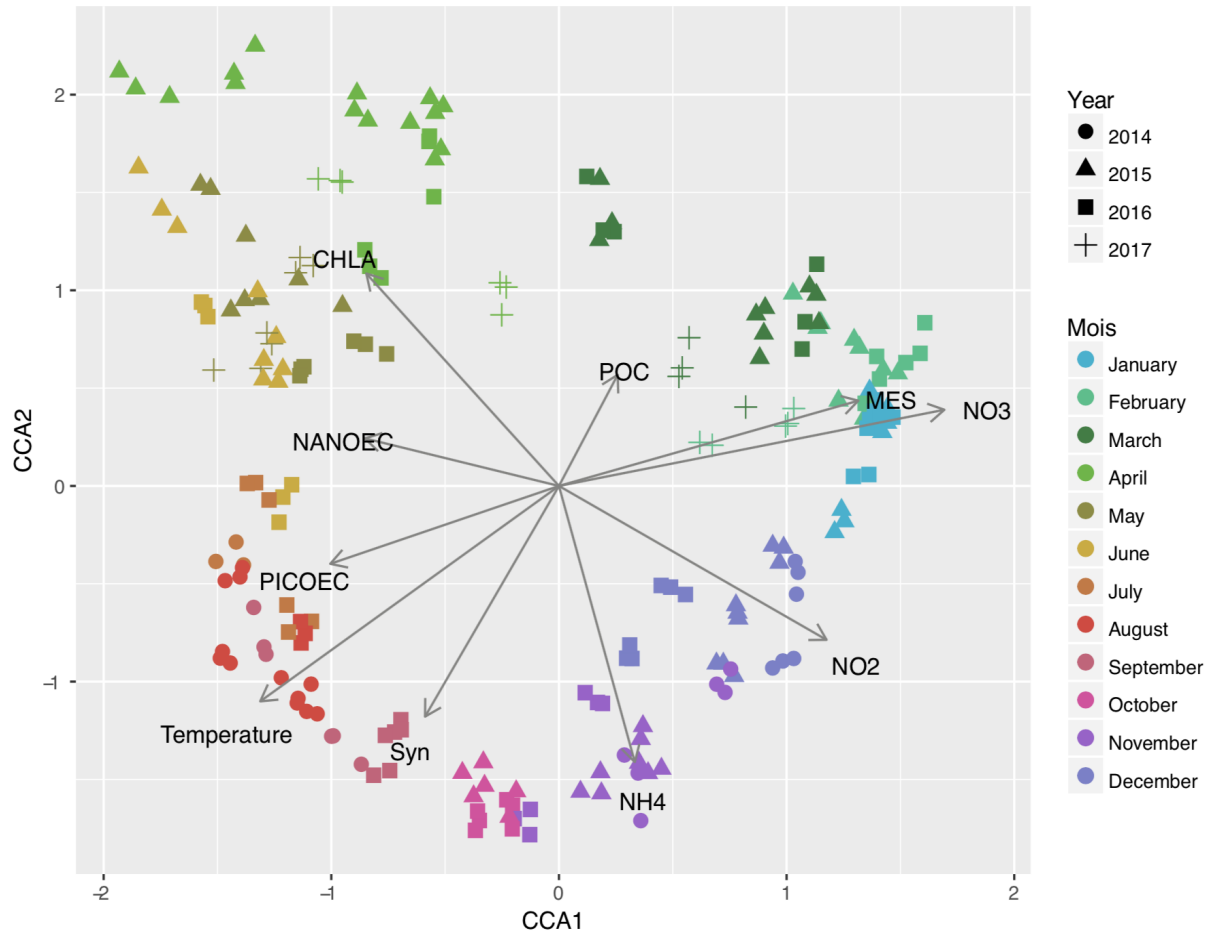


Figure 13 | Seasonality of the bacterioplankton communities. Canonical Correspondence Analysis (CCA) of the bacterial community composition across the samples and in relation with the environmental parameters: chlorophyll a (CHLA), particulate organic carbon (POC), suspended material (MES), nitrates (NO3), nitrites (NO2), ammonium (NH4), temperature and cytometric counts of *Synechococcus* (Syn), Picoeucaryotes (PICOEC) and Nanoeucaryotes (NANOEC).

5.3. SEASONALITY OF BACTERIAL COMMUNITIES

Changes in bacterioplankton community composition and their relation with biogeochemical variables were investigated using a CCA (Figure 13). Bacterial communities in the Bay of Brest presented a strong seasonality which was significantly related to different environmental parameters. Among the most important environmental variable, there were temperature, inorganic nutrients and Chlorophyll a (Chla). Yet, together, biogeochemical variables explained only 24% of the variation in bacterioplankton composition. Moreover, this seasonality presented more variability in between March and June, linked with the set-up of spring blooms (Chla parameter).

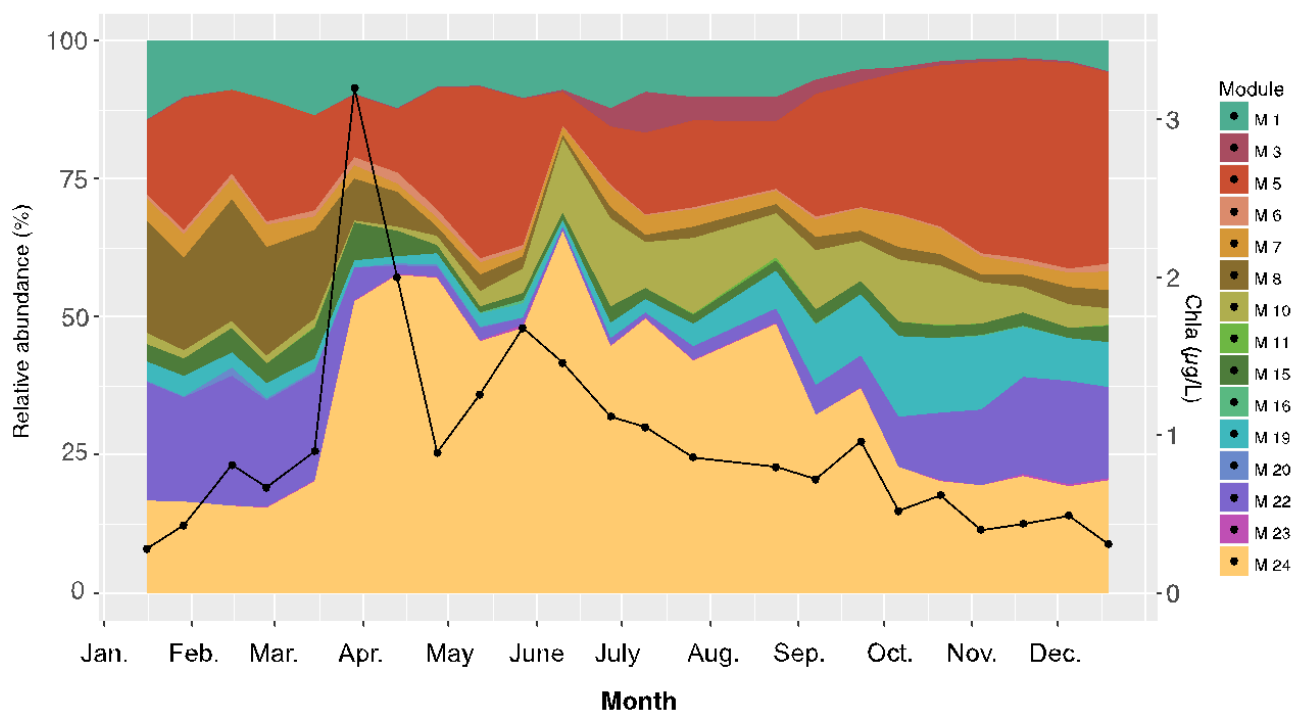


Figure 14 | Variation of the different modules over one year (2016). Cumulative curves of the different modules of co-varying OTUs defined for the year 2016. Modules with more than 1 OTU are plotted. The variation of chlorophyll a is also represented.

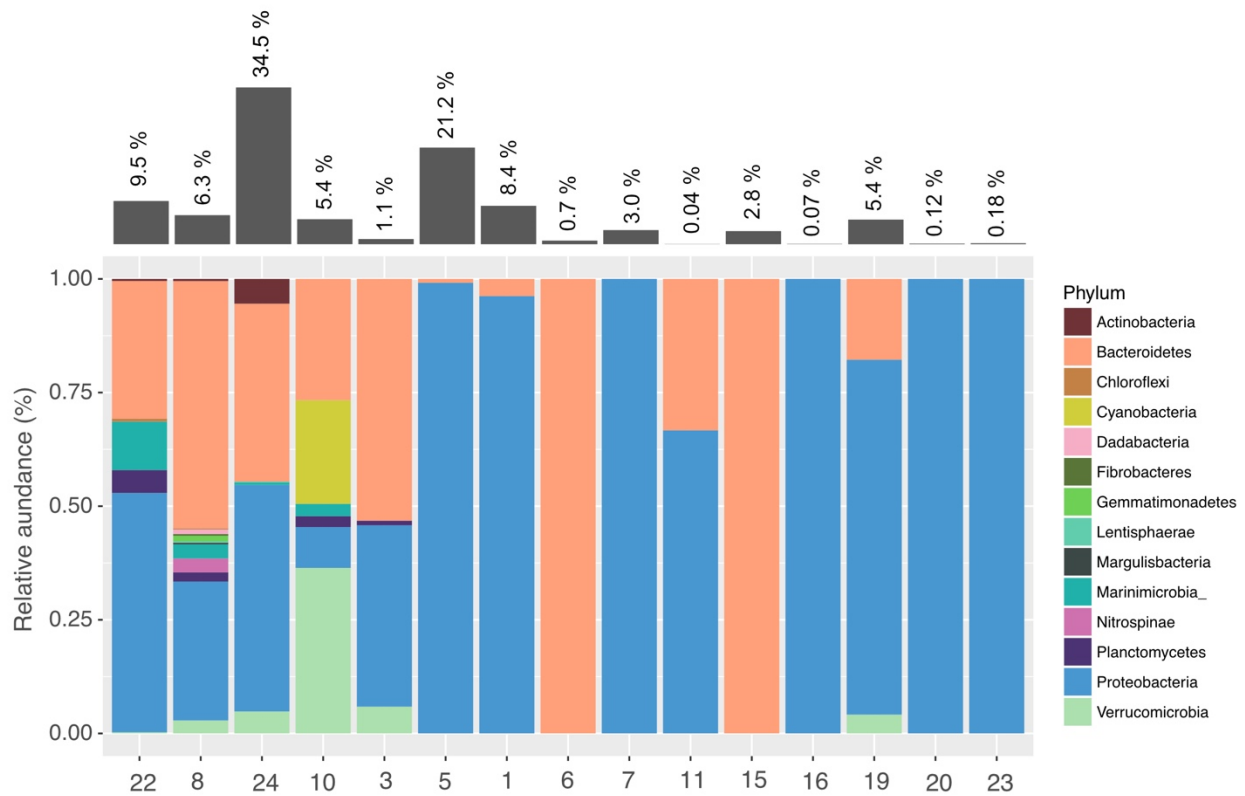
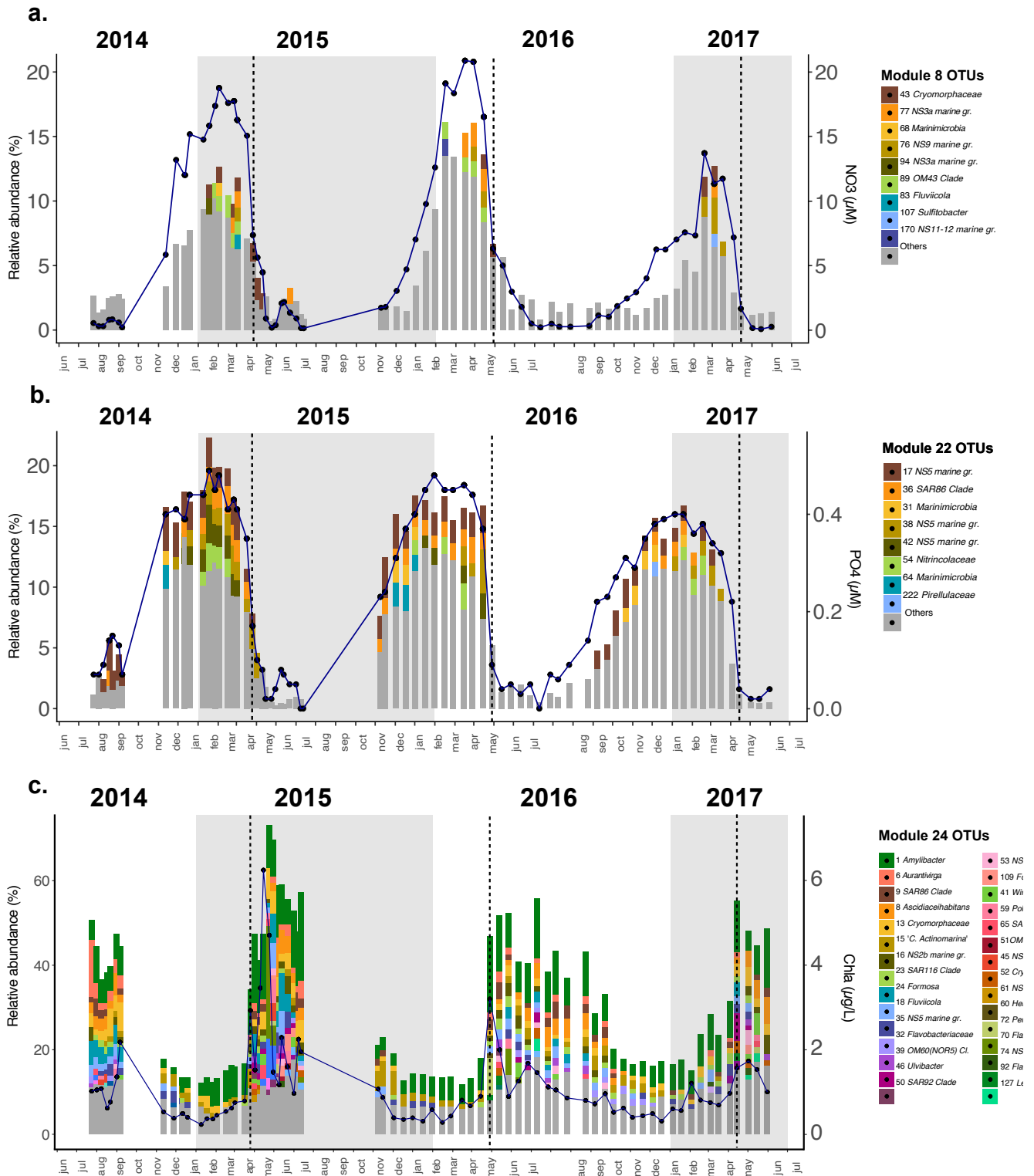


Figure 15 | Taxonomy composition of the different modules. Phylums affiliation within different modules and the relative proportion of each module in the dataset (in % of sequences).

To get more insight into the bacterioplankton community composition specific of the different seasons we used a network analysis on a complete year, from January to December 2016. Based on this network, we defined groups of highly co-varying OTUs that are likely found in the same environmental conditions and thus, likely share the same ecological niche. We obtained 15 modules with more than one OTU that presented temporal dynamic across the year (Figure 14). Furthermore, we looked at the dynamic of the main modules across the entire dataset in order to see if these seasonal patterns defined in 2016 were likely repeatable from one year to another (Figure 16), yet it is not a robust information on the seasonality of each OTU independently. In our study, a complete analysis of seasonality was complex to investigate due to large gaps in the dataset that range from 2 to 3 months in 2014 and 2015 respectively (see material and methods section).

CHAPTER II: TEMPORAL DYNAMIC OF BACTERIAL COMMUNITIES IN THE BAY OF BREST



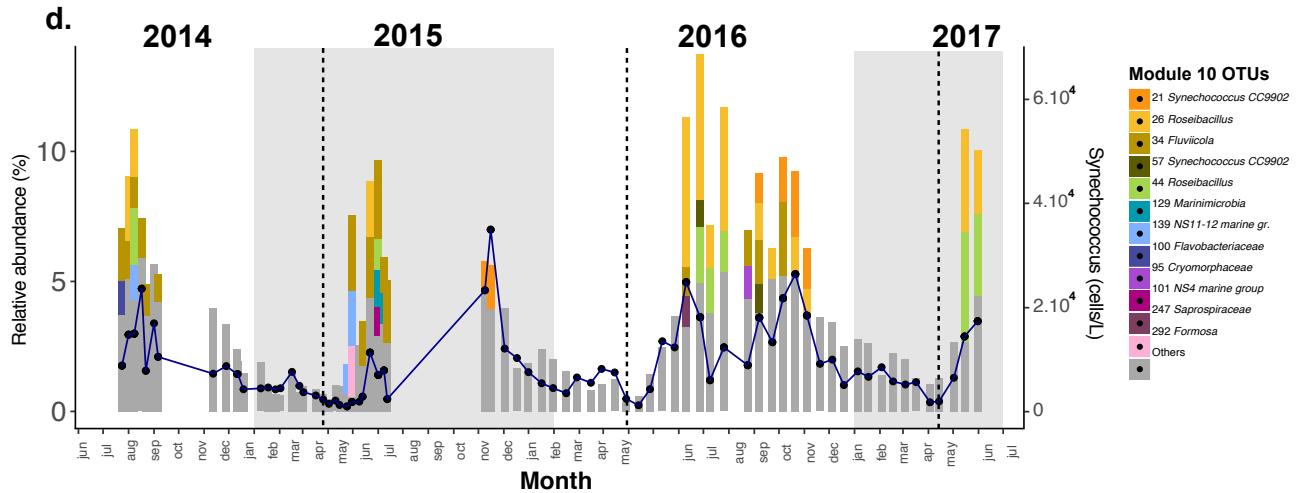


Figure 16 | Variation of the dominant OTUs of the main modules across the entire survey. Main OTUs (>1% in relative abundance in a sample) of modules with a great seasonality are plotted throughout all the survey. The environmental variable that correlated the most to their eigenvector are plotted: nitrates for module 8 (a.), phosphates for module 22 (b.), chlorophyll a for module 24 (c.) and *Synechococcus* cytometric counts for module 10 (d.). The dashed lines correspond to the first peak of chlorophyll a for each year.

Two modules (1 and 5) were abundant across all the year (Figure 14). Their taxonomy shows that they were composed of highly abundant *Proteobacteria*, but of different families: in module 1, OTUs were affiliated to *Rhodobacteraceae* with *Planktomarina* as the dominant genus (60% of the module's sequences), while most OTUs of module 5 were affiliated to *SAR11 Clade* (75% of the module) and particularly group Ia.

On the contrary, 4 modules presented a clear seasonal dynamic (Figure 14). Two modules (8 and 22) were present in winter, between December and March, and correlated strongly with numerous parameters that correspond to winter conditions: inorganic nutrients, and cold temperatures (Figure S 10). Module 8 was negatively correlated with salinity and presented a large diversity of different phyla, with OTUs affiliated to *Marinimicrobia*, *Planctomycetes*, *Nitrospinae*, *Gammatimonadetes* or *Dadabacteria* in addition to the dominant *Proteobacteria* and *Bacteroidetes* phyla (Figure 15). Its

dominant OTUs were affiliated to *Cryomorphaceae*, *NS3a marine group* and *Marinimicrobia* (Figure 16.A). Module 22 was correlated to inorganic nutrients and presented a lower diversity of *Bacteroidetes* and more *Proteobacteria* as well as *Marinimicrobia* and *Planctomycetes* (Figure 15). Its main OTUs were affiliated to *NS5 marine group*, *SAR86 Clade*, *Marinimicrobia* and *Nitrincolaceae* (Figure 16.B). OTUs affiliated to these two modules presented a very recurrent dynamic across the different years. They followed the seasonality of inorganic nutrients that peak in winter and drastically fall with the onset of the first phytoplankton bloom in a short period of time.

On contrary, one module (module 24) became highly dominant right after the first bloom in spring and remained abundant during the whole photoproduction period in 2016 (until the end of September). It correlated positively with Chla and with the main groups of phytoplankton. It is dominated by *Bacteroidetes* and *Proteobacteria* phyla, and in a less extent, of *Actinobacteria* (Figure 15). Patterns of this modules' OTUs among the whole time series showed that they increased sharply in response to the first spring bloom of each year (Figure 16.C). Specific dynamic of the main OTUs of this module is presented in Figure 17 and reveal two main rough patterns. In one hand, some OTUs seemed to be recurrently enhanced after each spring phytoplankton bloom of the survey and presented peaks in relative abundance during the whole period of productivity (Figure 17.A). On the other hand, some OTUs presented a more specific dynamic, being sharply enhanced at one time point during the blooms (Figure 17.B).

Finally, one module was more present in late summer. Module 10 stayed until December and grouped abundant members of *Cyanobacteria* (*Synechococcus* CC9902) and *Verrucomicrobia* among with *Bacteroidetes* and *Proteobacteria* (Figure 15). The main OTUs of this module were affiliated to *Synechococcus* CC9902, *Roseibacillus* and *Fluviicola* (Figure 16.D).

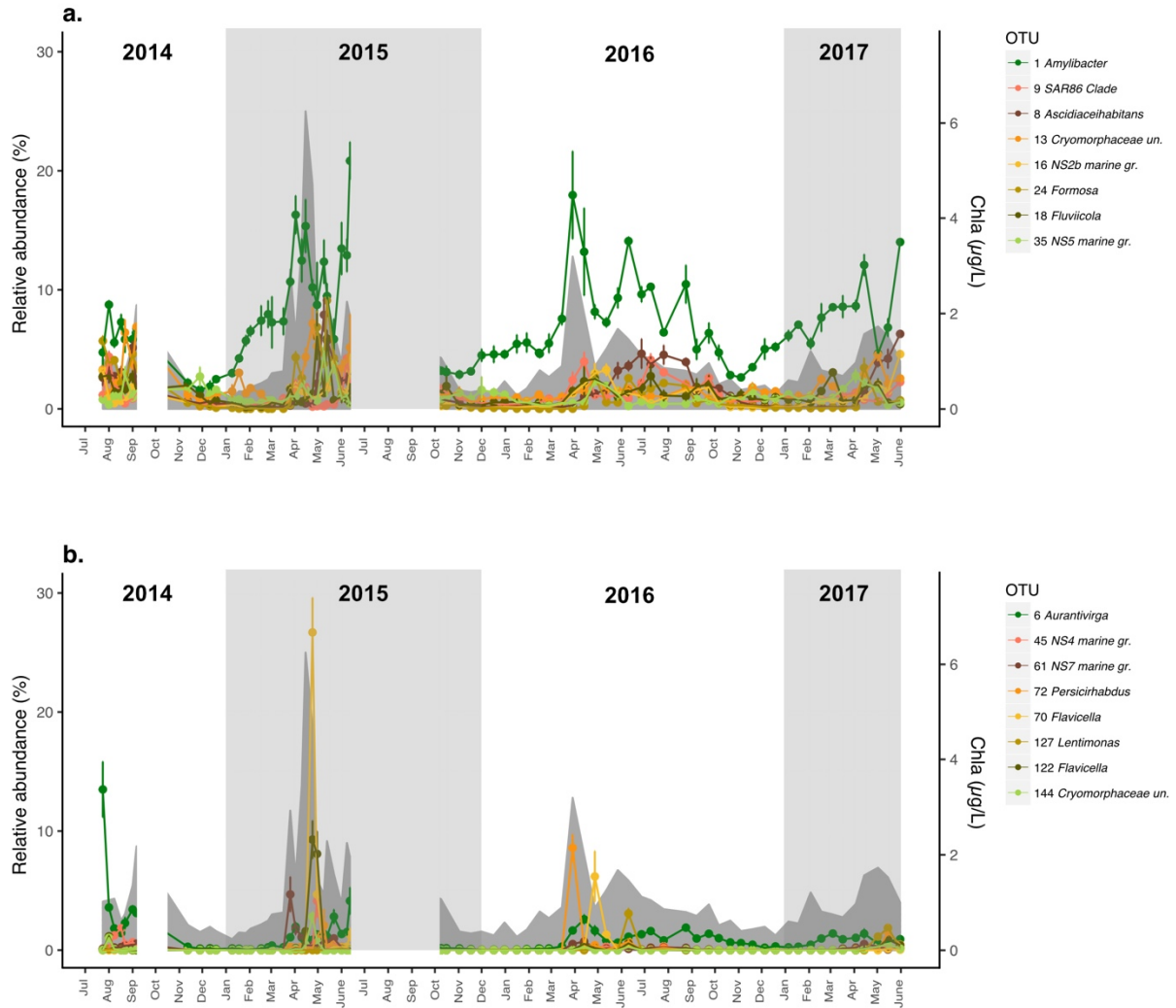


Figure 17 | Variation of OTUs from the spring module (module 24). A. OTUs with recurrent variations after each spring blooms. B. OTUs with highly variable and intense variations in response to the bloom.

6. DISCUSSION

6.1. SEASONAL DYNAMICS OF BACTERIAL COMMUNITIES

The Bay of Brest is a macrotidal coastal system submitted to local forcing that are known to affect the biology and physicochemical characteristics of this environment. Exceptional rain rates lead to winter river discharge anomalies that dramatically impact both salinity and inorganic nutrients concentrations (Tréguer et al., 2014). These winter nutrients loads in addition with tidally and wind-induced mixing of the water column are an important control of the onset of phytoplankton spring blooms (Ragueneau et al., 1996). These variations in salinity, nutrient concentrations and bloom intensities were observed during our sampling period and lead to notable changes between years. Nevertheless, free-living bacterial community composition in the Bay of Brest showed marked and repeatable seasonal variations (CCA, Figure 13). Temporal dynamic observed at seasonal scale are widely found in numerous marine environments and are supposed to remain remarkably stable years after years (Fuhrman et al., 2015; Gilbert et al., 2012). We used network analysis over one year to see patterns of covarying OTUs that presented different seasonal dynamic. These patterns defined in 2016 were particularly coherent from one year to another, comforting the finding that seasonality of microbial communities rely on recurrent assemblages (Chafee et al., 2018; Ward et al., 2017).

6.2. BREAK OF SEASONALITY: SPRING BLOOMS

Seasonality of bacterioplankton community composition shift drastically with the onset of spring blooms (Figure 13) as it was previously reported in other rich coastal environments and mesocosms experiments (Luria et al., 2017; Pinhassi et al., 2004; Riemann et al., 2000; Teeling et al., 2012). Each spring bloom, irrespectively of its intensity, lead to a rapid change between winter communities (module 8 and 22) and spring communities (module 24, Figure 14 and Figure 16). These sharp changes were driven by a rapid increase in relative abundance of OTUs in spring module. There were mainly affiliated to *Flavobacteriaceae* and *Rhodobacteraceae* families, which possess a wide range of hydrolytic enzymes

and membrane transporters to degrade and assimilate organic matter, mostly saccharides, released by the phytoplankton (Pinhassi et al., 2004; Riemann et al., 2000; Teeling et al., 2012). Some of them were specifically enhanced at one time point in less than a week, up to 28% of the whole community. The most abundant are affiliated to *Flavicella* and *Persicirhabdus* members of the *Flavobacteriaceae* and *Verrucomicrobiaceae* families respectively. Members of these families are thought to be highly specialized in the acquisition of complex molecules of organic matter (Buchan et al., 2014; Martinez-Garcia et al., 2012). They were not present every year, suggesting that local conditions are necessary for their development such as phytoplankton diversity of the bloom, that determine the type of organic matter present in the environment (Biddanda and Benner, 1997; Zhang et al., 2015b). Interestingly, the intensity of these taxa response seems to be proportional to the bloom intensity itself, indicating that substrate quantity could be important in their development (Hellweger, 2018).

Other taxa presented recurrent pattern after each spring bloom observed in the survey and last during the whole photoproduction period in 2016. In temperate coastal waters, the first spring blooms, which are the most important in Chla concentrations (Figure 12) are followed by successive blooms of medium intensities that succeed until the end of summer (Beucher et al., 2004). The dominant OTU (OTU1) represent the most successful Bacteria of the dataset, affiliated to *Amylibacter* (formerly *Roseobacter* NAC11-7 lineage) of the *Rhodobacteraceae* family. It reached up to 15-20% of the sequences under bloom conditions but remained relatively abundant, around 2-5%, throughout the year. This genus was previously found to be part of the dominant bacteria of the North sea coastal waters and to increase in response to diatoms substrates (Chafee et al., 2018; Giebel et al., 2011; Taylor and Cunliffe, 2017). Its year-around high relative abundance could be linked to a highly versatile genome that allow this taxon to exploit a large variety of resources (Hahnke et al., 2013; Moran et al., 2007), similarly to another *Rhodobacteraceae*, *Planktomarina* (Voget et al., 2015) that was also abundant permanently in our temporal survey (Sup. Figure S3). The other OTUs are affiliated to taxa previously found associated with phytoplankton such as *SAR86 Clade*, *Cryomorphaceae*, *Formosa* or *Asciadiaceihabitans* of the former *Roseobacter* OCT lineage (Choi et al., 2015; Hahnke et al., 2013; Lucas et al., 2015). Teeling et al., 2016, hypothesized that recurrent patterns of certain taxa following the spring blooms every year suggests

that they could target substrates that are commonly released by the phytoplankton. Free-living members of *Formosa* for instance are very specialized in the acquisition of laminarin, a key substrate of diatoms (Unfried et al., 2018) and phytoplankton communities in the Bay of Brest are largely dominated by Diatoms (Del Amo et al., 1997). Overall, the complex and variable dynamic of these taxa linked to the phytoplankton developments could be due to fluctuation in their resources availability and top-down controls such as grazing and viral lysis (Needham et al., 2017).

6.3. LATE SUMMER COMMUNITIES

After the spring blooms, a secondary bloom period generally took place between June and October, characterized by weaker peaks of Chla (Figure 12). Inorganic nutrients are depleted in the water column and the river flow are their lowest during this period. Phytoplankton growth thus mainly rely on an intense remineralization by both zooplankton and bacteria (Del Amo et al., 1997). A few taxa were specifically abundant during this period, such as *Cyanobacteria*. These *Cyanobacteria*, all affiliated to *Synechococcus* CC9902 were present between September and November in 2016 (Figure 16) which was rather in accordance with cytometry counts data (sup Figure 1b). The presence of *Cyanobacteria* in the end of summer is typical of temperate coastal environment, due to nutrient-depleted and stable conditions that favor small phytoplankton (Li et al., 2006). However, their abundance was rather limited in our study site, with a maximum of 2.6% in October 2016, this low abundance is probably due to a constant mixing of the water column with the tide currents, as these cyanobacteria were more abundant in the upper stratified water masses (Lemonnier et al., Chap.I) .

6.4. MULTIPLE POSSIBLE DRIVERS IN WINTER.

Contrary to the spring community that is likely mainly driven by phytoplankton blooms, environmental parameters shaping winter communities are less evident. Interestingly, the OTUs found in winter separated into two modules that presented a slightly different dynamic, module 22 was already abundant in the beginning of October, while module 8 was present since January. The former strongly correlated with phosphates which is tightly coupled with phytoplankton growth in this environment as

being one of their main limiting nutrient (Del Amo et al., 1997). This suggests that bacteria of this module could be more adapted in environment with a low primary production. It indeed presented dominant members of *Marinimicrobia*, that are generally described in sub-euphotic zone (Agogu e et al., 2011). But it was dominated by abundant heterotrophic taxa such as *SAR86 Clade* or *NS5 marine group*. Other drivers could explain their dynamic such as annual variations in temperatures that define different ecological niches (Eren et al., 2013a; Lucas et al., 2015). Module 8 was strongly negatively correlated to salinity and positively to suspended material. It could gather bacteria particularly influenced by river outflow and storm conditions that are maximized in January and February. During winter at the SOMLIT Station, up to 50% of the organic matter may have a terrestrial origin (Li enart et al., 2017). Little is known on the influence of such type of organic matter on marine bacteria, but some studies showed that terrestrial OM can stimulate numerous typical members of the dominant heterotrophic marine taxa, such as *Bacteroidetes* or *Alphaproteobacteria* (Lindh et al., 2015a). *Bacteroidetes* were particularly abundant in the module 8, with members of *NS9 marine group*, *NS3a marine group* and uncultured *Cryomorphaceae*. In addition, some typical phylum described in marine sediments *Dadabacteria*, *Gemmatimonadetes* with the *BD2-11 terrestrial group* (Hanada and Sekiguchi, 2014; Hug et al., 2016) or non-marine taxa such as *Enterobacteraceae* or *Rhodanobacteraceae* (Naushad et al., 2015) represented a low fraction of this module. This suggest that both river outflows and mixing during winter storms can bring exogenous taxa in the water column, but of minor influence in the global community. Finally, this module gathered all the nitrite-oxidizing bacteria of the Nitrospinae family. These bacteria, as well as the ammonium-oxidizing archaea or bacteria are supposed to be outcompete by the phytoplankton for the nitrogen uptake in summer (Zakem et al., 2018) or limited by the light (Lomas and Lipschultz, 2006) that could explain why they are more abundant in winter.

7. CONCLUSION

Here we presented the seasonal variations of free-living bacterial communities in the Bay of Brest. As in other marine environments, we found coherent and recurrent patterns of seasonality over the period studied. Numerous environmental variables could explain their dynamic, among which the most evident were the phytoplankton developments that lead to drastic changes in bacterioplankton community composition. In winter, multiple parameters such as temperature, high inorganic nutrients, low primary production or the influence of terrestrial input could be important drivers of the bacterioplankton communities. The full comprehension of the dynamic of these communities and likely their role will need further investigations of both poorly characterized taxa as members of *Bacteroidetes* (*NS5 marine group*, *NS3a marine group* and *NS9 marine group*) that were important in our survey and the sampling of other parameters possibly critical in the bacterioplankton dynamic such as organic matter quantity and quality. Finally, the addition of other complete years of sampling would be critical to confirm those first observations.

8. SUPPLEMENTARY DATA

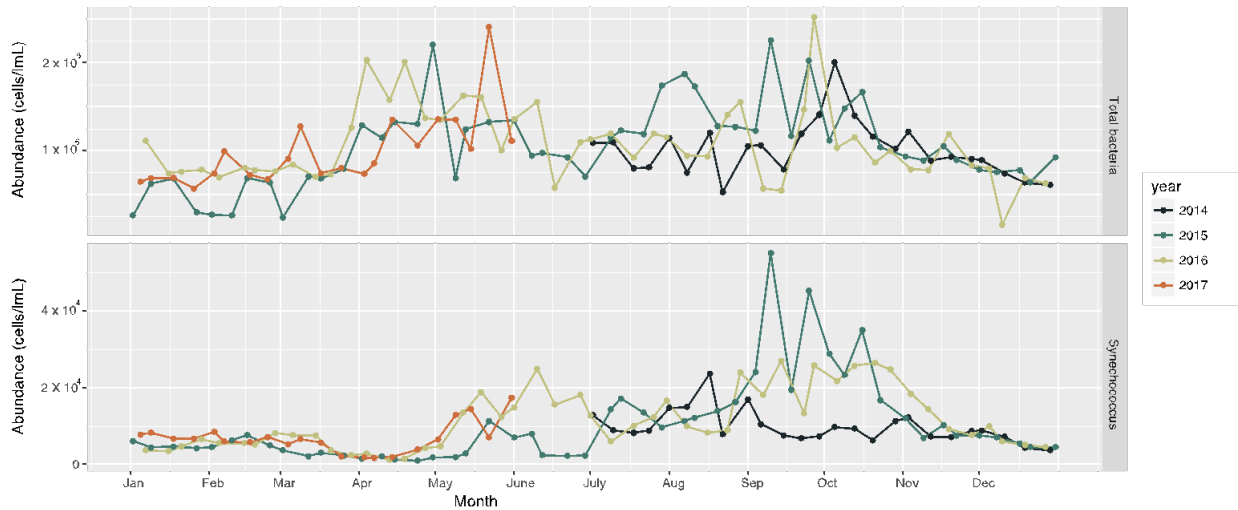


Figure S 7 | Cytometry counts for total bacteria and *Synechococcus* at SOMLIT station in the Bay of Brest, between 2014 and 2018.

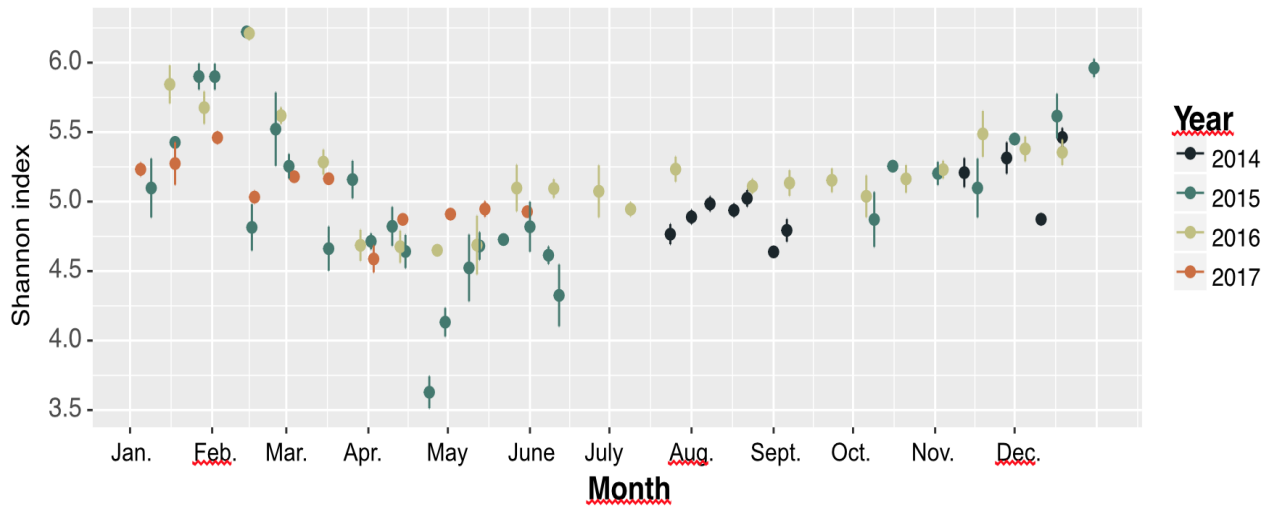


Figure S 8 | Variation in alpha diversity of the bacterial communities measured with Shannon index. Each dot represent the mean of Shannon value obtained for the replicates of a time-point, with their standard deviation.

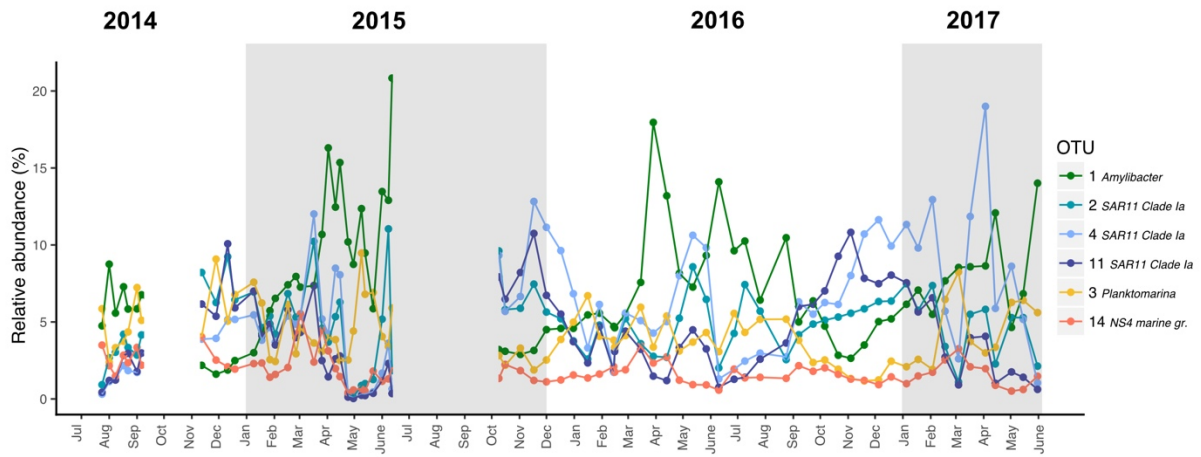


Figure S 9 | Relative abundance of the most abundant OTUs; *i.e.* with a relative abundance upper than 1% in more than 75% of all the samples.

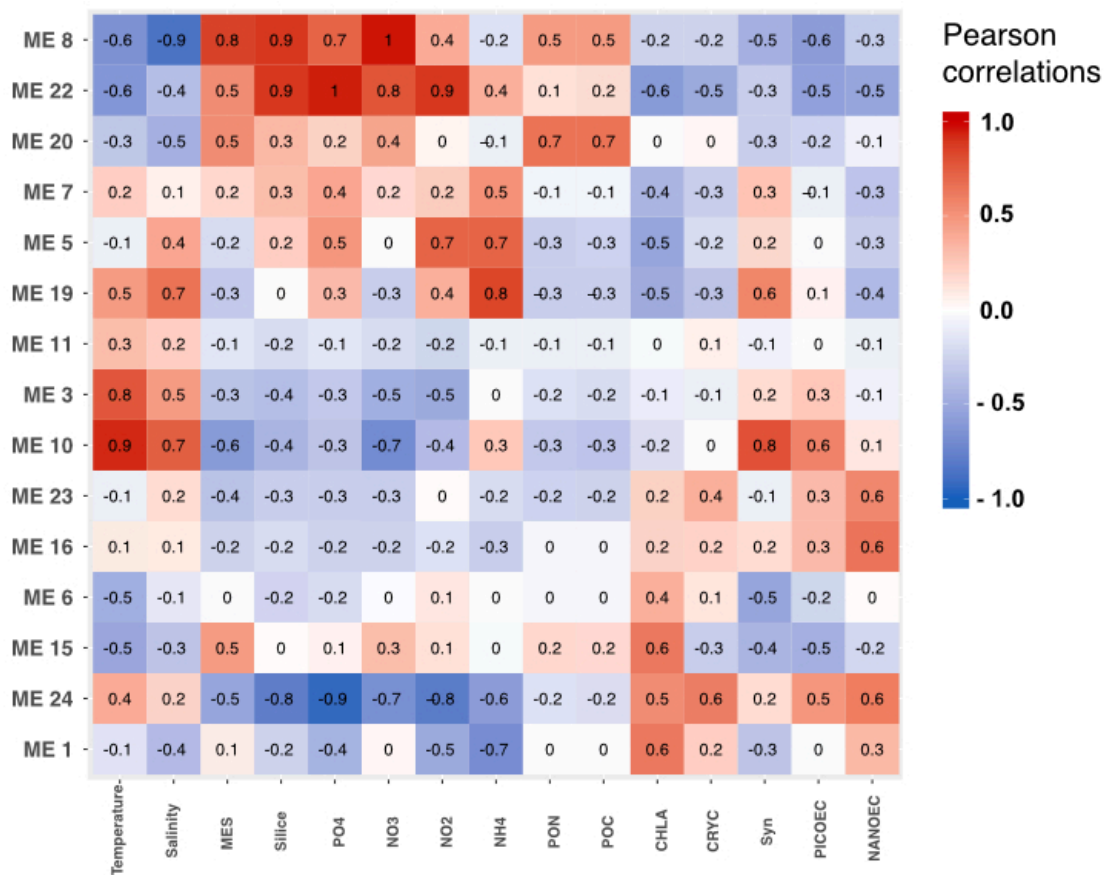


Figure S 10 | Pearson correlation of each module eigenvector with the environmental variables.

Chapter III.

Into the genome of
an abundant
Rhodobacteraceae

1. RÉSUMÉ EN FRANÇAIS

Bien qu'essentielles pour comprendre la dynamique des communautés microbiennes et leur lien étroit avec leur environnement, les approches de séquençage de l'ADN 16S utilisées dans les deux chapitres précédents sont limitées par nos connaissances de ces diversités et les limites de résolution d'un gène marqueur pour appréhender leur rôle et potentiel fonctionnel. Il est intéressant de compléter ces données par des analyses de métagénomique, permettant notamment de reconstruire des génomes, donnant une information précieuse sur les potentiels métabolismes et stratégies écologiques des bactéries présentes dans l'environnement. Pour cela, nous avons préparé 48 banques métagénomiques à partir d'échantillons des deux suivis, en mer d'Iroise et en Rade de Brest. A partir de ces métagénomes, nous avons pu reconstruire 43 génomes issus d'assemblages métagénomique (MAG) de très grande qualité. Ce troisième chapitre décrit un MAG affilié à la famille des *Rhodobacteraceae*, acteurs majeurs dans l'environnement marin. Ce MAG est particulièrement exceptionnel sur de nombreux points. Tout d'abord de par la taille de son génome (1,8Mb) et la proportion de GC dans son génome (37%) tout deux les plus faibles valeurs reportées pour une *Rhodobacteraceae*. Cette taille, ainsi que les gènes retrouvés dans son génome indiquent une adaptation suivant la théorie de la réduction de génome chez les bactéries pélagiques (streamlined genome). Mais contrairement à une vision classique des génomes réduits, les données d'abondance de ce génome indiquent qu'il est particulièrement compétitif en période de blooms. Enfin, il correspond probablement au génome de l'OTU le plus abondant retrouvé dans nos deux suivis en Rade de Brest et Mer d'Iroise, affilié à *Amylibacter sp.* Ces résultats offrent de nouvelles perspectives sur les adaptations des bactéries des eaux côtières tempérées.

Metagenome Assembled Genome of an abundant *Rhodobacteraceae* in the Bay of Brest and the Iroise Sea

C. Lemonnier^{1,2}, Johanne Aube¹, M. Perennou², H. Morrison³, C. Paillard² and L. Maignien^{1,3}

¹ Univ Brest, CNRS, IFREMER, Laboratoire de Microbiologie des Environnements Extrêmes, F-29280 Plouzané, France

² Univ Brest, CNRS, IRD, IFREMER, Laboratoire des Sciences de l'Environnement Marin, F-29280 Plouzané, France

³ Josephine Bay Paul Center for Comparative Molecular Biology and Evolution, Marine Biological Laboratory, Woods Hole, USA

Draft paper, additional results required before submission to publication

2. ABSTRACT

While being essential in the understanding of microbial communities dynamic and their close relationship to the environment, 16S DNA sequencing approaches used in the previous two chapters are limited by our knowledge of these diversities and the limits of resolution of a marker gene to understand their role and functional potential. It is interesting to supplement these data with metagenomic analyses, in particular to reconstruct genomes, giving valuable information on the potential metabolism and ecological strategies of bacteria present in the environment. For this purpose, we prepared 48 metagenomic banks from samples of the two surveys, in the Iroise Sea and in the Bay of Brest. From these metagenomes, we were able to reconstruct 43 very high quality metagenomic assemblies (MAGs). This third chapter describes a MAG affiliated with the family *Rhodobacteraceae*, widespread bacteria of marine environments. This MAG was particularly exceptional in many respects. First of all, the size of its genome (1.8Mb) and the proportion of GC in its genome (37%) are the two lowest values reported for a *Rhodobacteraceae*. Its size, as well as its gene content indicated an adaptation according to the streamlined genome theory in pelagic bacteria. But unlike a conventional view of reduced genomes, abundance data for this genome indicated that it was particularly competitive during bloom periods. Finally, it probably corresponds to the most abundant OTU genome found in our two surveys in the Bay of Brest and the Iroise Sea, affiliated with *Amylibacter sp.*

3. INTRODUCTION

Marine *Rhodobacteraceae* or *Roseobacter group* (Simon et al., 2017) is one of the major clades of heterotrophic bacteria inhabiting oceanic surface waters (Moran et al., 2007), grouping more than 70 genera and 170 species (Pujalte et al., 2014). Members of this clade represent ecological generalists, able to thrive within a wide variety of habitats including associations with different organisms as phytoplankton, corals and animals (Aprill et al., 2009; Grossart et al., 2005; Rao et al., 2007). Their capacity to exploit such different environments rely on large and versatile genomes giving them very diverse metabolic capacities (Moran et al., 2007; Polz et al., 2006).

However, numerous studies shows that the predominance of versatile capacities among *Rhodobacteraceae* could be culture biased. Metagenomics and Single cells genome sequencing approaches revealed that natural populations of this group could be more adapted to a free-living lifestyle with streamline genomes (Luo et al., 2014; Swan et al., 2013). This genome streamlining could have emerged in the evolution of a basal lineage, represented by *Rhodobacterales* bacterium HTCC2255 (Luo et al., 2013). Evolutionary studies showed that this bacteria likely encompass a gene reduction during its evolution, as other streamlined *Alphaproteobacteria* such as *SAR11 Clade* (Luo et al., 2013). Their genome content and life strategy suggest that they can play distinct roles in marine biogeochemical cycles than the other *Rhodobacteraceae* (Luo and Moran, 2014). But our knowledge on these streamlined *Rhodobacteraceae* is still scarce with few cultivated members and genomes completely reconstructed (Zhang et al., 2016). Moreover, most analysis of these environmental genomes relied on large spatial sampling over different oceans, giving few information on their dynamics and potential ecological niches in coastal environments.

In this study, we did metagenomic analysis on both a spatio-temporal and temporal survey in coastal water of the North-East Atlantic Ocean (Brittany, France) in order to reconstruct different genomes (MAGs) and get insight the genome characteristics of the abundant members of *Rhodobacteraceae* present in the Bay of Brest and the Iroise Sea.

4. MATERIAL AND METHODS

4.1. STUDY SITE AND SAMPLING

The 48 metagenomes were retrieved in the North-East Atlantic ocean, from a spatio-temporal survey over one year in the Iroise sea, between September 2014 and September 2015 and a time series at the SOMLIT Station in the Bay of Brest from July 2014 to April 2017. The sampling strategy, filtration of seawater and DNA extractions was the same as previously described in the first two chapter. The samples for metagenome libraries preparation were selected based on their interest regarding 16S sequencing results for the Iroise Sea (e.g. samples dominated by a particular module) and every month for the temporal survey in the Bay of Brest (Figure 18). Metagenomes samples were prepared from the same extracted DNA used for the 16S amplicon sequencing.

4.2. LIBRARIES PREPARATION FOR METAGENOMES

Metagenomes libraries were prepared following the Illumina kit TruSeq Nano DNA library prep®. DNA fragmentation was done using a Covaris® with an optimized time of 91s. Sequencing of libraries was done on a, Illumina NextSeq sequencer at the Josephine Bay Paul Center (Woods Hole, USA).

4.3. RECONSTRUCTION OF METAGENOME ASSEMBLED GENOMES (MAGs)

Our bioinformatic pipeline followed in large part the snakemake (Köster and Rahmann, 2012) proposed in Anvi'o v5.2 (Eren et al., 2015) as well as previous pipeline of MAGs reconstruction done in the Tara ocean dataset (Delmont et al., 2018 ; Tully et al., 2018):

4.3.1. SEQUENCES QUALITY FILTERING AND ASSEMBLY

Raw-reads were quality filtered based on Minoche et al., 2011, recommendations using Illumina-Utils (Eren et al., 2013c). High-quality reads were assembled using IDBA-UD v.1.1.3 (Peng et al., 2012) in each of the 48 samples to reduce computational limits. Contigs with a length < 1000 bp were discarded

for the downstream analysis. We then co-assembled those contigs in 5 groups of co-assembly (ID01, ID02, ID03, ID04 and negative control ; Figure S 11) defined based on k-mer similarities using Simka algorithm (Benoit et al., 2016). We performed a secondary assembly of the contigs within each groups of co-assembly. First, contigs with an identity of 99% or more were grouped together using CD-HIT-EST v4.6.8 (Li and Godzik, 2006) in order to reduce computational time. These contigs were then assembled using Minimus2 (Sommer et al., 2007) to form secondary contigs. Profiling of secondary contigs and the primary contigs that did not assemble with Minimus2 followed the Anvi'o method. Mapping of high quality reads on contigs was done using Bowtie2 v2.3.4.3 (Langmead and Salzberg, 2012) to assess their distribution among the samples. Putative genes present on contigs were detected using Prodigal v2.6.3 (Hyatt et al., 2010), and the single copy core genes of Bacteria (Campbell et al., 2013), Archaea (Rinke et al., 2013) and Eukarya (Simão et al., 2015) were retrieved with HMMER v3.2.1 (Mistry et al., 2013). These single copy core genes were then used to define genome completion and redundancy. Finally, taxonomy affiliation of the different contigs was done with Centrifuge v1.0.4 (Kim et al., 2016). All statistics of these first steps are summarized in Table S 2.

4.3.2. BINNING AND MAG CHARACTERIZATION

Contigs were automatically binned within each co-assembly group using CONCOCT (Alneberg et al., 2014). To avoid fragmentation error, we ran CONCOCT with a maximum number of clusters (ID01 : 146 , ID02 : 333, ID03 : 349, ID04 : 413 and negative control : 2) of about half of the expected number of microbial genomes based on single copy core gene profiles. Each bin was then manually curated using Anvi'o interface and the `anvi-refine-bin` command. Bins of good quality were considered according to these criteria : completion of single copy core genes above 70% or of a size > 2 Mb and a redundancy under 10% (Delmont et al., 2018). These bins were then grouped within a new database and redundant MAGs reconstructed in the different co-assemblies were groups together with Drep v2.2.4 (Olm et al., 2017) if they presented an average nucleotide identity (ANI) score of more than 98% and a minimum alignment of 75%. A complete profiling of the non-redundant MAGs was then done to assess their mean coverage and detection statistics within each of the 48 metagenomes. Functional assignation of

the genes within each MAG was performed with the NCBI COG database (Galperin et al., 2015) using `anvi-run-ncbi-cog`. We also used the KEGG database with GhostKOALA (Kanehisa et al., 2016). Taxonomy affiliation of the different MAGs was done using CheckM v1.0.13 (Parks et al., 2015).

4.3.3. PHYLOGENOMIC

Phylogenomic analysis of MAG ID03_00003 was conducted with 92 sequenced *Rhodobacteraceae* genomes retrieved from NCBI. Using Anvi'o, 114 single-copy genes from the bacterial collection were concatenated and aligned with MUSCLE v3.8.1551 (Edgar, 2004). Phylogenetic trees were created with MegaX with maximum likelihood method and using a JTT model ((Kumar et al., 2018)). Bootstrapping was executed using rapid bootstrap analysis algorithm with 100 bootstraps.

1. RESULTS AND DISCUSSION

1.1. RECOVERY OF MAGS AND IDENTIFICATION OF THE DOMINANT MAG AFFILIATED TO THE RHODOBACTERACEAE FAMILY

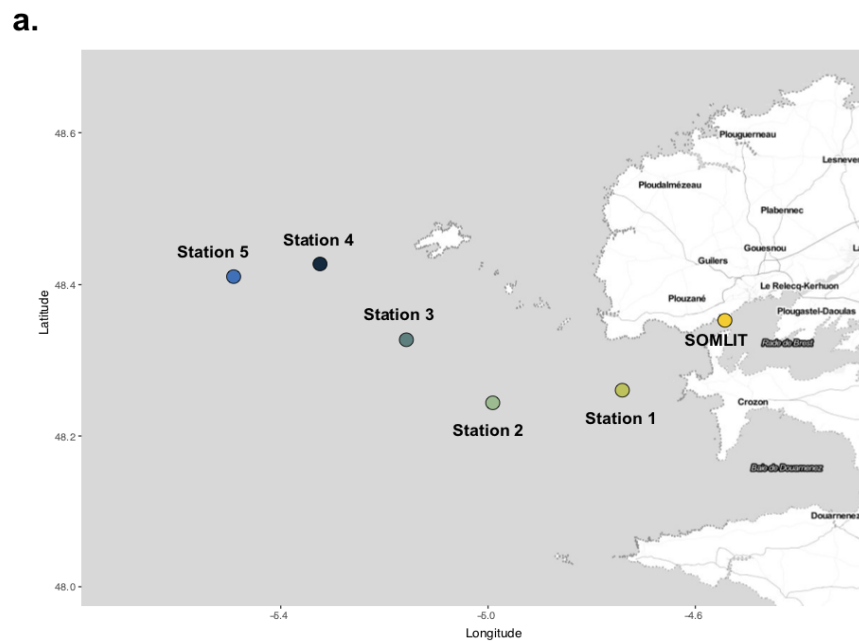
After sequencing we obtained 2 652 926 650 raw reads over the 48 metagenomes, with 2 472 893 193 that passed quality filtering. A total of 1 816 943 contigs with more than 1000 bp length were retrieved from the different co-assembly. We were able to assemble 396 manually curated, good quality genomes. They belonged to the Bacteria (n = 365) and Archaea (n = 31) domains. Bacteria were mostly affiliated to *Proteobacteria* (n = 184) and *Bacteroidetes* (n = 100) and Archaea to *Euryarchaeota* (n = 27). These MAGs represented 8.8% of the initial contigs length. We focused on the most abundant MAG affiliated to *Rhodobacteraceae*, that recruited the highest percentage of reads across the samples (Figure S 12). It was also the most abundant MAG we were able to reconstruct (Figure S 12). This MAG was redundant, indicating that we were able to assemble it from the 4 co-assembly. MAG ID03_00003 had the best quality scores: a completion of 94.24%, a redundancy of 0.72% and it was composed of only 35 contigs (Table S 2). This corresponds to a high quality, likely complete MAG (Bowers et al., 2017).

1.2. GENOMIC CHARACTERISTICS OF MAG ID03_00003

The genome size (1.8 Mb) and GC content (36.5%) of MAG ID03_00003 is typical of a streamlined genome and is smaller than any reported values for marine *Rhodobacteraceae* including uncultivated members. *Rhodobacteraceae* retrieved from Single Cell Genome sequencing in different studies presented a genome size ranging from 2.6 Mb to 3.5 Mb and a GC content from 37% to 40% (Luo et al., 2014; Swan et al., 2013; Zhang et al., 2016). The closest genome to MAG ID03_00003 is the cultivated strain *Rhodobacterales* bacterium HTCC2255, which present as genome size of 2.3 Mb and a GC percent of 38.9%. Genome streamlining is supposed to be an adaptation to nutrient-poor environments for pelagic marine bacteria (Giovannoni et al., 2005b). MAG ID03_00003 indeed presents

a more oligotrophic signature in its genome according to Lauro et al., 2009, criteria : for instance a low number of genes affiliated to COG categories for cell motility, signal transduction and defense mechanisms (Table S 3). Our MAG thus offers an exceptional example of uncultured pelagic marine *Rhodobacteraceae* harboring a streamlined genome and a typical oligotrophic, free-living lifestyle.

Indeed, a *Rhodobacteraceae* is usually considered to have a small genome when its size is around 3 Mb (Durham et al., 2014). This is the case for *Planktomarina temperata* RCA23 isolated in the North Sea, which was also described as presenting a streamlined genome for *Rhodobacteraceae* and an oligotrophic signature with a genome size of 3.29 Mb and GC content of 54% (Voget et al., 2015).



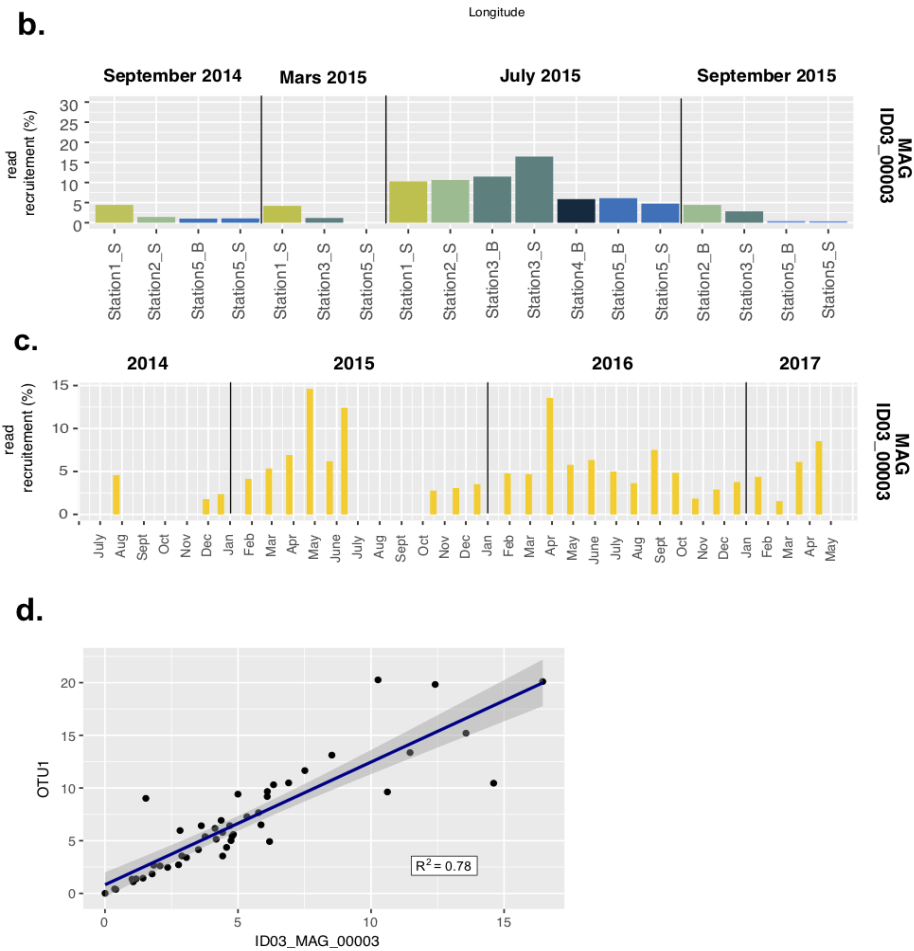


Figure 18 | Sampling points and distribution of MAG ID03_00003 across the samples.

a. Map showing the location of the sampling points. Station 1 to 5 were sampled during 4 cruises in the Iroise sea between September 2014 and September 2015. The SOMLIT Station was sampled between July 2014 and May 2017. b. Distribution of MAG ID03_00003 represented by the percentage of reads it recruited in a sample, across the Iroise sea. For each station, different depth could have been sampled: deep (B) and surface (S). c. Distribution of MAG ID03_00003 at the SOMLIT Station. d. Correlation between ID03_00003 distribution (in % of read recruited by MAG ID03_00003 in a sample) and OTU1 (in % of sequences in a sample), the most abundant *Rhodobacteraceae* OTU detected by 16S sequencing in the same samples, affiliated to *Amylibacter* sp.

Table 2 | Main pathway and genes present in different genomes of *Rhodobacteraceae*. Selection of the different pathways and genes and their presence in *Rhodobacterales* bacterium HTCC2255 and *Planktomarina temperata* RCA23 were given in previous studies (Newton et al., 2010; Voget et al., 2015). Presence of these pathways/gene in MAG ID03_00003 was retrieved based on KEGG and COG functional annotations. *A dehydrogenase was found for Glutamine and Leucine only.

		Rhodobacterales bacterium HTCC2255	ID03_00003	Planktomarina temperata RCA23
Trophic strategy	Photosynthesis cluster			
	rhodopsin			
Aromatic degradation	B-ketoadipate			
	benzoate			
	phenylacetic acid			
	homoprotocatechuate			
Carbon monoxide utilization	Group I CO DH			
	Group II CO DH			
C1 compound utilization	C1 incorp (serine)			
	MeOH oxidation			
	TMA oxidation			
	formaldehyde oxidation			
	formate oxidation			
C2 compound utilization	ethylmalonyl pathway			
	glyoxylate shunt			
Mobility, sensing and attachment	Chemotactic ability			
	motility			
	fliC pilus (type IV)			
	lux quorum sensing			
	virB system (type IV SS)			
Secondary metabolite production	antibiotic production (TDA)			
	NRPS/PKS			
Vitamine related	biotin synthesis			
	biotin uptake			
	cobalamin synthesis			
	Thiamin synthesis			
Iron related	heme uptake			
	siderophore uptake			
	Fe3+ uptake			
	Fe-S synthesis			
Nitrogen related	Carbamate kinase			
	Glutamine synthetase			
	Glu/leu/phenyl/val DH		*	
	Urease			
Phosphorus related	Alkaline phosphatase			
	High affinity Pi transport			
	Low affinity Pi transport			
	Phosphonate utilization			
	polyphos storage			
Sulfure related	DMSP demethyl (dmdA)			
	DMSP lyase (dddL)			
	DMSP cleavage (dddD)			
	DMSP cleavage (dddP)			
	sulfur oxidation (sox)			

1.3. GENE CONTENT AND PUTATIVE METABOLISM

MAG ID03_00003 contained only 1777 coding sequences with 1636 that presented a COG and/or KEGG functional annotation (Table S 3). This is small compared to the *Rhodobacterales* bacterium HTCC2255 (n = 2240) and *Planktomarina temperata* RCA23 (n = 3101) but still higher than typical streamlined bacteria such as *Pelagibacter ubique* HTCC1062 (n = 1354) (Giovannoni et al., 2005b; Luo et al., 2013; Voget et al., 2015). Comparison in presence or absence of key pathway or genes among *Planktomarina temperata* RCA23, *Rhodobacterales* bacterium HTCC2255 and our MAG is presented in Table 2. This figure shows that our MAG and *Rhodobacterales* bacterium HTCC2255 present close characteristics with a reduced number of key pathway and genes. They both lack functions of motility and attachment, confirming a strict pelagic lifestyle. They also both present a rhodopsin coding gene (COG5524), which can harvest energy from light (Giovannoni et al., 2005a).

In addition, MAG ID03_00003 presents abilities to use C1 and C2 compounds. There are five C1 utilization pathway found in marine *Roseobacter* (Newton et al., 2010). MAG ID03_00003 present two of them: Trimethylamine (TMA) and Formate oxidation. These two compounds correspond to common molecules produced in large amount by many organisms in the surface ocean (Lidbury et al., 2014). For instance, TMA comes from the degradation of proteins and other osmolytes such as glycine betaine or choline (Lidbury et al., 2015). The trimethylamine monooxygenase [EC:1.14.13.148] that can oxygenate TMA to TMAO was present in the genome of ID03_00003 but it lack genes to catabolize TMAO and assimilate it into biomass. The genome present the formate dehydrogenase [EC:1.17.1.9] that catalyzes the oxidation of formate into CO₂. This suggest an utilization of C1 compounds for energy metabolism rather than biomass as for some members of *SAR11 Clade* (Sun et al., 2011). Our MAG also present all the genes involved in the glyoxylate shunt, an important pathway for microorganisms to use C2 carbon for their growth and energy (Harder, 1973).

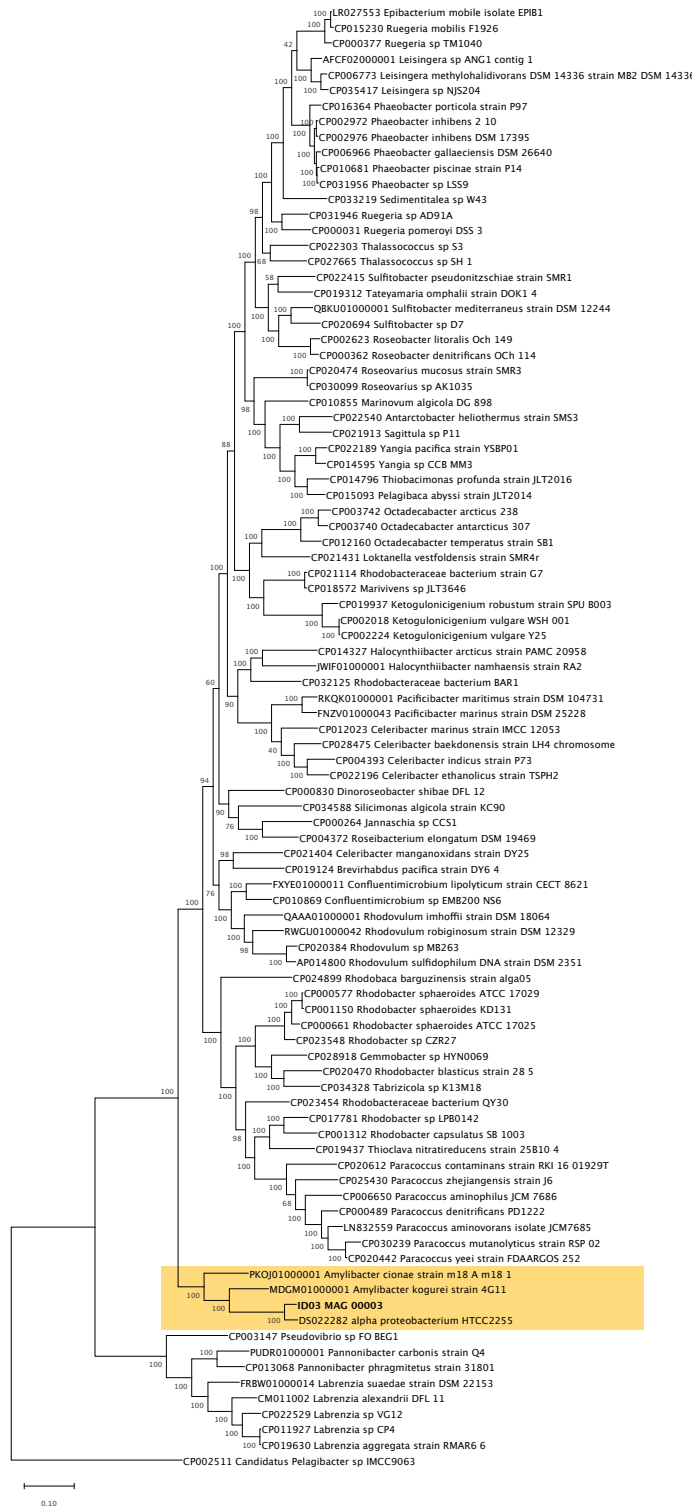
The genome of MAG ID03_00003 encodes 208 transport protein with a large part of the ABC or ABC-like family (n = 130). This correspond to a total of 71.8 of ABC transporter genes per Mb in ID03_00003, which is similar to the proportion found in *Planktomarina temperata* RCA23 and other pelagic

Rhodobacteraceae but significantly higher than the proportion found in *Rhodobacterales* bacterium HTCC2255 (47.6) or *C. Pelagibacter ubique* (51.1). These high-affinity transporters in MAG ID03_00003 included the uptake of amino acids, taurine or spermidine. Other transporters of the Drug/Metabolite superfamily (n = 21) or TRAP-type C4 dicarboxylate transport system were detected in our MAG. The latter transports different dicarboxylic acids such as malate, fumarate or succinate (Forward et al., 1997).

Regarding its nitrogen source, MAG ID03_00003 seems to rely on reduced or organic nitrogen only as it lacks genes encoding enzymes for reducing nitrates and nitrites but possess numerous transporters for ammonium (*amt*) and different nitrogen-rich compounds discussed previously (e.g. taurine, amino acids). This is coherent with other *Rhodobacteraceae* that are unable to use directly inorganic nitrate or nitrite (Voget et al., 2015). Finally, ID03_00003 present the gene coding to the dimethylsulfoniopropionate (DMSP) demethylase (*dmdA*). DMSP is an abundant osmolyte produced by microalgae in the marine environment (van Duyl et al., 1998) that represent an important source of reduced sulfur and carbon for marine bacteria (Yoch, 2002). This seems to be its only source of reduced sulfur as it lack the complete set of gene for assimilatory sulfate reduction (*cysDNCHIJ*).

Altogether these results suggest that MAG ID03_00003 can utilize and assimilate a wide variety of labile common carbon compounds in its environment and is possibly auxotroph for certain molecules such as DMSP. This metabolism is similar to the one found in *SAR11 Clade* and in line with streamlining theory (Giovannoni et al., 2005b; Tripp, 2013). Indeed, streamlined organisms likely occur within niches requiring minimal functional complexity (Giovannoni et al., 2014).

CHAPTER III: INTO THE GENOME OF AN ABUNDANT RHODOBACTERACEAE



a.

b.

	MAG ID03_00003
Rhodobacterales_bacterium_HTCC2255	0.85020596
Amylibacter kogurei 4G11	0.84171623
Amylibacter cionae m18	0.85089076

Figure 19 | Phylogenomic and taxonomic analysis of MAG ID03_00003. a. Maximum likelihood tree inferred from a core gene matrix of 114 single copy core genes of 92 genomes of *Rhodobacteraceae*. Average Nucleotide Identity (ANI) scores between MAG ID03_00003 and other members of the same cluster in the phylogenetic tree.

1.4. TAXONOMIC AFFILIATION AND PHYLOGENY

MAG ID03_00003 thus seems to be an original member of the uncultured *Rhodobacteraceae* with, to the best of our knowledge, the smallest genome, GC content and number of genes in any previously reported genome of *Rhodobacteraceae*. We thus investigated its phylogenomic position within the marine *Rhodobacteraceae* by a phylogenomic analysis with 93 sequenced genomes (Figure 19.A). The tree obtained shows that MAG ID03_00003 is included in the subclade of *Rhodobacterales* bacterium HTCC2255. This subclade was previously referred as NAC11-7 lineage (Zhang et al., 2016) and include two genomes of *Amylibacter*. These results confirm that MAG ID03_00003 represent a close relative of *Rhodobacterales* bacterium HTCC2255 and could explain their many shared characteristics highlighted in our study.

We further used a pairwise comparison of average nucleotide identity between the different members of this subclade in order to see their genome similarity (Figure 19.B). This metric is used to define taxonomic relations between two genomes (Figueras et al., 2014). For instance, two genomes are affiliated to the same specie if they present an ANI above 95%, corresponding to the DNA-DNA hybridization cutoff of 70% (Goris et al., 2007). MAG ID03_00003 present a maximum ANI score of 85.02% with the genome of *Amylibacter cionae* (Figure 19.B). This indicates that this genome belong to a novel specie of the *Rhodobacteraceae*.

1.5. IMPORTANCE OF MAG ID03_00003 IN THE BAY OF BREST AND THE IROISE SEA.

The distribution of MAG ID03_00003 across our spatio-temporal surveys in the Bay of Brest and the Iroise sea is very close to the dominant OTU in our environment (OTU1) found from 16S studies in the same samples (Figure 18.D). This OTU was affiliated to *Amylibacter* genus against Silva v132 database. Regarding their similar distribution and their close phylogenetic affiliation, OTU1 and the MAG ID03_00003 are thus likely to represent the same organism in the environment. Both 16S and metagenomic studies found that this organism is very successful in our environment, more than the dominant *SAR11 Clade* members. This confirms a previous large-scale metagenomic analysis over different oceans that showed that NAC11-7 lineage is highly abundant at the L4 station in the English Channel, a close sampling point to our studied area (Zhang et al., 2016).

Interestingly, ID03_00003 was particularly dominant in April at SOMLIT station and in July during M2BiPAT survey that corresponded to periods of intense bloom. This dynamic is totally in line with numerous 16S ecological studies of the free-living bacterioplankton that found Rhodobacteraceae often associated with phytoplankton bloom (Buchan et al., 2014; Hahnke et al., 2013), including members of NAC11-7 lineage (Teeling et al., 2012).

In temperate coastal waters typically, phytoplankton blooms release large amount of organic matter that is readily available for heterotrophic bacteria. Thus, the dominance of MAG ID03_00003 under these conditions is in contradiction with the common view of genome streamlining as an adaptation to oligotrophic environment (Giovannoni et al., 2005b). This importance under bloom conditions was also reported for *Planktomarina temperata* RCA23 in the North Sea (Voget et al., 2015). In this study, authors showed that it was also very active under these conditions, transcribing simultaneously almost its entire genome. Our result emphasizes the hypothesis that genome streamlining can be an adaptation in dynamic and rich coastal waters (Giovannoni et al., 2014; Yooseph et al., 2010). This characteristic, combined with a great number of high-affinity membrane transporters, obviously gives a great advantage for MAG ID03_00003 to be the most successful organism in our environment.

2. CONCLUSION

In this study we were able to assemble the genome of the most successful *Rhodobacteraceae* present in the Bay of Brest and the Iroise Sea. MAG ID03_00003 represent a new member of the pelagic *Rhodobacteraceae* that was phylogenetically close to *Rhodobacterales* bacterium HTCC2255. This MAG was exceptional within the *Rhodobacteraceae* family with a highly streamlined genome of 1.8 Mb and GC content of 36.9%, the smallest reported to date within this clade. Its streamlined genome, coupled with an adaptive niche for common and labile small substrates and a large number of high-affinity transporters could be explanation of its success all over the year and especially under bloom conditions. This result offers a new comprehension of the metabolisms of successful temperate coastal bacteria.

3. SUPPLEMENTARY DATA

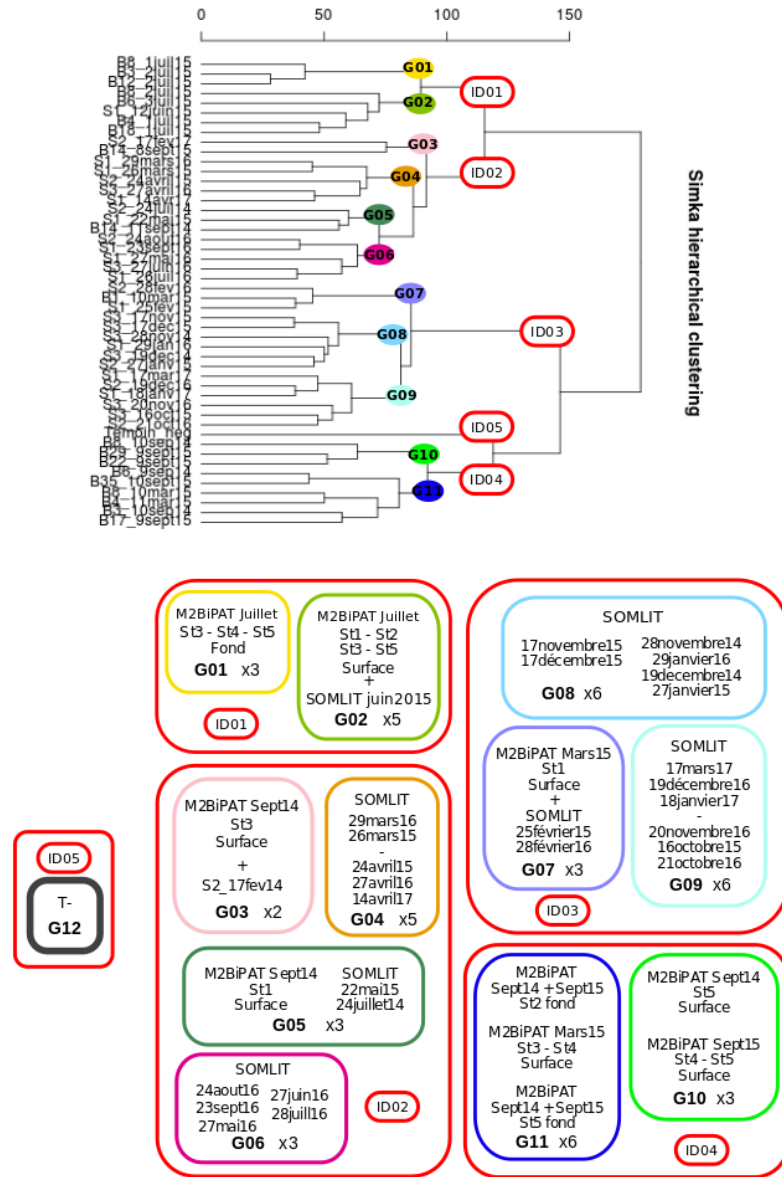


Figure S 11 | Hierarchical clustering with Bray-Curtis distances based on k-mer frequency with simka between samples. Groups of co-assembly used in this study are defined by red circle.

Table S 2 | Statistics of the contigs within each co-assembly group.

	ID01	ID02	ID03	ID04	Neg
Total Length	942596156	2178489610	2230582355	1948950826	13262041
Num. Contigs	223582	516526	586476	487162	3197
Num. Genes (prodigal)	1013737	2362376	2590670	2154041	14582
Longest Contig	496719	595969	555500	867007	307531
Shortest Contig	1000	1000	1000	1000	1099
N50	22722	54892	71472	48689	181
N75	77679	182123	227475	175068	960
N90	149356	346111	408873	331975	2113
L50	7634	7509	5897	6911	9997
L75	2674	2697	2386	2454	2281
L90	1510	1524	1463	1465	1378
Rinke_et_al	29542	62620	64759	72186	332
Campbell_et_al	42108	90017	90020	102442	410
Ribosomal_RNAs	319	518	568	511	10
BUSCO_83_Protista	2131	4279	4626	4677	24
Archaea (Rinke_et_al)	30	46	74	73	1
Bacteria (Campbell_et_al)	257	607	611	743	2
Eukaria (BUSCO_83_Protista)	5	13	14	11	0

Table S 3 | Characteristics of the different redundant bins of MAG ID03_00003.

Bins	length	N. of contigs	N50	GC (%)	Completion	Redundancy
ID01_MAG_00039	1704411	123	22795	36.46	89.20	2.15
ID02_MAG_00019	1890973	54	57808	36.47	93.52	0.72
ID03_MAG_00003	1812852	35	72247	36.51	94.24	0.72
ID04_MAG_00039	1600459	21	86817	36.47	87.76	1.43

Table S 4 | Comparison of some significant COG categories linked with trophic lifestyle according to Lauro et al., 2009.

	copiotroph	oligotroph	MAG ID03_00003
genome size	4.7 Mb	3.8 Mb	1.81 Mb
COG category N (cell motility)	high (3.3%)	low (1%)	Low (0%)
COG category T (signal transduction mechanisms)	high (7.07%)	low (3.63%)	low (1,66%)
COG category V (defense mechanisms)	high (1.47%)	low (1.23%)	low (0,98%)
COG category K (transduction)	high (7.52%)	low (6.62%)	low (4,06%)
COG category Q (secondary metabolite biosynthesis, transport and catabolisme)	low (2.17%)	High (3.64%)	low (1,84%)

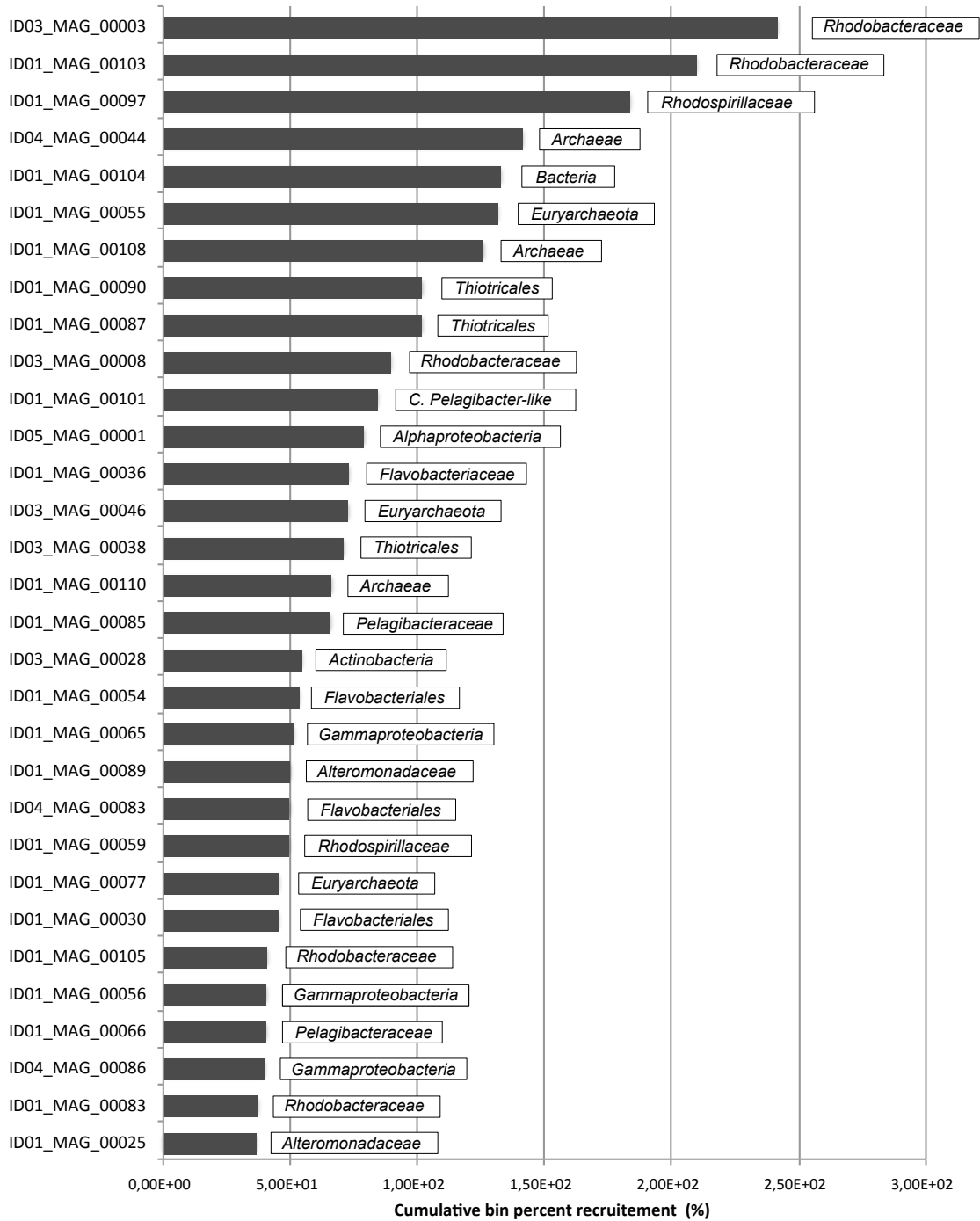


Figure S 12 | 30 most abundant manually refined, non-redundant and high quality MAGs with their CheckM affiliation. They were defined abundant based on their cumulative percentage of read recruitment across the samples.

Discussion and Perspectives

1. THE BREST BAY TIMES SERIES IN THE CONTEXT OF COASTAL GENOMIC OBSERVATORIES

In this work, we introduce new genomic time series in the Bay of Brest and the Iroise Sea in Brittany (France) with the goal to create a long-term coastal microbial observatory. An interesting question is to examine to what extent this observatory is similar to other actual microbial long-term surveys, in term of general characteristics and bacterial diversity.

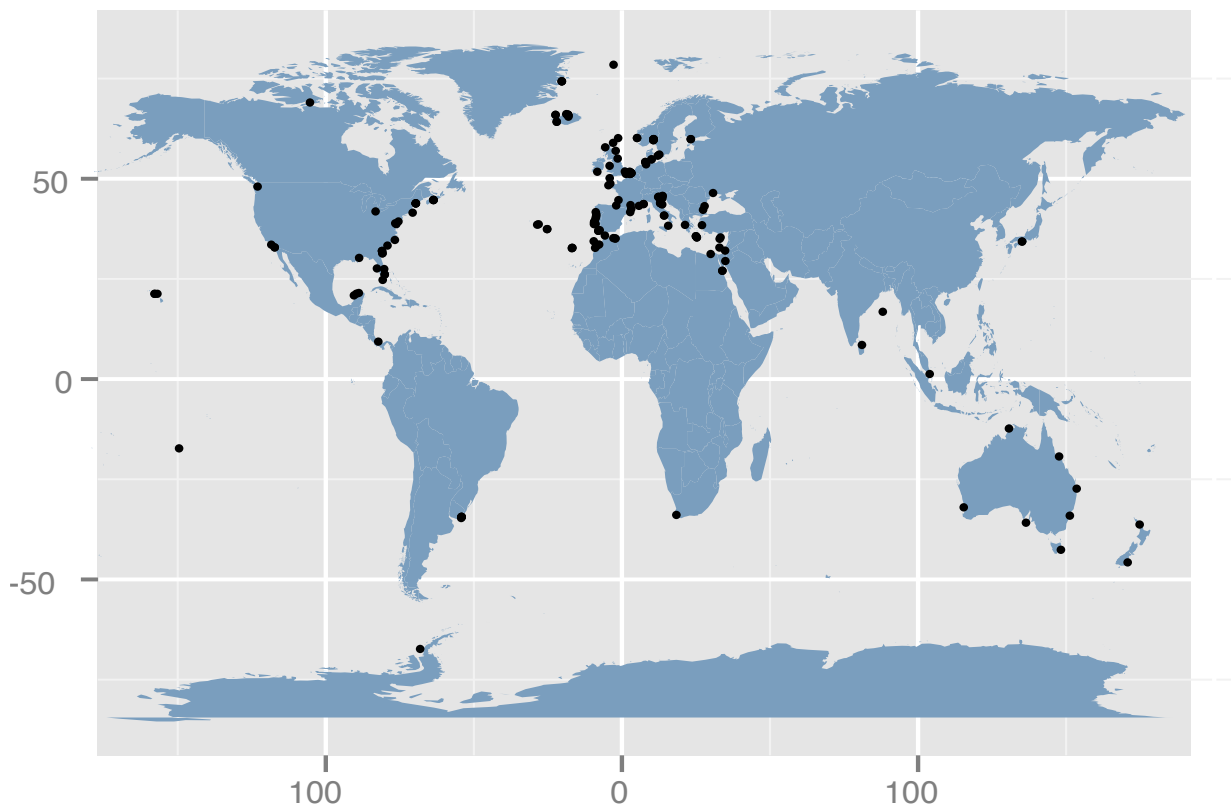


Figure 20 | Localization of the different stations sampled during the OSD consortium the 21st of June 2014.

To assess this question, we investigated the biogeography of microbial communities of the global coastal ocean through the Ocean Sampling Day data (OSD) (Kopf et al., 2015). On the 21st of June 2014, 191 coastal stations were sampled at the same time, and processed using a reproducible

workflow to assess the bacterial diversity at each station, using 16S amplicons sequencing (Figure 20). This allows to directly compare the biodiversity of numerous coastal stations throughout the world, including the SOMLIT station in Brest that was part of the consortium.

Based on OSD data, the global coastal bacterial communities were strongly structured along a seawater temperature gradient (Figure 21). This fits well previous observations of the surface ocean, that showed the importance of latitude and temperature in shaping bacterial communities at large biogeographical scales (Brown et al., 2012; Fuhrman et al., 2008; Swan et al., 2013). Ubiquitous members of the bacterioplankton such as *SAR11 Clade* were demonstrated to present adaptive radiations toward different temperatures (Delmont et al., 2017), emphasizing a niche adaptation to different environmental conditions globally defined by the polar, temperate and tropical biomes (Brown et al., 2012). This is also illustrated by the dominance of *Synechococcus* at our station and the absence of *Prochlorococcus*, which are known to present specific latitudinal distribution (Flombaum et al., 2013). The bacterial diversity observed in the Bay of Brest thus is related to general environmental conditions of a temperate coastal zone.

Within sampling sites of similar temperatures, we can see that the bacterial communities at the SOMLIT station in Brest is very close to the bacterial community at Roscoff, the English Channel L4 station and the Helgoland station in the North Sea (Figure 21). These stations were geographically the closest to our sampling point. In addition to an adaptation toward global temperatures, dispersal limitation could strongly affect the global bacterial communities distribution (Sul et al., 2013). Thus, even if this comparison of microbial community is based on a single sampling station, bacterial communities in the Bay of Brest and the Iroise Sea likely shares common actors, or characteristics with the English Channel L4 station and Helgoland Station in the North Sea. It is widely recognized that we lack of temporal surveys to enhance our comprehension of bacterial communities dynamic (Giovannoni and Vergin, 2012). The presence of our genomic observatory gives a new comparative point for the global ocean and more particularly in the North-East Atlantic shelve provinces. This offers a great opportunity to compare and increase our understanding of the global communities dynamics. As an example, we

found that a heterotroph, affiliated to *Balneatrix* (further affiliated to *Nitrincolaceae* with Silva v132. As in Chap.II) was a biomarker of winter coastal communities in M2BiPAT survey (Chap.I Figure S4 B.) and presented a recurrent cyclic pattern in winter at SOMLIT station (Chap.II Figure 5 b.). The same genus was found to peak in relative abundance, associated to phytoplankton spring blooms at Helgoland road (Teeling et al., 2012). This station is characterized by mean annual low temperature and by strong influence of freshwater (Wiltshire et al., 2010). The combination of these two results suggests that members of *Balneatrix* could be adapted to low temperature and low salinity.

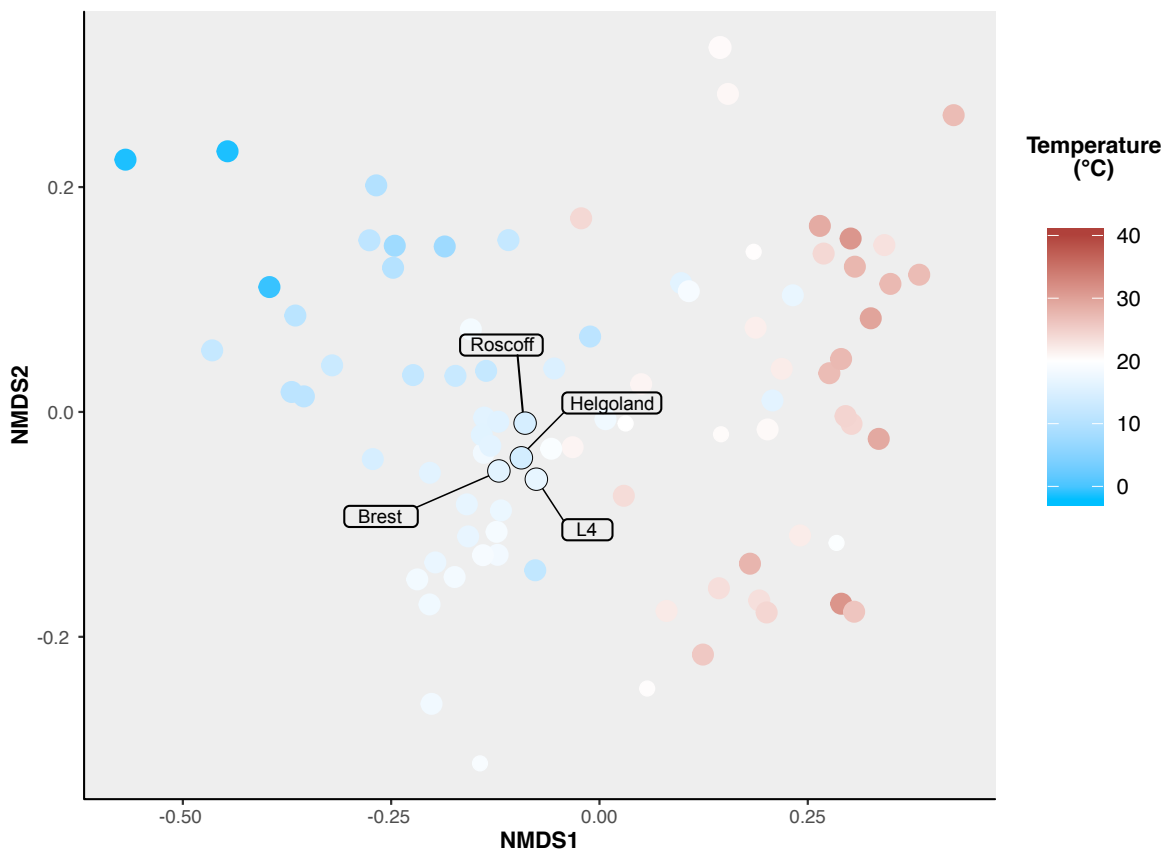


Figure 21 | Non-Metric Dimensional Scaling ordination based on the Bray-Curtis dissimilarity index for the bacterial communities found at the different stations sampled during OSD 2014.

1.1. HOW TO BRIDGE THE GAP BETWEEN PHYSICAL AND CHEMICAL OCEANOGRAPHY AND MICROBIAL ECOLOGY IN SUB-MESOSCALE STRUCTURES?

Understanding the complexity of bacterial communities and their link to global geochemical cycles is one of the main challenges for marine microbial ecology. The use of genomic observatories to decipher such complex roles is primordial. Correlations between variations of specific taxa relative abundances and geochemical parameters permits to generate hypothesis on the activity of different microorganisms in the ecosystem. However, few studies have investigated bacterial dynamic associated to complex sub-mesoscale features (Baltar et al., 2016). These areas represent key drivers of phytoplankton dynamic and production (Lévy et al., 2015) and are well described by physical models that help to precise their role in the coastal ecosystem (Chenillat et al., 2016). Such models have been developed since the 1980s for the Ushant front (Cadier et al., 2017; Mariette and Le Cann, 1985). Bacterial communities are rarely integrated in models, because they have been insufficiently characterized and due to the complexity of their structure making microbial OTUs intractable inputs for these models (Treseder et al., 2012). Moreover, their dynamic is rarely studied at both spatial, bathyal and temporal scale simultaneously. In this work we offer a new perspective for this challenge by extensively sampling a frontal area while reducing the community complexity using a network analysis. Based on this analysis, we suggest that bacteria could be separated in coherent guilds of co-varying OTU, which are distinguished by their trophic strategy. Next step towards integration of molecular data in these biogeochemical models could be, first, a better understanding of the functional role of these modules (or traits) and validate the effective activity of functional guilds using for instance genome-resolved metagenomics (Chap III). Our definition of the different module's strategies relied on the bibliography of few described heterotrophs taxa. It is clear that we still lack knowledge on the actors of the microbial loop (Fenchel, 2008) as illustrated by the discovery of unexpected genomic features of the most abundant heterotroph in our environment, *Amylibacter sp.* To confirm the hypothesis of trophic strategy of the different members of the modules, a better characterization of their genomic content should be continued. In addition to MAG ID03_00003, 43 high-quality MAGs (completion > 95% and redundancy < 5%) were recovered in our study, among which 17 were affiliated to the heterotrophic *Flavobacteriales* order.

Numerous members of this order were abundant in our survey, but poorly characterized as *NS5 marine group*, *NS4 marine group*, *NS9 marine group* or uncultured *Cryomorphaceae*. Different genomic characteristics could help to decipher the trophic strategy of these abundant bacteria. In addition to the oligotrophy or copiotrophy index suggested by Lauro et al., (2009) and utilized in Chapter III, other genes are indicative of a particular metabolism. CAZymes are enzymes involved in polysaccharides degradation that can be separated in different families targeting different specific polysaccharides (Kabisch et al., 2014). Bacteria that possess these systems are often highly specialists for a particular substrate and will present a dynamic strongly related to the presence of their substrate in the environment (Unfried et al., 2018). Secondly, establishing links between genomic data and effective metabolic using novel stable isotope labelled substrate incubations coupled with flow-cytometry and cell sorting (Lee et al., 2019). Or finally to test the predictive power of network module abundances in time and space as input variable in these models. However, the reduction in community's complexity doesn't go without limitations. Network analysis probably clustered taxa with different trophic strategies but that shares complex interactions (Lima-Mendez et al., 2015), or are influenced similarly by environmental parameters as we hypothesized for the presence of nitrite-oxidizing bacteria (*Nitrospina sp*) in the oligotrophic module. The temporal resolution of the sampling could not be sufficient to catch the complex dynamic of numerous taxa as marine microbial communities are highly dynamic in short period of times (Needham et al., 2018). Better delineation of meaningful modules will come at a cost of increased sampling efforts, both spatially and temporally (Martin-Platero et al., 2018).

In addition, this study strongly supports the view that the availability and type of organic matter could represent a key driver of bacterial communities associated to a tidal front. This link between bacterial communities and organic matter is one of the fundamental observations of marine microbial ecology that led to the microbial loop paradigm in the early 1970 (Azam et al., 1983; Azam and Malfatti, 2007). But we are just starting to understand the complex link between distinct heterotrophic bacteria, and organic matter, mainly due to the technical limitations of both fields that evolved quickly only in the last decade (Hertkorn et al., 2013; Koch et al., 2008). In our studies, we used chlorophyll a, phytoplankton cells counts and Particulate Organic Matter levels as a proxy to define copiotrophic (or oligotrophic)

environment for heterotrophic bacteria. While these parameters are inter-dependant – phytoplankton is one of the important producers of DOM in surface pelagic environment (Thornton, 2014) – they do not unambiguously reflect the availability of dissolved organic matter for heterotrophic bacteria. A PhD work carried by Johann Breitenstein (LEMAR) has been investigating the quantification and characterization of dissolved organic matter at the SOMLIT station since 2017. Preliminary results show that both low molecular weight molecules (e.g. glucose, amino acids) and High Molecular Weight molecules (e.g. proteins, polysaccharides) present a seasonal variation that can depart from chlorophyll a profile (Figure 22.A). Unfortunately there are only few data shared (February to May 2017) with the microbial survey (Figure 22), but the continued sampling for microbial diversity throughout 2017 and 2018 at the SOMLIT station will provide interesting data to closely compare the microbial diversity and the ambient dissolved organic matter concentrations.

Finally, relative abundance measures do not reflect the true influence of a taxa and biogeochemical cycles. For instance, less abundant taxa can contribute to a substantial fraction of nutrient uptake {Citation}. To reveal the importance of marine bacteria on the processing of organic matter in the Iroise Sea, further studies must include quantification of their abundance or biomass using qPCR and cytometry, and the measure of substrates incorporation by Stable Isotope probing

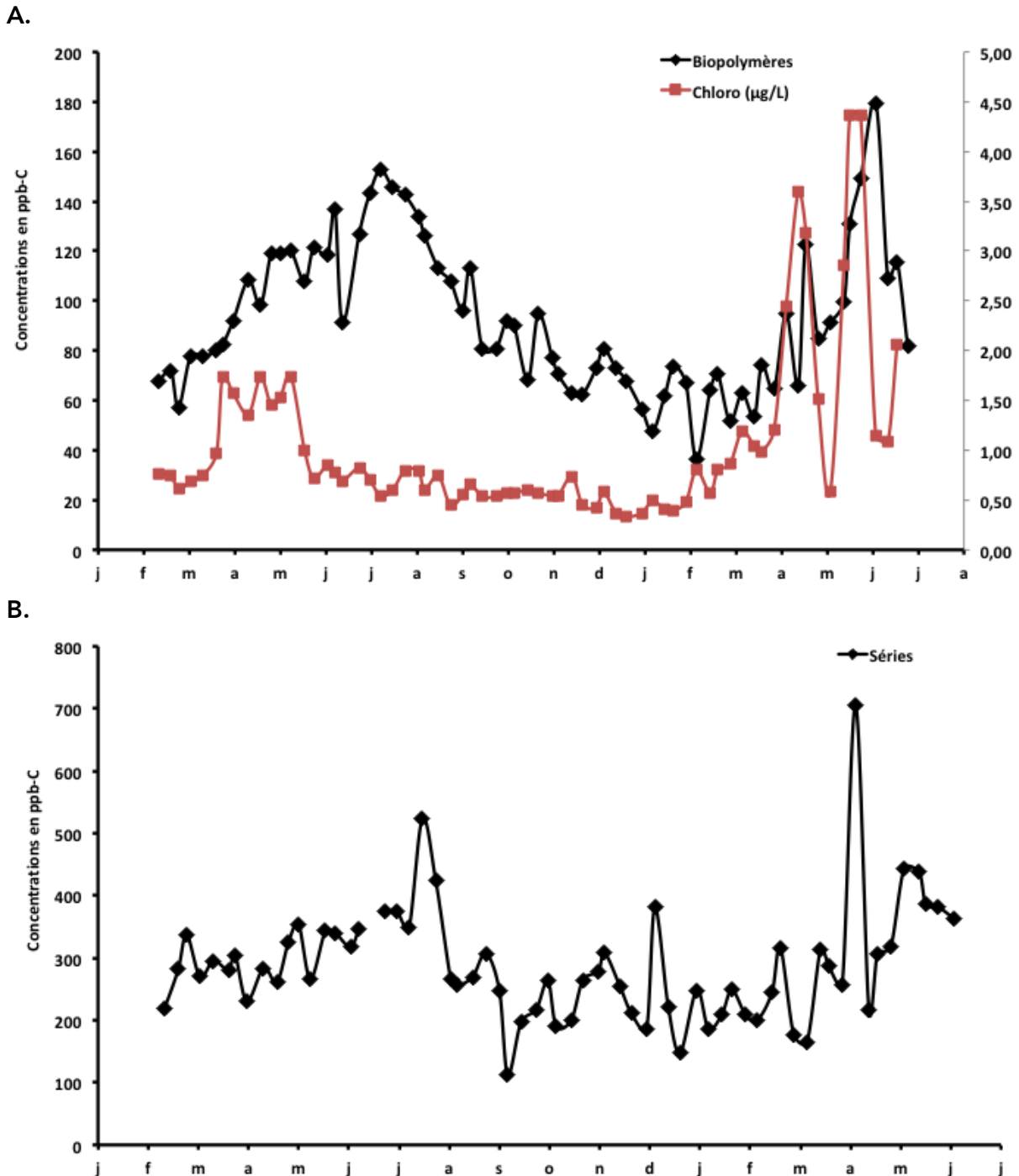


Figure 22 | Variation in quantity of dissolved organic matter (A.) high molecular weight (in black) and Chlorophyll a (in red) and low molecular weight (B.) molecules of Dissolved Organic Matter measured at the SOMLIT station between February 2017 and July 2018. (Johann Breitenstein, personal communication)

2. FROM ABIOTIC TO BIOTIC DRIVERS

Within the complexity of the pelagic microbial world, this work only analyzed the dynamic of free-living bacteria and their relation to the environmental, mostly abiotics, parameters. However, there is a growing interest concerning the importance of the biotic compartment on bacterial communities composition dynamic (Needham et al., 2018; Zhou et al., 2018). Microbial food web present in pelagic waters are organized in complex networks of trophic, symbiotic, or competitive interactions (Lima-Mendez et al., 2015; Worden et al., 2015).

One of the intriguing observations that lead to this interest is the stability of seasonal patterns years after years (Fuhrman et al., 2015). This cyclical pattern was recovered at the SOMLIT temporal sampling as illustrated with the CCA (Chap. II, Figure 8). This seasonality relied on coherent communities, represented by modules that were recurrent from one year to another, as observe at Helgoland Station in the North Sea (Chafee et al., 2018). Also, we hypothesized that some heterotrophs taxa that peaked each year after phytoplankton blooms could follow deterministic patterns and respond to metabolic cues released by phytoplankton (Teeling et al., 2016). The recovering of *Amylibacter* genome further strengthened this hypothesis: this recurrent taxon likely competes for small and common substrates that are produced in high quantity under bloom conditions. But the investigation of other members of the microbial communities could highlight different drivers, as the importance of top-down controls bacterial communities dynamic (Needham et al., 2018).

During this work we produced data for particle-attached Bacteria and Archaea for the spatio-temporal survey in the Iroise Sea, as well as pelagic Archaea for the two surveys. Bacteria and Archaea can be linked through metabolic interactions (Bayer et al., 2019). Ammonia-oxidizing archaea were found to release significant amount of labile dissolved organic matter, that can fuel heterotrophic bacteria (Bayer et al., 2019). In addition, a recent work found that ammonia-oxidizing archaea are closely associated with pelagic nitrite-oxidizing bacteria such as *Nitrospina* sp. that perform the second step of nitrification (Parada and Fuhrman, 2017). By combining the two datasets (bacteria and archaea), we could verify these observations in our environment, but also prospect for interactions that are still unsuspected.

3. A PERSPECTIVE: BETTER UNDERSTANDING OF ANTHROPIC PRESSURES ON COASTAL ECOSYSTEMS

In a changing environment, another important role of long-term survey is to monitor anthropic disturbance in sensitive environments. The Bay of Brest represents a key area where a rich ecosystem with a remarkable biodiversity meets important anthropic activities. It undergoes numerous sources of pollutions with, among others, the most important commercial harbour of Brittany, an increasing agricultural nitrogen releases from its two dominant catchment areas and the presence of two wastewater treatment plants of a capacity of 170 000 inhabitants. The observation of microorganisms in such environments are essential to 1/ monitor the evolution of natural communities and the impact of different events related to human activity on them 2/ monitor the emergence of possible pathogenic bacteria associated with ecosystem changes.

Intensive researches are led on the diversity and ecology of pathogenic *Vibrio* present in the environment of the Bay of Brest through two axes: human pathogens (Ifremer - LSEM) and bivalves pathogens (LEMAR). A monitoring of specific human pathogenic *Vibrio* was investigated in the three-years SOMLIT samples, by qPCR on the same extracted DNA (unpublished data). In a context of climate change, pioneering studies have led to a better understanding of seasonality and interannual patterns in the cycles of oceanic (Fuhrman et al., 2015; Giovannoni and Vergin, 2012) and coastal (Cram et al., 2015) microbial communities. Recent retrospective long-term monitoring studies (1961-2004) in the southern North Sea have shown that a major change occurred in the composition of bacterial communities, and that most of this change was attributed to an increase in the abundance of the Vibrionaceae family, including human pathogens (Vezzulli et al., 2012). In addition, a positive and significant relationship between global warming and the spread of vibrios in the environment have also been demonstrated (Vezzulli et al., 2013). Global warming and the emergence of vibriosis in molluscs have also been highlighted, for example, the links between abalone mortality and global warming (Travers et al., 2009) and the spread of brown ring disease in clams with a theoretical individual-based model using increases of 1 and 2°C by simulation (Paillard et al., 2014). The Bay of Brest represents an

important site for mollusc and fish farming and fisheries and therefore pathogen survey is essential to prevent epizooties. In this way, the objective will be to integrate definitively this monitoring of prokaryotic biodiversity in the IUEM observatory. In the longer term, microorganism observatory could act in synergy with the INEE "Zone Atelier" ZA network, and in coordination with IFREMER observations networks (REPHY, REMI, REPAMO...).

To further tackle the issue of anthropogenic disturbance in our coastal ecosystem, we also participated to the World Harbours project (<http://www.worldharbourproject.org>) initiated by the Sydney Institute of Marine Sciences, which brings together different research groups worldwide to study issues surrounding the multiple uses of harbours and ports. Samples in the Bay of Brest were taken in the specific project lead by Anthony Chariton (Mcquarie University), Sandra McLellan (University of Wisconsin-Milwaukee) and Peter Steinberg (University of New South Wales) to address two questions: first, decipher the source (i.e. sewage, urban stormwater and agricultural runoff) of faecal contaminants in harbours around the world using qPCR and high throughput sequencing and second, examine the relationship between sewage inputs and genetic elements associated with antibiotic resistance. We sampled the wastewater treatment plant of Brest (both in and outflow streams), the commercial harbour basin and the SOMLIT station, which was taken as a referent point. We anticipate that the results of this study, in comparison with those performed in chap I and II, will document the impact and persistence of wastewater and industrial site water microorganisms in the Bay of Brest and Iroise Sea.

Finally, the contamination of marine environments by small plastic particles has emerged as a huge issue globally. For microorganisms, it raises concerns regarding microplastics capacity to concentrate and propagate pathogenic bacteria beyond their natural dispersal range (Dussud et al., 2018). Bacterial communities attached to microplastic in the Bay of Brest were investigated during the thesis of Frère, 2017. During this work, she observed that microplastics present a statistically different community of bacteria compared to seawater, with two biomarkers including *Vibrio* and potentially pathogenic species *V. splendidus* detected by qPCR. However, in this study authors highlighted the lack of

“particle” control, to verify the difference colonization between microplastics and common particles present in the surrounding waters (Frère et al., 2018). Our temporal survey, and particularly the study of the particle-attached fraction should give a better understanding as to which extent this microplastics harbour a specific diversity compared to naturally occurring marine particles.

General conclusion

In conclusion, we report in this work the successful set-up of a new marine microbial observatory in the North-East Atlantic temperate coastal waters. The analysis of the first data provided by this new survey gave a first view of the global bacterial communities dynamic in this area. In the Bay of Brest, bacterial communities seem to be typical of temperate coastal waters. They present strong and recurrent seasonal variations that drastically change with the set-up of the spring blooms. In the Iroise Sea, bacterial communities are strongly shaped by the presence of a frontal area and particularly by the primary producers and the availability of organic matter in this environment. The reconstruction of a new and important genome of the global bacterial communities underlines how these bacterial communities of these temperate coastal waters are still insufficiently described and that the set-up of new surveys is highly relevant. Next steps include the analysis of new genomes and characterization of the organic matter, to better understand the role of these communities in global biogeochemical cycles. Secondly, to consider Archaea and other size fractions to better understand other drivers such as biotic drivers. Finally, this work provides a set of data allowing the identification of potential bacterial pathogens to human and marine organisms and also those potentially useful as environmental biomarkers (global warming, toxic algae, pollutants, microplastics...). This microbial time-series will have to overcome the most complex problems for all time surveys: to continue as long as possible. As underlined in a paradox: long term surveys as often the shortest experiments in oceanography (Ducklow et al., 2009). And it is only at long-time scales that numerous key ecological processes can be observed (e.g. episodic or complex phenomena). This is also only at long scales that long-term survey can help to understand detect and predict climate changes (Ducklow et al., 2009).

References

REFERENCES

- Agogu , H., Lamy, D., Neal, P. R., Sogin, M. L., and Herndl, G. J. (2011). Water mass-specificity of bacterial communities in the North Atlantic revealed by massively parallel sequencing. *Mol Ecol* 20, 258–274. doi:10.1111/j.1365-294X.2010.04932.x.
- Albertsen, M., Hugenholtz, P., Skarshewski, A., Nielsen, K. L., Tyson, G. W., and Nielsen, P. H. (2013). Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. *Nature Biotechnology* 31, 533–538. doi:10.1038/nbt.2579.
- Allen, L. Z., Allen, E. E., Badger, J. H., McCrow, J. P., Paulsen, I. T., Elbourne, L. D., et al. (2012). Influence of nutrients and currents on the genomic composition of microbes across an upwelling mosaic. *The ISME Journal* 6, 1403–1414. doi:10.1038/ismej.2011.201.
- Alneberg, J., Bjarnason, B. S., de Bruijn, I., Schirmer, M., Quick, J., Ijaz, U. Z., et al. (2014). Binning metagenomic contigs by coverage and composition. *Nature Methods* 11, 1144–1146. doi:10.1038/nmeth.3103.
- Alonso, C., and Pernthaler, J. (2006). Roseobacter and SAR11 dominate microbial glucose uptake in coastal North Sea waters. *Environmental Microbiology* 8, 2022–2030. doi:10.1111/j.1462-2920.2006.01082.x.
- Alonso-S ez, L., D az-P rez, L., and Mor n, X. A. G. (2015). The hidden seasonality of the rare biosphere in coastal marine bacterioplankton. *Environmental Microbiology* 17, 3766–3780. doi:10.1111/1462-2920.12801.
- Alonso-S ez, L., S nchez, O., Gasol, J. M., Balagu , V., and Pedr s-Alio, C. (2008). Winter-to-summer changes in the composition and single-cell activity of near-surface Arctic prokaryotes. *Environ. Microbiol.* 10, 2444–2454. doi:10.1111/j.1462-2920.2008.01674.x.
- Aminot, A., and K rouel, R. (2007). *Dosage automatique des nutriments dans les eaux marines: m thodes en flux continu*. Editions Quae.
- Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome biology* 11, R106.
- Apprill, A., Marlow, H. Q., Martindale, M. Q., and Rapp , M. S. (2009). The onset of microbial associations in the coral *Pocillopora meandrina*. *The ISME Journal* 3, 685–699. doi:10.1038/ismej.2009.3.
- Azam, F. (1998). Microbial Control of Oceanic Carbon Flux: The Plot Thickens. *Science* 280, 694–696. doi:10.1126/science.280.5364.694.

-
- Azam, F., Fenchel, T., Field, J., Gray, J., Meyer-Reil, L., and Thingstad, F. (1983). The Ecological Role of Water-Column Microbes in the Sea. *Marine Ecology Progress Series* 10, 257–263. doi:10.3354/meps010257.
- Azam, F., and Long, R. A. (2001). Oceanography: Sea snow microcosms. *Nature*. Available at: <http://link.galegroup.com/apps/doc/A187934204/AONE?sid=googlescholar> [Accessed March 20, 2019].
- Azam, F., and Malfatti, F. (2007). Microbial structuring of marine ecosystems. *Nature Reviews Microbiology* 5, 782–791. doi:10.1038/nrmicro1747.
- Azam, F., Smith, D. C., Steward, G. F., and Hagström, Å. (1994). Bacteria-organic matter coupling and its significance for oceanic carbon cycling. *Microb Ecol* 28, 167–179. doi:10.1007/BF00166806.
- Azam, F., and Worden, A. Z. (2004). Microbes, molecules, and marine ecosystems. *Science* 303, 1622–1624.
- Baas-Becking, L. G. M. (1934). Geobiologie; of inleiding tot de milieukunde. *WP Van Stockum & Zoon NV*.
- Baines, S. B., and Pace, M. L. (1991). The production of dissolved organic matter by phytoplankton and its importance to bacteria: Patterns across marine and freshwater systems. *Limnology and Oceanography* 36, 1078–1090. doi:10.4319/lo.1991.36.6.1078.
- Baltar, F., Arístegui, J., Gasol, J. M., Sintes, E., Aken, H. M. van, and Herndl, G. J. (2010). High dissolved extracellular enzymatic activity in the deep central Atlantic Ocean. *Aquatic Microbial Ecology* 58, 287–302. doi:10.3354/ame01377.
- Baltar, F., Currie, K., Stuck, E., Roosa, S., and Morales, S. E. (2016). Oceanic fronts: transition zones for bacterioplankton community composition: Fronts delimit bacterioplankton communities. *Environmental Microbiology Reports* 8, 132–138. doi:10.1111/1758-2229.12362.
- Baltar, F., Stuck, E., Morales, S., and Currie, K. (2015). Bacterioplankton carbon cycling along the Subtropical Frontal Zone off New Zealand. *Progress in Oceanography* 135, 168–175. doi:10.1016/j.pocean.2015.05.019.
- Barberán, A., and Casamayor, E. O. (2010). Global phylogenetic community structure and β -diversity patterns in surface bacterioplankton metacommunities. *Aquatic Microbial Ecology* 59, 1–10. doi:10.3354/ame01389.
- Barnes, R. S. K., and Hughes, R. N. (1999). *An Introduction to Marine Ecology, Third Edition*. Wiley-Blackwell.

REFERENCES

- Barra Caracciolo, A., Topp, E., and Grenni, P. (2015). Pharmaceuticals in the environment: Biodegradation and effects on natural microbial communities. A review. *Journal of Pharmaceutical and Biomedical Analysis* 106, 25–36. doi:10.1016/j.jpba.2014.11.040.
- Bastian, M., Heymann, S., and Jacomy, M. (2009). Gephi: an open source software for exploring and manipulating networks. *Icwsn* 8, 361–362.
- Bayer, B., Hansman, R. L., Bittner, M. J., Noriega-Ortega, B. E., Niggemann, J., Dittmar, T., et al. (2019). Ammonia-oxidizing archaea release a suite of organic compounds potentially fueling prokaryotic heterotrophy in the ocean. *bioRxiv*, 558726. doi:10.1101/558726.
- Béjà, O., Aravind, L., Koonin, E. V., Suzuki, M. T., Hadd, A., Nguyen, L. P., et al. (2000). Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. *Science* 289, 1902–1906.
- Benoit, G., Peterlongo, P., Mariadassou, M., Drezen, E., Schbath, S., Lavenier, D., et al. (2016). Multiple comparative metagenomics using multiset k-mer counting. *PeerJ Comput. Sci.* 2, e94. doi:10.7717/peerj-cs.94.
- Beucher, C., Tréguer, P., Corvaisier, R., Hapette, A., and Elskens, M. (2004). Production and dissolution of biosilica, and changing microphytoplankton dominance in the Bay of Brest (France). *Marine Ecology Progress Series* 267, 57–69. doi:10.3354/meps267057.
- Biddanda, B., and Benner, R. (1997). Carbon, nitrogen, and carbohydrate fluxes during the production of particulate and dissolved organic matter by marine phytoplankton. *Limnology and Oceanography* 42, 506–518.
- Birrien, J. L., Wafar, M. V. M., Corre, P. L., and Riso, R. (1991). Nutrients and primary production in a shallow stratified ecosystem in the Iroise Sea. *Journal of Plankton Research* 13, 721–742. doi:10.1093/plankt/13.4.721.
- Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *J. Stat. Mech.* 2008, P10008. doi:10.1088/1742-5468/2008/10/P10008.
- Bonnet, S., Caffin, M., Berthelot, H., and Moutin, T. (2017). Hot spot of N₂ fixation in the western tropical South Pacific pleads for a spatial decoupling between N₂ fixation and denitrification. *Proceedings of the National Academy of Sciences* 114, E2800–E2801.
- Bowers, R. M., Kyrpides, N. C., Stepanauskas, R., Harmon-Smith, M., Doud, D., Reddy, T. B. K., et al. (2017). Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nature Biotechnology* 35, 725–731. doi:10.1038/nbt.3893.

-
- Bowman, J. P. (2014). "The Family Cryomorphaceae," in *The Prokaryotes*, eds. E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt, and F. Thompson (Berlin, Heidelberg: Springer Berlin Heidelberg), 539–550. doi:10.1007/978-3-642-38954-2_135.
- Brochier-Armanet, C., Boussau, B., Gribaldo, S., and Forterre, P. (2008). Mesophilic Crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nat. Rev. Microbiol.* 6, 245–252. doi:10.1038/nrmicro1852.
- Broecker, W. S. (1991). THE GREAT OCEAN CONVEYOR. *Oceanography* 4, 11.
- Brown, E. A., Chain, F. J. J., Crease, T. J., Maclsaac, H. J., and Cristescu, M. E. (2015). Divergence thresholds and divergent biodiversity estimates: can metabarcoding reliably describe zooplankton communities? *Ecol Evol* 5, 2234–2251. doi:10.1002/ece3.1485.
- Brown, M. V., Lauro, F. M., DeMaere, M. Z., Muir, L., Wilkins, D., Thomas, T., et al. (2012). Global biogeography of SAR11 marine bacteria. *Molecular Systems Biology* 8, 595. doi:10.1038/msb.2012.28.
- Bryant, J. A., Aylward, F. O., Eppley, J. M., Karl, D. M., Church, M. J., and DeLong, E. F. (2016). Wind and sunlight shape microbial diversity in surface waters of the North Pacific Subtropical Gyre. *The ISME Journal* 10, 1308–1322. doi:10.1038/ismej.2015.221.
- Bryson, S., Li, Z., Chavez, F., Weber, P. K., Pett-Ridge, J., Hettich, R. L., et al. (2017). Phylogenetically conserved resource partitioning in the coastal microbial loop. *The ISME Journal* 11, 2781–2792. doi:10.1038/ismej.2017.128.
- Buchan, A., LeClerc, G. R., Gulvik, C. A., and González, J. M. (2014). Master recyclers: features and functions of bacteria associated with phytoplankton blooms. *Nature Reviews Microbiology* 12, 686–698. doi:10.1038/nrmicro3326.
- Bunse, C., Bertos-Fortis, M., Sassenhagen, I., Sildever, S., Sjöqvist, C., Godhe, A., et al. (2016). Spatio-Temporal Interdependence of Bacteria and Phytoplankton during a Baltic Sea Spring Bloom. *Front. Microbiol.* 7. doi:10.3389/fmicb.2016.00517.
- Bunse, C., and Pinhassi, J. (2017). Marine Bacterioplankton Seasonal Succession Dynamics. *Trends in Microbiology* 25, 494–505. doi:10.1016/j.tim.2016.12.013.
- Cadier, M., Gorgues, T., Sourisseau, M., Edwards, C. A., Aumont, O., Marié, L., et al. (2017). Assessing spatial and temporal variability of phytoplankton communities' composition in the Iroise Sea ecosystem (Brittany, France): A 3D modeling approach. Part 1: Biophysical control over plankton functional types succession and distribution. *Journal of Marine Systems* 165, 47–68. doi:10.1016/j.jmarsys.2016.09.009.

REFERENCES

- Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Meth* advance online publication. Available at: <http://dx.doi.org/10.1038/nmeth.3869>.
- Campbell, B. J., Yu, L., Heidelberg, J. F., and Kirchman, D. L. (2011). Activity of abundant and rare bacteria in a coastal ocean. *Proceedings of the National Academy of Sciences* 108, 12776–12781. doi:10.1073/pnas.1101405108.
- Campbell, J. H., O'Donoghue, P., Campbell, A. G., Schwientek, P., Sczyrba, A., Woyke, T., et al. (2013). UGA is an additional glycine codon in uncultured SR1 bacteria from the human microbiota. *PNAS* 110, 5540–5545. doi:10.1073/pnas.1303090110.
- Caron, D. A., Goldman, J. C., and Dennett, M. R. (1988). Experimental demonstration of the roles of bacteria and bacterivorous protozoa in plankton nutrient cycles. *Hydrobiologia* 159, 27–40. doi:10.1007/BF00007365.
- Casotti, R., Brunet, C., Aronne, B., and d'Alcala, M. R. (2000). Mesoscale features of phytoplankton and planktonic bacteria in a coastal area as induced by external water masses. *Marine Ecology Progress Series* 195, 15–27. doi:10.3354/meps195015.
- Chafee, M., Fernández-Guerra, A., Buttigieg, P. L., Gerdts, G., Eren, A. M., Teeling, H., et al. (2018). Recurrent patterns of microdiversity in a temperate coastal marine environment. *The ISME Journal* 12, 237–252. doi:10.1038/ismej.2017.165.
- Charlson, R. J., Lovelock, J. E., Andreae, M. O., and Warren, S. G. (1987). Oceanic phytoplankton, atmospheric sulphur, cloud albedo and climate. *nature* 326, 655.
- Chenillat, F., Franks, P. J. S., and Combes, V. (2016). Biogeochemical properties of eddies in the California Current System: California Current Ecosystem Eddies. *Geophysical Research Letters* 43, 5812–5820. doi:10.1002/2016GL068945.
- Choi, D. H., Kwon, Y. M., Chiura, H. X., Yang, E. C., Bae, S. S., Kang, S. G., et al. (2015). Extracellular Vesicles of the Hyperthermophilic Archaeon “*Thermococcus onnurineus*” NA1T. *Appl. Environ. Microbiol.* 81, 4591–4599. doi:10.1128/AEM.00428-15.
- Chow, C.-E. T., Sachdeva, R., Cram, J. A., Steele, J. A., Needham, D. M., Patel, A., et al. (2013). Temporal variability and coherence of euphotic zone bacterial communities over a decade in the Southern California Bight. *The ISME journal* 7, 2259.
- Christie-Oleza, J. A., Fernandez, B., Nogales, B., Bosch, R., and Armengaud, J. (2012). Proteomic insights into the lifestyle of an environmentally relevant marine bacterium. *ISME J* 6, 124–135. doi:10.1038/ismej.2011.86.

-
- Cloern, J. E., Foster, S. Q., and Kleckner, A. E. (2014). Phytoplankton primary production in the world's estuarine-coastal ecosystems. *Biogeosciences* 11, 2477–2501. doi:<https://doi-org.scd-proxy.univ-brest.fr/10.5194/bg-11-2477-2014>.
- Costanza, R., d'Arge, R., Groot, R. de, Farber, S., Grasso, M., Hannon, B., et al. (1997). The value of the world's ecosystem services and natural capital. *Nature* 387, 253. doi:10.1038/387253a0.
- Costanza, R., de Groot, R., Sutton, P., Van der Ploeg, S., Anderson, S. J., Kubiszewski, I., et al. (2014). Changes in the global value of ecosystem services. *Global environmental change* 26, 152–158.
- Cotner, J. B., and Biddanda, B. A. (2002). Small Players, Large Role: Microbial Influence on Biogeochemical Processes in Pelagic Aquatic Ecosystems. *Ecosystems* 5, 105–121. doi:10.1007/s10021-001-0059-3.
- Cottrell, M. T., and Kirchman, D. (2003). Contribution of major bacterial groups to bacterial biomass production (thymidine and leucine incorporation) in the Delaware estuary. *Limnology and Oceanography* 48, 168–178. doi:10.4319/lo.2003.48.1.0168.
- Cram, J. A., Chow, C.-E. T., Sachdeva, R., Needham, D. M., Parada, A. E., Steele, J. A., et al. (2015). Seasonal and interannual variability of the marine bacterioplankton community throughout the water column over ten years. *ISME J* 9, 563–580. doi:10.1038/ismej.2014.153.
- Crespo, B. G., Pommier, T., Fernández-Gómez, B., and Pedrós-Alió, C. (2013). Taxonomic composition of the particle-attached and free-living bacterial assemblages in the Northwest Mediterranean Sea analyzed by pyrosequencing of the 16S rRNA. *MicrobiologyOpen* 2, 541–552. doi:10.1002/mbo3.92.
- Croft, M. T., Lawrence, A. D., Raux-Deery, E., Warren, M. J., and Smith, A. G. (2005). Algae acquire vitamin B12 through a symbiotic relationship with bacteria. *Nature* 438, 90–93. doi:10.1038/nature04056.
- De Meester, L. (2011). "A metacommunity perspective on the phylo- and biogeography of small organisms," in *Biogeography of Microscopic Organisms: Is Everything Small Everywhere?* Systematics Association Special Volume Series., ed. D. Fontaneto (Cambridge: Cambridge University Press), 324–334. doi:10.1017/CBO9780511974878.017.
- Del Amo, Y., Le Pape, O., Tréguer, P., Quéguiner, B., Ménesguen, A., and Aminot, A. (1997). Impacts of high-nitrate freshwater inputs on macrotidal ecosystems. I. Seasonal evolution of nutrient limitation for the diatom-dominated phytoplankton of the Bay of Brest (France). *Marine Ecology Progress Series* 161, 213–224.
- Delmas, R. (1981). Etude de l'évolution saisonniere des sels nutritifs dans la rade de Brest en fonction des apports fluviaux et des échanges avec l'Iroise.

REFERENCES

- Delmas, R., and Treguer, P. (1983). Évolution saisonnière des nutriments dans un écosystème eutrophe d'Europe occidentale (la rade de Brest). *Interactions marines et terrestres. Oceanologica Acta* 6, 345–356. Available at: <https://archimer.ifremer.fr/doc/00120/23122/> [Accessed March 21, 2019].
- Delmont, T. O., Kiefl, E., Kilinc, O., Esen, O. C., Uysal, I., Rappe, M. S., et al. (2017). The global biogeography of amino acid variants within a single SAR11 population is governed by natural selection. *bioRxiv*, 170639. doi:10.1101/170639.
- Delmont, T. O., Quince, C., Shaiber, A., Esen, Ö. C., Lee, S. T., Rappé, M. S., et al. (2018). Nitrogen-fixing populations of Planctomycetes and Proteobacteria are abundant in surface ocean metagenomes. *Nature Microbiology* 3, 804. doi:10.1038/s41564-018-0176-9.
- DeLong, E. F. (1992). Archaea in coastal marine environments. *Proceedings of the National Academy of Sciences* 89, 5685–5689.
- DeLong, E. F., and Pace, N. R. (2001). Environmental diversity of bacteria and archaea. *Systematic biology* 50, 470–478.
- Doney, S. C., Ruckelshaus, M., Emmett Duffy, J., Barry, J. P., Chan, F., English, C. A., et al. (2012). Climate Change Impacts on Marine Ecosystems. *Annual Review of Marine Science* 4, 11–37. doi:10.1146/annurev-marine-041911-111611.
- Duarte, C. M. (2015). Seafaring in the 21st Century: The Malaspina 2010 Circumnavigation Expedition. *Limnology and Oceanography Bulletin* 24, 11–14. doi:10.1002/lob.10008.
- Ducklow, H. W., Doney, S. C., and Steinberg, D. K. (2009). Contributions of Long-Term Research and Time-Series Observations to Marine Ecology and Biogeochemistry. *Annu. Rev. Mar. Sci.* 1, 279–302. doi:10.1146/annurev.marine.010908.163801.
- Ducklow, H. W., Steinberg, D. K., and Buesseler, K. O. (2001). Upper ocean carbon export and the biological pump. *OCEANOGRAPHY-WASHINGTON DC-OCEANOGRAPHY SOCIETY-* 14, 50–58.
- Dupont, C. L., Rusch, D. B., Yooseph, S., Lombardo, M.-J., Alexander Richter, R., Valas, R., et al. (2012). Genomic insights to SAR86, an abundant and uncultivated marine bacterial lineage. *The ISME Journal* 6, 1186–1199. doi:10.1038/ismej.2011.189.
- DuRand, M. D., Olson, R. J., and Chisholm, S. W. (2001). Phytoplankton population dynamics at the Bermuda Atlantic Time-series station in the Sargasso Sea. *Deep Sea Research Part II: Topical Studies in Oceanography* 48, 1983–2003. doi:10.1016/S0967-0645(00)00166-1.

-
- Durham, B. P., Grote, J., Whittaker, K. A., Bender, S. J., Luo, H., Grim, S. L., et al. (2014). Draft genome sequence of marine alphaproteobacterial strain HIMB11, the first cultivated representative of a unique lineage within the Roseobacter clade possessing an unusually small genome. *Standards in Genomic Sciences* 9, 632. doi:10.4056/sigs.4998989.
- Dussud, C., Meistertzheim, A. L., Conan, P., Pujo-Pay, M., George, M., Fabre, P., et al. (2018). Evidence of niche partitioning among bacteria living on plastics, organic particles and surrounding seawaters. *Environmental Pollution* 236, 807–816. doi:10.1016/j.envpol.2017.12.027.
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32, 1792–1797. doi:10.1093/nar/gkh340.
- Eiler, A., Heinrich, F., and Bertilsson, S. (2012). Coherent dynamics and association networks among lake bacterioplankton taxa. *The ISME Journal*, 330–342.
- Eren, A. M., Esen, Ö. C., Quince, C., Vineis, J. H., Morrison, H. G., Sogin, M. L., et al. (2015). Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* 3, e1319. doi:10.7717/peerj.1319.
- Eren, A. M., Maignien, L., Sul, W. J., Murphy, L. G., Grim, S. L., Morrison, H. G., et al. (2013a). Oligotyping: Differentiating between closely related microbial taxa using 16S rRNA gene data. *Methods Ecol Evol* 4. doi:10.1111/2041-210X.12114.
- Eren, A. M., Morrison, H. G., Lescault, P. J., Reveillaud, J., Vineis, J. H., and Sogin, M. L. (2014). Minimum entropy decomposition: Unsupervised oligotyping for sensitive partitioning of high-throughput marker gene sequences. *The ISME Journal*. doi:10.1038/ismej.2014.195.
- Eren, A. M., Vineis, J. H., Morrison, H. G., and Sogin, M. L. (2013b). A filtering method to generate high quality short reads using Illumina paired-end technology. *PloS one* 8, e66643.
- Eren, A. M., Vineis, J. H., Morrison, H. G., and Sogin, M. L. (2013c). Correction: A Filtering Method to Generate High Quality Short Reads Using Illumina Paired-End Technology. *PLoS ONE* 8, 10.1371/annotation/afa5c40d-c604-46ae-84c4-82cb92193a5e. doi:10.1371/annotation/afa5c40d-c604-46ae-84c4-82cb92193a5e.
- Falkowski, P. G., Fenchel, T., and Delong, E. F. (2008). The Microbial Engines That Drive Earth's Biogeochemical Cycles. *Science* 320, 1034–1039. doi:10.1126/science.1153213.
- Fenchel, T. (2008). The microbial loop – 25 years later. *Journal of Experimental Marine Biology and Ecology* 366, 99–103. doi:10.1016/j.jembe.2008.07.013.

REFERENCES

- Fenchel, T. M., and Jørgensen, B. B. (1977). "Detritus Food Chains of Aquatic Ecosystems: The Role of Bacteria," in *Advances in Microbial Ecology Advances in Microbial Ecology.*, ed. M. Alexander (Boston, MA: Springer US), 1–58. doi:10.1007/978-1-4615-8219-9_1.
- Fernández-Gómez, B., Richter, M., Schüller, M., Pinhassi, J., Acinas, S. G., González, J. M., et al. (2013). Ecology of marine Bacteroidetes: a comparative genomics approach. *ISME J* 7, 1026–1037. doi:10.1038/ismej.2012.169.
- Field, null, Behrenfeld, null, Randerson, null, and Falkowski, null (1998). Primary production of the biosphere: integrating terrestrial and oceanic components. *Science* 281, 237–240.
- Fierer, N., and Jackson, R. B. (2006). The diversity and biogeography of soil bacterial communities. *Proc. Natl. Acad. Sci. U.S.A.* 103, 626–631. doi:10.1073/pnas.0507535103.
- Figueras, M. J., Beaz-Hidalgo, R., Hossain, M. J., and Liles, M. R. (2014). Taxonomic Affiliation of New Genomes Should Be Verified Using Average Nucleotide Identity and Multilocus Phylogenetic Analysis. *Genome Announcements* 2. doi:10.1128/genomeA.00927-14.
- Flombaum, P., Gallegos, J. L., Gordillo, R. A., Rincón, J., Zabala, L. L., Jiao, N., et al. (2013). Present and future global distributions of the marine Cyanobacteria *Prochlorococcus* and *Synechococcus*. *PNAS* 110, 9824–9829. doi:10.1073/pnas.1307701110.
- Forward, J. A., Behrendt, M. C., Wyborn, N. R., Cross, R., and Kelly, D. J. (1997). TRAP transporters: a new family of periplasmic solute transport systems encoded by the *dctPQM* genes of *Rhodobacter capsulatus* and by homologs in diverse gram-negative bacteria. *Journal of Bacteriology* 179, 5482–5493. doi:10.1128/jb.179.17.5482-5493.1997.
- Franks, P. (1992). Sink or swim, accumulation of biomass at fronts. *Marine Ecology Progress Series* 82, 1–12. doi:10.3354/meps082001.
- Frère, L. (2017). Les microplastiques: une menace en rade de Brest? Available at: <http://www.theses.fr/2017BRES0046> [Accessed March 22, 2019].
- Frère, L., Maignien, L., Chalopin, M., Huvet, A., Rinnert, E., Morrison, H., et al. (2018). Microplastic bacterial communities in the Bay of Brest: Influence of polymer type and size. *Environmental Pollution* 242, 614–625. doi:10.1016/j.envpol.2018.07.023.
- Frère, L., Paul-Pont, I., Rinnert, E., Petton, S., Jaffré, J., Bihannic, I., et al. (2017). Influence of environmental and anthropogenic factors on the composition, concentration and spatial distribution of microplastics: A case study of the Bay of Brest (Brittany, France). *Environmental Pollution* 225, 211–222. doi:10.1016/j.envpol.2017.03.023.

-
- Fuhrman, J. A. (2009). Microbial community structure and its functional implications. *Nature* 459, 193–199. doi:10.1038/nature08058.
- Fuhrman, J. A., and Caron, D. A. (2016). Heterotrophic Planktonic Microbes: Virus, Bacteria, Archaea, and Protozoa. *Manual of Environmental Microbiology, Fourth Edition*, 4.2.2-1-4.2.2-34. doi:10.1128/9781555818821.ch4.2.2.
- Fuhrman, J. A., Cram, J. A., and Needham, D. M. (2015). Marine microbial community dynamics and their ecological interpretation. *Nature Reviews Microbiology* 13, 133–146. doi:10.1038/nrmicro3417.
- Fuhrman, J. A., McCallum, K., and Davis, A. A. (1992). Novel major archaeobacterial group from marine plankton. *Nature* 356, 148.
- Fuhrman, J. A., Steele, J. A., Hewson, I., Schwalbach, M. S., Brown, M. V., Green, J. L., et al. (2008). A latitudinal diversity gradient in planktonic marine bacteria. *Proceedings of the National Academy of Sciences* 105, 7774–7778. doi:10.1073/pnas.0803070105.
- Galand, P. E., Casamayor, E. O., Kirchman, D. L., and Lovejoy, C. (2009). Ecology of the rare microbial biosphere of the Arctic Ocean. *Proc Natl Acad Sci U S A* 106, 22427–22432. doi:10.1073/pnas.0908284106.
- Galand, P. E., Pereira, O., Hochart, C., Auguet, J. C., and Debroas, D. (2018). A strong link between marine microbial community composition and function challenges the idea of functional redundancy. *The ISME Journal*. doi:10.1038/s41396-018-0158-1.
- Galperin, M. Y., Makarova, K. S., Wolf, Y. I., and Koonin, E. V. (2015). Expanded microbial genome coverage and improved protein family annotation in the COG database. *Nucleic Acids Res* 43, D261–D269. doi:10.1093/nar/gku1223.
- Garneau, M.-È., Vincent, W. F., Terrado, R., and Lovejoy, C. (2009). Importance of particle-associated bacterial heterotrophy in a coastal Arctic ecosystem. *Journal of Marine Systems* 75, 185–197. doi:10.1016/j.jmarsys.2008.09.002.
- Gasol, J. M., Cardelús, C., Morán, X. A. G., Balagué, V., Forn, I., Marrasé, C., et al. (2016). Seasonal patterns in phytoplankton photosynthetic parameters and primary production at a coastal NW Mediterranean site. *Scientia Marina* 80, 63–77. doi:10.3989/scimar.04480.06E.
- Ghiglione, J. F., Palacios, C., Marty, J. C., Mével, G., Labrune, C., Conan, P., et al. (2008). Role of environmental factors for the vertical distribution (0–1000 m) of marine bacterial communities in the NW Mediterranean Sea. 35.

REFERENCES

- Ghiglione, J.-F., Galand, P. E., Pommier, T., Pedros-Alio, C., Maas, E. W., Bakker, K., et al. (2012). Pole-to-pole biogeography of surface and deep marine bacterial communities. *Proceedings of the National Academy of Sciences* 109, 17633–17638. doi:10.1073/pnas.1208160109.
- Giebel, H.-A., Kalhoefer, D., Lemke, A., Thole, S., Gahl-Janssen, R., Simon, M., et al. (2011). Distribution of *Roseobacter* RCA and SAR11 lineages in the North Sea and characteristics of an abundant RCA isolate. *The ISME Journal* 5, 8–19. doi:10.1038/ismej.2010.87.
- Gilbert, J. A., Steele, J. A., Caporaso, J. G., Steinbrück, L., Reeder, J., Temperton, B., et al. (2012). Defining seasonal marine microbial community dynamics. *The ISME Journal* 6, 298–308. doi:10.1038/ismej.2011.107.
- Giovannoni, S. J., Bibbs, L., Cho, J.-C., Stapels, M. D., Desiderio, R., Vergin, K. L., et al. (2005a). Proteorhodopsin in the ubiquitous marine bacterium SAR11. *Nature* 438, 82–85. doi:10.1038/nature04032.
- Giovannoni, S. J., Britschgi, T. B., Moyer, C. L., and Field, K. G. (1990). Genetic diversity in Sargasso Sea bacterioplankton. *Nature* 345, 60.
- Giovannoni, S. J., Cameron Thrash, J., and Temperton, B. (2014). Implications of streamlining theory for microbial ecology. *ISME J* 8, 1553–1565. doi:10.1038/ismej.2014.60.
- Giovannoni, S. J., and Stingl, U. (2005). Molecular diversity and ecology of microbial plankton. *Nature* 437, 343–348. doi:10.1038/nature04158.
- Giovannoni, S. J., Tripp, H. J., Givan, S., Podar, M., Vergin, K. L., Baptista, D., et al. (2005b). Genome Streamlining in a Cosmopolitan Oceanic Bacterium. *Science* 309, 5.
- Giovannoni, S. J., and Vergin, K. L. (2012). Seasonality in Ocean Microbial Communities. *Science* 335, 671–676. doi:10.1126/science.1198078.
- Glover, H. E., Campbell, L., and Prézelin, B. B. (1986). Contribution of *Synechococcus* spp. to size-fractionated primary productivity in three water masses in the Northwest Atlantic Ocean. *Mar. Biol.* 91, 193–203. doi:10.1007/BF00569435.
- Goberville, E., Beaugrand, G., Sautour, B., Tréguer, P., and Team, S. (2010). Climate-driven changes in coastal marine systems of western Europe. *Marine Ecology Progress Series* 408, 129–148.
- Gómez-Consarnau, L., González, J. M., Coll-Lladó, M., Gourdon, P., Pascher, T., Neutze, R., et al. (2007). Light stimulates growth of proteorhodopsin-containing marine Flavobacteria. *Nature* 445, 210–213. doi:10.1038/nature05381.

-
- Goris, J., Konstantinidis, K. T., Klappenbach, J. A., Coenye, T., Vandamme, P., and Tiedje, J. M. (2007). DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int. J. Syst. Evol. Microbiol.* 57, 81–91. doi:10.1099/ijs.0.64483-0.
- Graham, E. D., Heidelberg, J. F., and Tully, B. J. (2017). BinSanity: unsupervised clustering of environmental microbial assemblies using coverage and affinity propagation. *PeerJ* 5, e3035. doi:10.7717/peerj.3035.
- Gregg, W. W., and Rousseaux, C. S. (2016). Directional and Spectral Irradiance in Ocean Models: Effects on Simulated Global Phytoplankton, Nutrients, and Primary Production. *Front. Mar. Sci.* 3. doi:10.3389/fmars.2016.00240.
- GREPMA (1988). A Physical, Chemical and Biological Characterization of the Ushant Tidal Front. *Internationale Revue der gesamten Hydrobiologie und Hydrographie* 73, 511–536. doi:10.1002/iroh.19880730503.
- Grossart, H.-P., Levold, F., Allgaier, M., Simon, M., and Brinkhoff, T. (2005). Marine diatom species harbour distinct bacterial communities. *Environmental Microbiology* 7, 860–873.
- Grote, J., Thrash, J. C., Huggett, M. J., Landry, Z. C., Carini, P., Giovannoni, S. J., et al. (2012). Streamlining and Core Genome Conservation among Highly Divergent Members of the SAR11 Clade. *mBio* 3, e00252-12. doi:10.1128/mBio.00252-12.
- Haggerty, J. M., and Dinsdale, E. A. (2017). Distinct biogeographical patterns of marine bacterial taxonomy and functional genes. *Global Ecology and Biogeography* 26, 177–190. doi:10.1111/geb.12528.
- Hahnke, S., Brock, N. L., Zell, C., Simon, M., Dickschat, J. S., and Brinkhoff, T. (2013). Physiological diversity of Roseobacter clade bacteria co-occurring during a phytoplankton bloom in the North Sea. *Systematic and Applied Microbiology* 36, 39–48. doi:10.1016/j.syapm.2012.09.004.
- Hanada, S., and Sekiguchi, Y. (2014). The phylum gemmatimonadetes. *The Prokaryotes: Other Major Lineages of Bacteria and The Archaea*, 677–681.
- Harder, W. (1973). Microbial metabolism of organic C1 and C2 compounds. *Antonie van Leeuwenhoek* 39, 650–652.
- Haro-Moreno, J. M., Rodriguez-Valera, F., López-García, P., Moreira, D., and Martin-Cuadrado, A.-B. (2017). New insights into marine group III Euryarchaeota, from dark to light. *The ISME journal* 11, 1102.

REFERENCES

- Hatosy, S. M., Martiny, J. B. H., Sachdeva, R., Steele, J., Fuhrman, J. A., and Martiny, A. C. (2013). Beta diversity of marine bacteria depends on temporal scale. *Ecology* 94, 1898–1904. doi:10.1890/12-2125.1.
- Hausmann, B., Pelikan, C., Rattei, T., Loy, A., and Pester, M. (2018). Long-term transcriptional activity at zero growth by a cosmopolitan rare biosphere member. *bioRxiv*. doi:10.1101/284430.
- Hedges, J. I., Hatcher, P. G., Ertel, J. R., and Meyers-Schulte, K. J. (1992). A comparison of dissolved humic substances from seawater with Amazon River counterparts by ¹³C-NMR spectrometry. *Geochimica et Cosmochimica Acta* 56, 1753–1757.
- Hedges, S. B., Chen, H., Kumar, S., Wang, D. Y., Thompson, A. S., and Watanabe, H. (2001). A genomic timescale for the origin of eukaryotes. *BMC Evol Biol* 1, 4. doi:10.1186/1471-2148-1-4.
- Heinänen, A., Kononen, K., Kuosa, H., Kuparinen, J., and Mäkelä, K. (1995). Bacterioplankton growth associated with physical fronts during a cyanobacterial bloom. *Marine Ecology Progress Series*, 233–245.
- Hellweger, F. L. (2018). Heterotrophic substrate specificity in the aquatic environment: The role of microscale patchiness investigated using modelling. *Environmental microbiology* 20, 3825–3835.
- Herlemann, D. P., Labrenz, M., Jürgens, K., Bertilsson, S., Waniek, J. J., and Andersson, A. F. (2011). Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *ISME J* 5, 1571–1579. doi:10.1038/ismej.2011.41.
- Hertkorn, N., Harir, M., Koch, B. P., Michalke, B., and Schmitt-Kopplin, P. (2013). High-field NMR spectroscopy and FTICR mass spectrometry: powerful discovery tools for the molecular level characterization of marine dissolved organic matter. *Biogeosciences* 10, 1583–1624. doi:https://doi.org/10.5194/bg-10-1583-2013.
- Hewson, I., Steele, J. A., Capone, D. G., and Fuhrman, J. A. (2006). Temporal and spatial scales of variation in bacterioplankton assemblages of oligotrophic surface waters. *Marine Ecology Progress Series* 311, 67–77. doi:10.3354/meps311067.
- Ho, A., Di Lonardo, D. P., and Bodelier, P. L. E. (2017). Revisiting life strategy concepts in environmental microbial ecology. *FEMS Microbiol Ecol* 93. doi:10.1093/femsec/fix006.
- Hobbie, J. E., Daley, R. J., and Jasper, S. (1977). Use of nuclepore filters for counting bacteria by fluorescence microscopy. *Appl. Environ. Microbiol.* 33, 1225–1228.
- Holm-Hansen, O., Lorenzen, C. J., Holmes, R. W., and Strickland, J. D. (1965). Fluorometric determination of chlorophyll. *ICES Journal of Marine Science* 30, 3–15.

-
- Hug, L. A., Baker, B. J., Anantharaman, K., Brown, C. T., Probst, A. J., Castelle, C. J., et al. (2016). A new view of the tree of life. *Nature Microbiology* 1, nmicrobiol201648. doi:10.1038/nmicrobiol.2016.48.
- Hugerth, L. W., Larsson, J., Alneberg, J., Lindh, M. V., Legrand, C., Pinhassi, J., et al. (2015). Metagenome-assembled genomes uncover a global brackish microbiome. *Genome Biology* 16, 279. doi:10.1186/s13059-015-0834-7.
- Huggett, M. J., Hayakawa, D. H., and Rappé, M. S. (2012). Genome sequence of strain HIMB624, a cultured representative from the OM43 clade of marine Betaproteobacteria. *Standards in Genomic Sciences* 6, 11. doi:10.4056/sigs.2305090.
- Hugoni, M., Taib, N., Debroas, D., Domaizon, I., Jouan Dufournel, I., Bronner, G., et al. (2013). Structure of the rare archaeal biosphere and seasonal dynamics of active ecotypes in surface coastal waters. *Proceedings of the National Academy of Sciences* 110, 6004–6009. doi:10.1073/pnas.1216863110.
- Hunt, D. E., Lin, Y., Church, M. J., Karl, D. M., Tringe, S. G., Izzo, L. K., et al. (2013). Relationship between Abundance and Specific Activity of Bacterioplankton in Open Ocean Surface Waters. *Appl. Environ. Microbiol.* 79, 177–184. doi:10.1128/AEM.02155-12.
- Hyatt, D., Chen, G.-L., LoCasio, P. F., Land, M. L., Larimer, F. W., and Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11, 119. doi:10.1186/1471-2105-11-119.
- Iriberry, J., Unanue, M., Barcina, I., and Egea, L. (1987). Seasonal Variation in Population Density and Heterotrophic Activity of Attached and Free-Living Bacteria in Coastal Waters. *APPL. ENVIRON. MICROBIOL.* 53, 7.
- Iverson, V., Morris, R. M., Frazar, C. D., Berthiaume, C. T., Morales, R. L., and Armbrust, E. V. (2012). Untangling Genomes from Metagenomes: Revealing an Uncultured Class of Marine Euryarchaeota. *Science* 335, 587–590. doi:10.1126/science.1212665.
- Jiao, N., Herndl, G. J., Hansell, D. A., Benner, R., Kattner, G., Wilhelm, S. W., et al. (2010). Microbial production of recalcitrant dissolved organic matter: long-term carbon storage in the global ocean. *Nature Reviews Microbiology* 8, 593.
- Jing, H., Xia, X., Suzuki, K., and Liu, H. (2013). Vertical Profiles of Bacteria in the Tropical and Subarctic Oceans Revealed by Pyrosequencing. *PLOS ONE* 8, e79423. doi:10.1371/journal.pone.0079423.
- Jones, S. E., and Lennon, J. T. (2010). Dormancy contributes to the maintenance of microbial diversity. *PNAS* 107, 5881–5886. doi:10.1073/pnas.0912765107.

REFERENCES

- Kabisch, A., Otto, A., König, S., Becher, D., Albrecht, D., Schüler, M., et al. (2014). Functional characterization of polysaccharide utilization loci in the marine *Bacteroidetes* 'Gramella forsetii' KT0803. *The ISME Journal* 8, 1492–1502. doi:10.1038/ismej.2014.4.
- Kan, J., Suzuki, M. T., Wang, K., Evans, S. E., and Chen, F. (2007). High Temporal but Low Spatial Heterogeneity of Bacterioplankton in the Chesapeake Bay. *Applied and Environmental Microbiology* 73, 6776–6789. doi:10.1128/AEM.00541-07.
- Kanehisa, M., Sato, Y., and Morishima, K. (2016). BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. *Journal of Molecular Biology* 428, 726–731. doi:10.1016/j.jmb.2015.11.006.
- Kang, D. D., Froula, J., Egan, R., and Wang, Z. (2015). MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ* 3, e1165. doi:10.7717/peerj.1165.
- Karl, D. M. (2007). Microbial oceanography: paradigms, processes and promise. *Nature Reviews Microbiology* 5, 759–769. doi:10.1038/nrmicro1749.
- Karner, M. B., DeLong, E. F., and Karl, D. M. (2001). Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature* 409, 507–510. doi:10.1038/35054051.
- Katz, M. E., Finkel, Z. V., Grzebyk, D., Knoll, A. H., and Falkowski, P. G. (2004). Evolutionary Trajectories and Biogeochemical Impacts of Marine Eukaryotic Phytoplankton. *Annu. Rev. Ecol. Evol. Syst.* 35, 523–556. doi:10.1146/annurev.ecolsys.35.112202.130137.
- Kennish, M. J. (1991). *Ecology of estuaries: anthropogenic effects*. CRC press.
- Kiene, R. P., Linn, L. J., and Bruton, J. A. (2000). New and important roles for DMSP in marine microbial communities. *Journal of Sea Research* 43, 209–224.
- Kim, D., Song, L., Breitwieser, F. P., and Salzberg, S. L. (2016). Centrifuge: rapid and sensitive classification of metagenomic sequences. *Genome Res.* 26, 1721–1729. doi:10.1101/gr.210641.116.
- Kirchman, D. L. (2016). Growth Rates of Microbes in the Oceans. *Annual Review of Marine Science* 8, 285–309. doi:10.1146/annurev-marine-122414-033938.
- Kits, K. D., Sedlacek, C. J., Lebedeva, E. V., Han, P., Bulaev, A., Pjevac, P., et al. (2017). Kinetic analysis of a complete nitrifier reveals an oligotrophic lifestyle. *Nature* 549, 269.
- Koch, A. L. (2001). Oligotrophs versus copiotrophs. *BioEssays* 23, 657–661. doi:10.1002/bies.1091.

-
- Koch, B. P., Ludwichowski, K.-U., Kattner, G., Dittmar, T., and Witt, M. (2008). Advanced characterization of marine dissolved organic matter by combining reversed-phase liquid chromatography and FT-ICR-MS. *Marine Chemistry* 111, 233–241. doi:10.1016/j.marchem.2008.05.008.
- Könneke, M., Bernhard, A. E., José, R., Walker, C. B., Waterbury, J. B., and Stahl, D. A. (2005). Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* 437, 543.
- Kopf, A., Kostadinov, I., Wichels, A., Quast, C., and Glöckner, F. O. (2015). Metatranscriptome of marine bacterioplankton during winter time in the North Sea assessed by total RNA sequencing. *Marine Genomics* 19, 45–46. doi:10.1016/j.margen.2014.11.001.
- Krause, S., Le Roux, X., Niklaus, P. A., Van Bodegom, P. M., Lennon, J. T., Bertilsson, S., et al. (2014). Trait-based approaches for understanding microbial biodiversity and ecosystem functioning. *Front. Microbiol.* 5. doi:10.3389/fmicb.2014.00251.
- Kumar, S., Stecher, G., Li, M., Nnyaz, C., and Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol* 35, 1547–1549. doi:10.1093/molbev/msy096.
- Kurtz, Z. D., Müller, C. L., Miraldi, E. R., Littman, D. R., Blaser, M. J., and Bonneau, R. A. (2015). Sparse and compositionally robust inference of microbial ecological networks. *PLoS computational biology* 11, e1004226.
- Kuypers, M. M., Marchant, H. K., and Kartal, B. (2018). The microbial nitrogen-cycling network. *Nature Reviews Microbiology* 16, 263.
- Lambert, S., Tragin, M., Lozano, J.-C., Ghiglione, J.-F., Vaulot, D., Bouget, F.-Y., et al. (2018). Rhythmicity of coastal marine picoeukaryotes, bacteria and archaea despite irregular environmental perturbations. *The ISME Journal*. doi:10.1038/s41396-018-0281-z.
- Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9, 559. doi:10.1186/1471-2105-9-559.
- Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi:10.1038/nmeth.1923.
- Lauro, F. M., McDougald, D., Thomas, T., Williams, T. J., Egan, S., Rice, S., et al. (2009). The genomic basis of trophic strategy in marine bacteria. *PNAS* 106, 15527–15533. doi:10.1073/pnas.0903507106.

REFERENCES

- Le, B., Cambon, G., Daniault, N., Herbette, S., Le, C., Marié, L., et al. (2009). Observations of the Ushant tidal front in September 2007. *Continental Shelf Research* 29, 1026–1037. doi:10.1016/j.csr.2008.12.020.
- Le Boyer, A., Cambon, G., Daniault, N., Herbette, S., Le Cann, B., Marié, L., et al. (2009). Observations of the Ushant tidal front in September 2007. *Continental Shelf Research* 29, 1026–1037. doi:10.1016/j.csr.2008.12.020.
- Le Fèvre, J. (1986). Aspect of the biology of frontal systems. *Advances in Marine Biology*.
- Le Fèvre, J. (1987). "Aspects of the Biology of Frontal Systems," in *Advances in Marine Biology*, eds. J. H. S. Blaxter and A. J. Southward (Academic Press), 163–299. doi:10.1016/S0065-2881(08)60109-1.
- Le Fèvre, J., Le Corre, P., Morin, P., and Birrien, J.-L. (1983). The pelagic ecosystem in frontal zones and other environments off the West coast of Brittany. *Oceanologica Acta*, 125–129.
- Le Pape, O., Del Amo, Y., Menesguen, A., Aminot, A., Quequiner, B., and Treguer, P. (1996). Resistance of a coastal ecosystem to increasing eutrophic conditions: the Bay of Brest (France), a semi-enclosed zone of Western Europe. *Continental Shelf Research* 16, 1885–1907. doi:10.1016/0278-4343(95)00068-2.
- Lee, K. S., Palatinszky, M., Pereira, F. C., Nguyen, J., Fernandez, V. I., Mueller, A. J., et al. (2019). An automated Raman-based platform for the sorting of live cells by functional properties. *Nature Microbiology*, 1. doi:10.1038/s41564-019-0394-9.
- Legendre, L., and Demers, S. (1984). Towards Dynamic Biological Oceanography and Limnology. *Can. J. Fish. Aquat. Sci.* 41, 2–19. doi:10.1139/f84-001.
- Legendre, L., and Le Fèvre, J. (1995). Microbial food webs and the export of biogenic carbon in oceans. *Aquatic Microbial Ecology* 9, 69–77. doi:10.3354/ame009069.
- Lennon, J. T., and Jones, S. E. (2011). Microbial seed banks: the ecological and evolutionary implications of dormancy. *Nature Reviews Microbiology* 9, 119–130. doi:10.1038/nrmicro2504.
- Lévy, M., Jahn, O., Dutkiewicz, S., Follows, M. J., and d'Ovidio, F. (2015). The dynamical landscape of marine phytoplankton diversity. *J R Soc Interface* 12, 20150481–20150481. doi:10.1098/rsif.2015.0481.
- LHelguen, S., Slawyk, G., and Le Corre, P. (2005). Seasonal patterns of urea regeneration by size-fractionated microheterotrophs in well-mixed temperate coastal waters. *Journal of Plankton Research*, 263–270.

-
- Li, W., and Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22, 1658–1659. doi:10.1093/bioinformatics/btl158.
- Li, W. K., Glen Harrison, W., and Head, E. J. (2006). Coherent assembly of phytoplankton communities in diverse temperate ocean ecosystems. *Proceedings of the Royal Society B: Biological Sciences* 273, 1953–1960.
- Li, Y., Sun, L.-L., Sun, M.-L., Su, H.-N., Zhang, X.-Y., Xie, B.-B., et al. (2018). Vertical and horizontal biogeographic patterns and major factors affecting bacterial communities in the open South China Sea. *Scientific Reports* 8, 8800. doi:10.1038/s41598-018-27191-w.
- Lidbury, I. D., Murrell, J. C., and Chen, Y. (2015). Trimethylamine and trimethylamine N-oxide are supplementary energy sources for a marine heterotrophic bacterium: implications for marine carbon and nitrogen cycling. *ISME J* 9, 760–769. doi:10.1038/ismej.2014.149.
- Lidbury, I., Murrell, J. C., and Chen, Y. (2014). Trimethylamine N-oxide metabolism by abundant marine heterotrophic bacteria. *Proceedings of the National Academy of Sciences* 111, 2710–2715. doi:10.1073/pnas.1317834111.
- Liénart, C., Savoye, N., Bozec, Y., Breton, E., Conan, P., David, V., et al. (2017). Dynamics of particulate organic matter composition in coastal systems: A spatio-temporal study at multi-systems scale. *Progress in Oceanography* 156, 221–239. doi:10.1016/j.pocean.2017.03.001.
- Lima-Mendez, G., Faust, K., Henry, N., Decelle, J., Colin, S., Carcillo, F., et al. (2015). Determinants of community structure in the global plankton interactome. *Science* 348, 1262073.
- Lindh, M. V., Lefébure, R., Degerman, R., Lundin, D., Andersson, A., and Pinhassi, J. (2015a). Consequences of increased terrestrial dissolved organic matter and temperature on bacterioplankton community composition during a Baltic Sea mesocosm experiment. *AMBIO* 44, 402–412. doi:10.1007/s13280-015-0659-3.
- Lindh, M. V., Sjöstedt, J., Andersson, A. F., Baltar, F., Hugerth, L. W., Lundin, D., et al. (2015b). Disentangling seasonal bacterioplankton population dynamics by high-frequency sampling. *Environmental microbiology* 17, 2459–2476.
- Loman, N., Goodwin, S., Jansen, H., and Loose, M. (2015). A disruptive sequencer meets disruptive publishing. *F1000Research*. doi:10.12688/f1000research.7229.1.
- Lomas, M. W., Glibert, P. M., Shiah, F.-K., and Smith, E. M. (2002). Microbial processes and temperature in Chesapeake Bay: current relationships and potential impacts of regional warming. *Global Change Biology* 8, 51–70. doi:10.1046/j.1365-2486.2002.00454.x.

REFERENCES

- Lomas, M. W., and Lipschultz, F. (2006). Forming the primary nitrite maximum: Nitrifiers or phytoplankton? *Limnology and Oceanography* 51, 2453–2467. doi:10.4319/lo.2006.51.5.2453.
- Louca, S., Parfrey, L. W., and Doebeli, M. (2016). Decoupling function and taxonomy in the global ocean microbiome. *Science* 353, 1272–1277. doi:10.1126/science.aaf4507.
- Lucas, J., Wichels, A., Teeling, H., Chafee, M., Scharfe, M., and Gerdts, G. (2015). Annual dynamics of North Sea bacterioplankton: seasonal variability superimposes short-term variation. *FEMS Microbiol Ecol* 91. doi:10.1093/femsec/fiv099.
- Lund, J. W. G., Kipling, C., and Le Cren, E. D. (1958). The inverted microscope method of estimating algal numbers and the statistical basis of estimations by counting. *Hydrobiologia* 11, 143–170.
- Luo, H., Csúros, M., Hughes, A. L., and Moran, M. A. (2013). Evolution of Divergent Life History Strategies in Marine Alphaproteobacteria. *mBio* 4, e00373-13. doi:10.1128/mBio.00373-13.
- Luo, H., and Moran, M. A. (2014). Evolutionary Ecology of the Marine Roseobacter Clade. *Microbiology and Molecular Biology Reviews* 78, 573–587. doi:10.1128/MMBR.00020-14.
- Luo, H., Swan, B. K., Stepanauskas, R., Hughes, A. L., and Moran, M. A. (2014). Evolutionary analysis of a streamlined lineage of surface ocean Roseobacters. *ISME J* 8, 1428–1439. doi:10.1038/ismej.2013.248.
- Luria, C. M., Amaral-Zettler, L. A., Ducklow, H. W., Repeta, D. J., Rhyne, A. L., and Rich, J. J. (2017). Seasonal Shifts in Bacterial Community Responses to Phytoplankton-Derived Dissolved Organic Matter in the Western Antarctic Peninsula. *Front. Microbiol.* 8. doi:10.3389/fmicb.2017.02117.
- MacLeod, R. A. (1965). The Question of the Existence of Specific Marine Bacteria¹. *Bacteriol Rev* 29, 9–23. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC441258/> [Accessed March 20, 2019].
- Madsen, E. L. (2011). Microorganisms and their roles in fundamental biogeochemical cycles. *Current Opinion in Biotechnology* 22, 456–464. doi:10.1016/j.copbio.2011.01.008.
- Mahé, F., Rognes, T., Quince, C., de Vargas, C., and Dunthorn, M. (2014). Swarm: robust and fast clustering method for amplicon-based studies. *PeerJ* 2, e593.
- Mahé, F., Rognes, T., Quince, C., de Vargas, C., and Dunthorn, M. (2015). Swarm v2: highly-scalable and high-resolution amplicon clustering. *PeerJ* 3, e1420. doi:10.7717/peerj.1420.
- Mariette, V., and Le Cann, B. (1985). Simulation of the formation of Ushant thermal front. *Continental Shelf Research* 4, 637–660. doi:10.1016/0278-4343(85)90034-2.

-
- Martinez-Garcia, M., Brazel, D. M., Swan, B. K., Arnosti, C., Chain, P. S. G., Reitenga, K. G., et al. (2012). Capturing Single Cell Genomes of Active Polysaccharide Degradors: An Unexpected Contribution of Verrucomicrobia. *PLOS ONE* 7, e35314. doi:10.1371/journal.pone.0035314.
- Martin-Platero, A. M., Cleary, B., Kauffman, K., Preheim, S. P., McGillicuddy, D. J., Alm, E. J., et al. (2018). High resolution time series reveals cohesive but short-lived communities in coastal plankton. *Nature Communications* 9, 266. doi:10.1038/s41467-017-02571-4.
- Matthews, H. D., Gillett, N. P., Stott, P. A., and Zickfeld, K. (2009). The proportionality of global warming to cumulative carbon emissions. *Nature* 459, 829–832. doi:10.1038/nature08047.
- Mayali, X., Weber, P. K., Brodie, E. L., Mabery, S., Hoepflich, P. D., and Pett-Ridge, J. (2012). High-throughput isotopic analysis of RNA microarrays to quantify microbial resource use. *ISME J* 6, 1210–1221. doi:10.1038/ismej.2011.175.
- Mayali, X., Weber, P. K., Mabery, S., and Pett-Ridge, J. (2014). Phylogenetic Patterns in the Microbial Response to Resource Availability: Amino Acid Incorporation in San Francisco Bay. *PLOS ONE* 9, e95842. doi:10.1371/journal.pone.0095842.
- McFall-Ngai, M., Hadfield, M. G., Bosch, T. C. G., Carey, H. V., Domazet-Lošo, T., Douglas, A. E., et al. (2013). Animals in a bacterial world, a new imperative for the life sciences. *Proc. Natl. Acad. Sci. U.S.A.* 110, 3229–3236. doi:10.1073/pnas.1218525110.
- Mestre, M., Borrull, E., Montserrat, S., and Gasol, J. M. (2017). Patterns of bacterial diversity in the marine planktonic particulate matter continuum. *The ISME Journal*, 999–1010.
- Middelboe, M., Søndergaard, M., Letarte, Y., and Borch, N. H. (1995). Attached and free-living bacteria: Production and polymer hydrolysis during a diatom bloom. *Microb Ecol* 29, 231–248. doi:10.1007/BF00164887.
- Minoche, A. E., Dohm, J. C., and Himmelbauer, H. (2011). Evaluation of genomic high-throughput sequencing data generated on Illumina HiSeq and Genome Analyzer systems. *Genome Biology* 12, R112. doi:10.1186/gb-2011-12-11-r112.
- Mistry, J., Finn, R. D., Eddy, S. R., Bateman, A., and Punta, M. (2013). Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. *Nucleic Acids Res* 41, e121–e121. doi:10.1093/nar/gkt263.
- Mitra, A., Flynn, K. J., Burkholder, J. M., Berge, T., Calbet, A., Raven, J. A., et al. (2014). The role of mixotrophic protists in the biological carbon pump. *Biogeosciences* 11, 995–1005. doi:10.5194/bg-11-995-2014.

REFERENCES

- Moisander, P. H., Benavides, M., Bonnet, S., Berman-Frank, I., White, A. E., and Riemann, L. (2017). Chasing after non-cyanobacterial nitrogen fixation in marine pelagic environments. *Frontiers in microbiology* 8, 1736.
- Moran, M. A., Belas, R., Schell, M. A., González, J. M., Sun, F., Sun, S., et al. (2007). Ecological Genomics of Marine Roseobacters. *Appl. Environ. Microbiol.* 73, 4559–4569. doi:10.1128/AEM.02580-06.
- Moran, M. A., Buchan, A., González, J. M., Heidelberg, J. F., Whitman, W. B., Kiene, R. P., et al. (2004). Genome sequence of *Silicibacter pomeroyi* reveals adaptations to the marine environment. *Nature* 432, 910.
- Morris, J. J., Lenski, R. E., and Zinser, E. R. (2012a). The Black Queen Hypothesis: Evolution of Dependencies through Adaptive Gene Loss. *mBio* 3. doi:10.1128/mBio.00036-12.
- Morris, R. M., Frazar, C. D., and Carlson, C. A. (2012b). Basin-scale patterns in the abundance of SAR11 subclades, marine Actinobacteria (OM1), members of the Roseobacter clade and OCS116 in the South Atlantic. *Environmental Microbiology* 14, 1133–1144. doi:10.1111/j.1462-2920.2011.02694.x.
- Morris, R. M., Rappé, M. S., Connon, S. A., Vergin, K. L., Siebold, W. A., Carlson, C. A., et al. (2002). SAR11 clade dominates ocean surface bacterioplankton communities. *Nature* 420, 806.
- Mou, X., Sun, S., Edwards, R. A., Hodson, R. E., and Moran, M. A. (2008). Bacterial carbon processing by generalist species in the coastal ocean. *Nature* 451, 708–711. doi:10.1038/nature06513.
- Muller, H., Dumas, F., Blanke, B., and Mariette, V. (2007). High-resolution atmospheric forcing for regional oceanic model: the Iroise Sea. *Ocean Dynamics* 57, 375–400. doi:10.1007/s10236-007-0115-4.
- Musat, N., Halm, H., Winterholler, B., Hoppe, P., Peduzzi, S., Hillion, F., et al. (2008). A single-cell view on the ecophysiology of anaerobic phototrophic bacteria. *PNAS* 105, 17861–17866. doi:10.1073/pnas.0809329105.
- Naushad, S., Adeolu, M., Wong, S., Sohail, M., Schellhorn, H. E., and Gupta, R. S. (2015). A phylogenomic and molecular marker based taxonomic framework for the order Xanthomonadales: proposal to transfer the families Algiphilaceae and Solimonadaceae to the order Nevskiales ord. nov. and to create a new family within the order Xanthomonadales, the family Rhodanobacteraceae fam. nov., containing the genus *Rhodanobacter* and its closest relatives. *Antonie Van Leeuwenhoek* 107, 467–485. doi:10.1007/s10482-014-0344-8.
- Nebel, M., Pfabel, C., Stock, A., Dunthorn, M., and Stoeck, T. (2011). Delimiting operational taxonomic units for assessing ciliate environmental diversity using small-subunit rRNA gene sequences. *Environ Microbiol Rep* 3, 154–158. doi:10.1111/j.1758-2229.2010.00200.x.

-
- Needham, D. M., Chow, C.-E. T., Cram, J. A., Sachdeva, R., Parada, A., and Fuhrman, J. A. (2013). Short-term observations of marine bacterial and viral communities: patterns, connections and resilience. *The ISME Journal* 7, 1274–1285. doi:10.1038/ismej.2013.19.
- Needham, D. M., Fichot, E. B., Wang, E., Berdjeb, L., Cram, J. A., Fichot, C. G., et al. (2018). Dynamics and interactions of highly resolved marine plankton via automated high-frequency sampling. *ISME J* 12, 2417–2432. doi:10.1038/s41396-018-0169-y.
- Needham, D. M., and Fuhrman, J. A. (2016). Pronounced daily succession of phytoplankton, archaea and bacteria following a spring bloom. *Nature Microbiology* 1, 16005. doi:10.1038/nmicrobiol.2016.5.
- Needham, D. M., Sachdeva, R., and Fuhrman, J. A. (2017). Ecological dynamics and co-occurrence among marine phytoplankton, bacteria and myoviruses shows microdiversity matters. *ISME J* 11, 1614–1629. doi:10.1038/ismej.2017.29.
- Nelson, M. C., Morrison, H. G., Benjamino, J., Grim, S. L., and Graf, J. (2014). Analysis, Optimization and Verification of Illumina-Generated 16S rRNA Gene Amplicon Surveys. *PLOS ONE* 9, e94249. doi:10.1371/journal.pone.0094249.
- Newton, R. J., Griffin, L. E., Bowles, K. M., Meile, C., Gifford, S., Givens, C. E., et al. (2010). Genome characteristics of a generalist marine bacterial lineage. *The ISME Journal* 4, 784–798. doi:10.1038/ismej.2009.150.
- Newton, R., and Shade, A. (2016). Lifestyles of rarity: understanding heterotrophic strategies to inform the ecology of the microbial rare biosphere. *Aquatic Microbial Ecology* 78, 51–63. doi:10.3354/ame01801.
- Nielsen, H. B., Almeida, M., Juncker, A. S., Rasmussen, S., Li, J., Sunagawa, S., et al. (2014). Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. *Nature Biotechnology* 32, 822–828. doi:10.1038/nbt.2939.
- Nixon, S. W., and Buckley, B. A. (2002). “A strikingly rich zone”—nutrient enrichment and secondary production in coastal marine ecosystems. *Estuaries* 25, 782–796.
- Oh, H.-M., Kwon, K. K., Kang, I., Kang, S. G., Lee, J.-H., Kim, S.-J., et al. (2010). Complete Genome Sequence of “*Candidatus Puniceispirillum marinum*” IMCC1322, a Representative of the SAR116 Clade in the Alphaproteobacteria. *Journal of Bacteriology* 192, 3240–3241. doi:10.1128/JB.00347-10.

REFERENCES

- Olm, M. R., Brown, C. T., Brooks, B., and Banfield, J. F. (2017). dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J* 11, 2864–2868. doi:10.1038/ismej.2017.126.
- Olson, D. B., and Backus, R. H. (1985). The concentrating of organisms at fronts: A cold-water fish and a warm-core Gulf Stream ring. *Journal of Marine Research*, 113–137.
- Olson, R. J., Chisholm, S. W., Zettler, E. R., Altabet, M. A., and Dusenberry, J. A. (1990). Spatial and temporal distributions of prochlorophyte picoplankton in the North Atlantic Ocean. *Deep Sea Research Part A. Oceanographic Research Papers* 37, 1033–1051. doi:10.1016/0198-0149(90)90109-9.
- Ouverney, C. C., and Fuhrman, J. A. (2000). Marine planktonic archaea take up amino acids. *Applied and Environmental Microbiology* 66, 4829–4833. Available at: <http://aem.asm.org/content/66/11/4829.short> [Accessed February 17, 2015].
- Pace, N. R., Stahl, D. A., Lane, D. J., and Olsen, G. J. (1986). "The Analysis of Natural Microbial Populations by Ribosomal RNA Sequences," in *Advances in Microbial Ecology Advances in Microbial Ecology.*, ed. K. C. Marshall (Boston, MA: Springer US), 1–55. doi:10.1007/978-1-4757-0611-6_1.
- Paillard, C., Jean, F., Ford, S. E., Powell, E. N., Klinck, J. M., Hofmann, E. E., et al. (2014). A theoretical individual-based model of Brown Ring Disease in Manila clams, *Venerupis philippinarum*. *Journal of Sea Research* 91, 15–34. doi:10.1016/j.seares.2014.03.005.
- Parada, A. E., and Fuhrman, J. A. (2017). Marine archaeal dynamics and interactions with the microbial community over 5 years from surface to seafloor. *ISME J* 11, 2510–2525. doi:10.1038/ismej.2017.104.
- Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., and Tyson, G. W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Research* 25, 1043–1055. doi:10.1101/gr.186072.114.
- Pedler, B. E., Aluwihare, L. I., and Azam, F. (2014). Single bacterial strain capable of significant contribution to carbon cycling in the surface ocean. *Proceedings of the National Academy of Sciences* 111, 7202–7207. doi:10.1073/pnas.1401887111.
- Pedrós-Alió, C. (2006). Marine microbial diversity: can it be determined? *Trends Microbiol.* 14, 257–263. doi:10.1016/j.tim.2006.04.007.
- Peng, Y., Leung, H. C. M., Yiu, S. M., and Chin, F. Y. L. (2012). IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* 28, 1420–1428. doi:10.1093/bioinformatics/bts174.

-
- Pinhassi, J., DeLong, E. F., Bèjà, O., González, J. M., and Pedrós-Alió, C. (2016). Marine Bacterial and Archaeal Ion-Pumping Rhodopsins: Genetic Diversity, Physiology, and Ecology. *Microbiol Mol Biol Rev* 80, 929–954. doi:10.1128/MMBR.00003-16.
- Pinhassi, J., Sala, M. M., Havskum, H., Peters, F., Guadayol, O., Malits, A., et al. (2004). Changes in bacterioplankton composition under different phytoplankton regimens. *Appl. Environ. Microbiol.* 70, 6753–6766.
- Polz, M. F., Hunt, D. E., Preheim, S. P., and Weinreich, D. M. (2006). Patterns and mechanisms of genetic and phenotypic differentiation in marine microbes. *Philosophical Transactions of the Royal Society B: Biological Sciences* 361, 2009–2021. doi:10.1098/rstb.2006.1928.
- Pomeroy, L., leB. Williams, P., Azam, F., and Hobbie, J. (2007). The Microbial Loop. *Oceanography* 20, 28–33. doi:10.5670/oceanog.2007.45.
- Pomeroy, L. R. (1974). The Ocean's Food Web, A Changing Paradigm. *BioScience* 24, 499–504. doi:10.2307/1296885.
- Porter, J., Deere, D., Pickup, R., and Edwards, C. (1996). Fluorescent probes and flow cytometry: New insights into environmental bacteriology. *Cytometry* 23, 91–96. doi:10.1002/(SICI)1097-0320(19960201)23:2<91::AID-CYTO1>3.0.CO;2-O.
- Pujalte, M. J., Lucena, T., Ruvira, M. A., Arahal, D. R., and Macián, M. C. (2014). “The Family Rhodobacteraceae,” in *The Prokaryotes: Alphaproteobacteria and Betaproteobacteria*, eds. E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt, and F. Thompson (Berlin, Heidelberg: Springer Berlin Heidelberg), 439–512. doi:10.1007/978-3-642-30197-1_377.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2012). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic acids research* 41, D590–D596.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, D590-596. doi:10.1093/nar/gks1219.
- Quéguiner, B., and Tréguer, P. (1984). Studies on the Phytoplankton in the Bay of Brest (Western Europe). Seasonal Variations in Composition, Biomass and Production in Relation to Hydrological and Chemical Features (1981 — 1982). *Botanica Marina* 27. doi:10.1515/botm.1984.27.10.449.
- Quinn, P. K., and Bates, T. S. (2011). The case against climate regulation via oceanic phytoplankton sulphur emissions. *Nature* 480, 51.

REFERENCES

- Rabalais, N. N., Turner, R. E., and Scavia, D. (2002). Beyond Science into Policy: Gulf of Mexico Hypoxia and the Mississippi River Nutrient policy development for the Mississippi River watershed reflects the accumulated scientific evidence that the increase in nitrogen loading is the primary factor in the worsening of hypoxia in the northern Gulf of Mexico. *BioScience* 52, 129–142. doi:10.1641/0006-3568(2002)052[0129:BSIPGO]2.0.CO;2.
- Raes, J., Letunic, I., Yamada, T., Jensen, L. J., and Bork, P. (2011). Toward molecular trait-based ecology through integration of biogeochemical, geographical and metagenomic data. *Molecular Systems Biology* 7, 473. doi:10.1038/msb.2011.6.
- Ragueneau, O., Quéguiner, B., and Tréguer, P. (1996). Contrast in biological responses to tidally-induced vertical mixing for two macrotidal ecosystems of Western Europe. *Estuarine, Coastal and Shelf Science* 42, 645–665.
- Ragueneau, O., and Tréguer, P. (1994). Determination of biogenic silica in coastal waters: applicability and limits of the alkaline digestion method. *Marine Chemistry* 45, 43–51.
- Ramesh, R., Chen, Z., Cummins, V., Day, J., D’Elia, C., Dennison, B., et al. (2015). Land–Ocean Interactions in the Coastal Zone: Past, present & future. *Anthropocene* 12, 85–98. doi:10.1016/j.ancene.2016.01.005.
- Rao, D., Webb, J. S., Holmstrom, C., Case, R., Low, A., Steinberg, P., et al. (2007). Low Densities of Epiphytic Bacteria from the Marine Alga *Ulva australis* Inhibit Settlement of Fouling Organisms. *Applied and Environmental Microbiology* 73, 7844–7852. doi:10.1128/AEM.01543-07.
- Rappé, M. S., Connon, S. A., Vergin, K. L., and Giovannoni, S. J. (2002). Cultivation of the ubiquitous SAR11 marine bacterioplankton clade. *Nature* 418, 630–633. doi:10.1038/nature00917.
- Reeder, J., and Knight, R. (2009). The “rare biosphere”: a reality check. *Nature Methods*. Available at: <http://link.galegroup.com/apps/doc/A207705589/AONE?sid=googlescholar> [Accessed March 20, 2019].
- Reveillaud, J., Maignien, L., Murat Eren, A., Huber, J. A., Apprill, A., Sogin, M. L., et al. (2014). Host-specificity among abundant and rare taxa in the sponge microbiome. *ISME J* 8, 1198–1209. doi:10.1038/ismej.2013.227.
- Riebesell, U., Körtzinger, A., and Oeschlies, A. (2009). Sensitivities of marine carbon fluxes to ocean change. *Proceedings of the National Academy of Sciences* 106, 20602–20609.
- Rieck, A., Herlemann, D. P. R., Jürgens, K., and Grossart, H.-P. (2015). Particle-Associated Differ from Free-Living Bacteria in Surface Waters of the Baltic Sea. *Front. Microbiol.* 6. doi:10.3389/fmicb.2015.01297.

-
- Riemann, L., Steward, G. F., and Azam, F. (2000). Dynamics of Bacterial Community Composition and Activity during a Mesocosm Diatom Bloom. *Appl. Environ. Microbiol.* 66, 578–587. doi:10.1128/AEM.66.2.578-587.2000.
- Rinke, C., Schwientek, P., Sczyrba, A., Ivanova, N. N., Anderson, I. J., Cheng, J.-F., et al. (2013). Insights into the phylogeny and coding potential of microbial dark matter. *Nature* 499, 431–437. doi:10.1038/nature12352.
- Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 4, e2584. doi:10.7717/peerj.2584.
- Rosenberg, R., and Loo, L.-O. (1988). Marine eutrophication induced oxygen deficiency: Effects on soft bottom Fauna, Western Sweden. *Ophelia* 29, 213–225. doi:10.1080/00785326.1988.10430830.
- Rusch, D. B., Halpern, A. L., Sutton, G., Heidelberg, K. B., Williamson, S., Yooseph, S., et al. (2007). The Sorcerer II Global Ocean Sampling Expedition: Northwest Atlantic through Eastern Tropical Pacific. *PLOS Biology* 5, e77. doi:10.1371/journal.pbio.0050077.
- Sachs, J. L., and Hollowell, A. C. (2012). The Origins of Cooperative Bacterial Communities. *mBio* 3, e00099-12. doi:10.1128/mBio.00099-12.
- Sarmiento, H., Morana, C., and Gasol, J. M. (2016). Bacterioplankton niche partitioning in the use of phytoplankton-derived dissolved organic carbon: quantity is more important than quality. *ISME J* 10, 2582–2592. doi:10.1038/ismej.2016.66.
- Sauret, C., Tedetti, M., Guigue, C., Dumas, C., Lami, R., Pujo-Pay, M., et al. (2016). Influence of PAHs among other coastal environmental variables on total and PAH-degrading bacterial communities. *Environ Sci Pollut Res* 23, 4242–4256. doi:10.1007/s11356-015-4768-0.
- Scanlan, D. J., Ostrowski, M., Mazard, S., Dufresne, A., Garczarek, L., Hess, W. R., et al. (2009). Ecological genomics of marine picocyanobacteria. *Microbiol. Mol. Biol. Rev.* 73, 249–299. doi:10.1128/MMBR.00035-08.
- Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., et al. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and environmental microbiology* 75, 7537–7541.
- Schultes, S., Sourisseau, M., Le Masson, E., Lunven, M., and Marié, L. (2013). Influence of physical forcing on mesozooplankton communities at the Ushant tidal front. *Journal of Marine Systems* 109–110, S191–S202. doi:10.1016/j.jmarsys.2011.11.025.

REFERENCES

- Schut, F., Prins, R., and Gottschal, J. (1997). Oligotrophy and pelagic marine bacteria: facts and fiction. *Aquatic Microbial Ecology* 12, 177–202. doi:10.3354/ame012177.
- Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W. S., et al. (2011). Metagenomic biomarker discovery and explanation. *Genome Biol.* 12, R60. doi:10.1186/gb-2011-12-6-r60.
- Semenov, A. M. (1991). Physiological bases of oligotrophy of microorganisms and the concept of microbial community. *Microb Ecol* 22, 239–247. doi:10.1007/BF02540226.
- Seyedsayamdost, M. R., Case, R. J., Kolter, R., and Clardy, J. (2011). The Jekyll-and-Hyde chemistry of *Phaeobacter gallaeciensis*. *Nat Chem* 3, 331–335. doi:10.1038/nchem.1002.
- Sherr, E. B., and Sherr, B. F. (2002). Significance of predation by protists in aquatic microbial food webs. *Antonie Van Leeuwenhoek* 81, 293–308. doi:10.1023/A:1020591307260.
- Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31, 3210–3212. doi:10.1093/bioinformatics/btv351.
- Simó, R., Archer, S. D., Pedrós-Alió, C., Gilpin, L., and Stelfox-Widdicombe, C. E. (2002). Coupled dynamics of dimethylsulfoniopropionate and dimethylsulfide cycling and the microbial food web in surface waters of the North Atlantic. *Limnology and Oceanography* 47, 53–61.
- Simon, M., Scheuner, C., Meier-Kolthoff, J. P., Brinkhoff, T., Wagner-Döbler, I., Ulbrich, M., et al. (2017). Phylogenomics of *Rhodobacteraceae* reveals evolutionary adaptation to marine and non-marine habitats. *The ISME Journal* 11, 1483–1499. doi:10.1038/ismej.2016.198.
- Sinclair, L., Osman, O. A., Bertilsson, S., and Eiler, A. (2015). Microbial Community Composition and Diversity via 16S rRNA Gene Amplicons: Evaluating the Illumina Platform. *PLOS ONE* 10, e0116955. doi:10.1371/journal.pone.0116955.
- Sintes, E., De Corte, D., Ouillon, N., and Herndl, G. J. (2015). Macroecological patterns of archaeal ammonia oxidizers in the Atlantic Ocean. *Molecular Ecology* 24, 4931–4942. doi:10.1111/mec.13365.
- Smith, J. M., Casciotti, K. L., Chavez, F. P., and Francis, C. A. (2014). Differential contributions of archaeal ammonia oxidizer ecotypes to nitrification in coastal surface waters. *The ISME journal* 8, 1704–1714. Available at: <http://www.nature.com/ismej/journal/v8/n8/abs/ismej201411a.html> [Accessed February 18, 2015].
- Smith, V. H., Tilman, G. D., and Nekola, J. C. (1999). Eutrophication: impacts of excess nutrient inputs on freshwater, marine, and terrestrial ecosystems. *Environ. Pollut.* 100, 179–196.

-
- Sogin, M. L., Morrison, H. G., Huber, J. A., Welch, D. M., Huse, S. M., Neal, P. R., et al. (2006). Microbial diversity in the deep sea and the underexplored "rare biosphere." *PNAS* 103, 12115–12120. doi:10.1073/pnas.0605127103.
- Sommer, D. D., Delcher, A. L., Salzberg, S. L., and Pop, M. (2007). Minimus: a fast, lightweight genome assembler. *BMC Bioinformatics* 8, 64. doi:10.1186/1471-2105-8-64.
- Sommer, U., Stibor, H., Katechakis, A., Sommer, F., and Hansen, T. (2002). Pelagic food web configurations at different levels of nutrient richness and their implications for the ratio fish production:primary production. *Hydrobiologia* 484, 11–20. doi:10.1023/A:1021340601986.
- Song, H.-K., Song, W., Kim, M., Tripathi, B. M., Kim, H., Jablonski, P., et al. (2017). Bacterial strategies along nutrient and time gradients, revealed by metagenomic analysis of laboratory microcosms. *FEMS Microbiol Ecol* 93. doi:10.1093/femsec/fix114.
- Sournia, A. (1994). Pelagic biogeography and fronts. *Progress in Oceanography* 34, 109–120. doi:10.1016/0079-6611(94)90004-3.
- Sournia, A., Brylinski, J.-M., Dallot, S., Le Corre, P., Leveau, M., Prieur, L., et al. (1987). Fronts hydrologiques au large des côtes françaises: Les sites-ateliers du programme Frontal. *Oceanologica Acta*.
- Stackebrandt, E., and Goebel, B. M. (1994). Taxonomic Note: A Place for DNA-DNA Reassociation and 16S rRNA Sequence Analysis in the Present Species Definition in Bacteriology. *International Journal of Systematic and Evolutionary Microbiology* 44, 846–849. doi:10.1099/00207713-44-4-846.
- Staley, J. T., and Konopka, A. (1985). Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. *Annu. Rev. Microbiol.* 39, 321–346. doi:10.1146/annurev.mi.39.100185.001541.
- Steele, J. H. (1974). *The Structure of Marine Ecosystems*. 3rd printing 1976. Reprint 2014. Cambridge: Harvard University Press doi:10.4159/harvard.9780674592513.
- Stocker, R. (2012). Marine microbes see a sea of gradients. *Science* 338, 628–633. doi:10.1126/science.1208929.
- Strickland, J. D., and Parsons, T. R. (1972). A practical handbook of seawater analysis.
- Sul, W. J., Oliver, T. A., Ducklow, H. W., Amaral-Zettler, L. A., and Sogin, M. L. (2013). Marine bacteria exhibit a bipolar distribution. *Proc. Natl. Acad. Sci. U.S.A.* 110, 2342–2347. doi:10.1073/pnas.1212424110.

REFERENCES

- Sun, J., Steindler, L., Thrash, J. C., Halsey, K. H., Smith, D. P., Carter, A. E., et al. (2011). One Carbon Metabolism in SAR11 Pelagic Marine Bacteria. *PLoS ONE* 6, e23973. doi:10.1371/journal.pone.0023973.
- Sunagawa, S., Coelho, L. P., Chaffron, S., Kultima, J. R., Labadie, K., Salazar, G., et al. (2015). Structure and function of the global ocean microbiome. *Science* 348, 1261359.
- Suttle, C. A. (2007). Marine viruses--major players in the global ecosystem. *Nat. Rev. Microbiol.* 5, 801–812. doi:10.1038/nrmicro1750.
- Sverdrup, H. U. (1953). On Conditions for the Vernal Blooming of Phytoplankton. *ICES J Mar Sci* 18, 287–295. doi:10.1093/icesjms/18.3.287.
- Swan, B. K., Tupper, B., Sczyrba, A., Lauro, F. M., Martinez-Garcia, M., Gonzalez, J. M., et al. (2013). Prevalent genome streamlining and latitudinal divergence of planktonic bacteria in the surface ocean. *Proceedings of the National Academy of Sciences* 110, 11463–11468. doi:10.1073/pnas.1304246110.
- Tada, Y., Taniguchi, A., Sato-Takabe, Y., and Hamasaki, K. (2012). Growth and succession patterns of major phylogenetic groups of marine bacteria during a mesocosm diatom bloom. *J Oceanogr* 68, 509–519. doi:10.1007/s10872-012-0114-z.
- Taylor, J. D., and Cunliffe, M. (2017). Coastal bacterioplankton community response to diatom-derived polysaccharide microgels. *Environmental Microbiology Reports* 9, 151–157. doi:10.1111/1758-2229.12513.
- Taylor, J. R., and Ferrari, R. (2011). Shutdown of turbulent convection as a new criterion for the onset of spring phytoplankton blooms. *Limnology and Oceanography* 56, 2293–2307. doi:10.4319/lo.2011.56.6.2293.
- Teeling, H., Fuchs, B. M., Becher, D., Klockow, C., Gardebrecht, A., Bennis, C. M., et al. (2012). Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science* 336, 608–611.
- Teeling, H., Fuchs, B. M., Bennis, C. M., Krüger, K., Chafee, M., Kappelmann, L., et al. (2016). Recurring patterns in bacterioplankton dynamics during coastal spring algae blooms. *eLife* 5. doi:10.7554/eLife.11888.
- Teeling, H., Meyerdierks, A., Bauer, M., Amann, R., and Glockner, F. O. (2004). Application of tetranucleotide frequencies for the assignment of genomic fragments. *Environmental Microbiology* 6, 938–947. doi:10.1111/j.1462-2920.2004.00624.x.

-
- Thingstad, T. F. (2000). Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. *Limnology and Oceanography* 45, 1320–1328. doi:10.4319/lo.2000.45.6.1320.
- Thornton, D. C. O. (2014). Dissolved organic matter (DOM) release by phytoplankton in the contemporary and future ocean. *European Journal of Phycology* 49, 20–46. doi:10.1080/09670262.2013.875596.
- Travers, M.-A., Bouffant, R. L., Friedman, C. S., Buzin, F., Cougard, B., Huchette, S., et al. (2009). Pathogenic *Vibrio harveyi*, in contrast to non-pathogenic strains, intervenes with the p38 MAPK pathway to avoid an abalone haemocyte immune response. *Journal of Cellular Biochemistry* 106, 152–160. doi:10.1002/jcb.21990.
- Tréguer, P., Goberville, E., Barrier, N., L’Helguen, S., Morin, P., Bozec, Y., et al. (2014). Large and local-scale influences on physical and chemical characteristics of coastal waters of Western Europe during winter. *Journal of Marine Systems* 139, 79–90. doi:10.1016/j.jmarsys.2014.05.019.
- Treguer, P., and Queguiner, B. (1989). Seasonal-variations in conservative and nonconservative mixing of nitrogen-compounds in a west european macrotidal estuary. *Oceanologica Acta* 12, 371–380. Available at: <https://archimer.ifremer.fr/doc/00106/21734/> [Accessed March 21, 2019].
- Treseder, K. K., Balsler, T. C., Bradford, M. A., Brodie, E. L., Dubinsky, E. A., Eviner, V. T., et al. (2012). Integrating microbial ecology into ecosystem models: challenges and priorities. *Biogeochemistry* 109, 7–18. doi:10.1007/s10533-011-9636-5.
- Tripp, H. J. (2013). The unique metabolism of SAR11 aquatic bacteria. *Journal of Microbiology* 51, 147–153. doi:10.1007/s12275-013-2671-2.
- Tripp, H. J., Kitner, J. B., Schwalbach, M. S., Dacey, J. W. H., Wilhelm, L. J., and Giovannoni, S. J. (2008). SAR11 marine bacteria require exogenous reduced sulphur for growth. *Nature* 452, 741–744. doi:10.1038/nature06776.
- Troade, P., and Le Goff, R. (1997). *Etat des lieux et des milieux de la rade de Brest et de son bassin versant. Phase préliminaire du Contrat de Baie de la rade de Brest*. Brest: Communauté Urbaine de Brest.
- Tully, B. J. (2019). Metabolic diversity within the globally abundant Marine Group II Euryarchaea offers insight into ecological patterns. *Nat Commun* 10. doi:10.1038/s41467-018-07840-4.
- Turley, C., Bianchi, M., Christaki, U., Conan, P., Harris, J., Psarra, S., et al. (2000). Relationship between primary producers and bacteria in an oligotrophic sea-the Mediterranean and biogeochemical implications. *Marine Ecology Progress Series* 193, 11–18. doi:10.3354/meps193011.

REFERENCES

- Turnbaugh, P. J., Ley, R. E., Hamady, M., Fraser-Liggett, C., Knight, R., and Gordon, J. I. (2007). The human microbiome project: exploring the microbial part of ourselves in a changing world. *Nature* 449, 804–810. doi:10.1038/nature06244.
- Turner, J. T. (2015). Zooplankton fecal pellets, marine snow, phytodetritus and the ocean's biological pump. *Progress in Oceanography* 130, 205–248. doi:10.1016/j.pocean.2014.08.005.
- Tyrrell, T. (1999). The relative influences of nitrogen and phosphorus on oceanic primary production. *Nature* 400, 525.
- Unanue, M., Ayo, B., Azúa, I., Barcina, I., and Iriberry, J. (1992). Temporal variability of attached and free-living bacteria in coastal waters. *Microb Ecol* 23, 27–39. doi:10.1007/BF00165905.
- Unfried, F., Becker, S., Robb, C. S., Hehemann, J.-H., Markert, S., Heiden, S. E., et al. (2018). Adaptive mechanisms that provide competitive advantages to marine bacteroidetes during microalgal blooms. *The ISME journal* 12, 2894.
- van Duyl, F. C., Gieskes, W. W. C., Kop, A. J., and Lewis, W. E. (1998). Biological control of short-term variations in the concentration of DMSP and DMS during a *Phaeocystis* spring bloom. *Journal of Sea Research* 40, 221–231. doi:10.1016/S1385-1101(98)00024-0.
- Varela, M. M., van Aken, H. M., Sintès, E., and Herndl, G. J. (2008). Latitudinal trends of Crenarchaeota and Bacteria in the meso- and bathypelagic water masses of the Eastern North Atlantic. *Environ. Microbiol.* 10, 110–124. doi:10.1111/j.1462-2920.2007.01437.x.
- Vaulot, D., Marie, D., Olson, R. J., and Chisholm, S. W. (1995). Growth of *Prochlorococcus*, a Photosynthetic Prokaryote, in the Equatorial Pacific Ocean. *Science* 268, 1480–1482. doi:10.1126/science.268.5216.1480.
- Venter, J. C., Remington, K., Heidelberg, J. F., Halpern, A. L., Rusch, D., Eisen, J. A., et al. (2004). Environmental Genome Shotgun Sequencing of the Sargasso Sea. *Science* 304, 66–74. doi:10.1126/science.1093857.
- Vergin, K., Done, B., Carlson, C., and Giovannoni, S. (2013). Spatiotemporal distributions of rare bacterioplankton populations indicate adaptive strategies in the oligotrophic ocean. *Aquatic Microbial Ecology* 71, 1–13. doi:10.3354/ame01661.
- Vernadsky, W. I. (1945). THE BIOSPHERE AND THE NOÖSPHERE. *American Scientist* 33, xxii–12. Available at: <https://www.jstor.org/stable/27826043> [Accessed March 20, 2019].
- Vezzulli, L., Brettar, I., Pezzati, E., Reid, P. C., Colwell, R. R., Höfle, M. G., et al. (2012). Long-term effects of ocean warming on the prokaryotic community: evidence from the vibrios. *ISME J* 6, 21–30. doi:10.1038/ismej.2011.89.

-
- Vezzulli, L., Colwell, R. R., and Pruzzo, C. (2013). Ocean warming and spread of pathogenic vibrios in the aquatic environment. *Microb. Ecol.* 65, 817–825. doi:10.1007/s00248-012-0163-2.
- Videau, C. (1987). Primary production and physiological state of phytoplankton at the Ushant tidal front (west coast of Brittany, France). *Marine Ecology Progress Series* 35, 141–151. doi:10.3354/meps035141.
- Voget, S., Wemheuer, B., Brinkhoff, T., Vollmers, J., Dietrich, S., Giebel, H.-A., et al. (2015). Adaptation of an abundant Roseobacter RCA organism to pelagic systems revealed by genomic and transcriptomic analyses. *ISME J* 9, 371–384. doi:10.1038/ismej.2014.134.
- Wan, X. S., Sheng, H.-X., Dai, M., Zhang, Y., Shi, D., Trull, T. W., et al. (2018). Ambient nitrate switches the ammonium consumption pathway in the euphotic ocean. *Nature Communications* 9, 915. doi:10.1038/s41467-018-03363-0.
- Ward, C. S., Yung, C.-M., Davis, K. M., Blinebry, S. K., Williams, T. C., Johnson, Z. I., et al. (2017). Annual community patterns are driven by seasonal switching between closely related marine bacteria. *The ISME journal* 11, 1412.
- Waycott, M., Duarte, C. M., Carruthers, T. J. B., Orth, R. J., Dennison, W. C., Olyarnik, S., et al. (2009). Accelerating loss of seagrasses across the globe threatens coastal ecosystems. *PNAS* 106, 12377–12381. doi:10.1073/pnas.0905620106.
- Weisberg, R. H., Zheng, L., and Liu, Y. (2015). "Chapter 4 - Basic Tenets for Coastal Ocean Ecosystems Monitoring," in *Coastal Ocean Observing Systems*, eds. Y. Liu, H. Kerkerling, and R. H. Weisberg (Boston: Academic Press), 40–57. doi:10.1016/B978-0-12-802022-7.00004-3.
- Whitman, W. B., Coleman, D. C., and Wiebe, W. J. (1998). Prokaryotes: The unseen majority. *Proceedings of the National Academy of Sciences* 95, 6578–6583. doi:10.1073/pnas.95.12.6578.
- Williams, R. J., and Martinez, N. D. (2000). Simple rules yield complex food webs. *Nature* 404, 180–183. doi:10.1038/35004572.
- Wiltshire, K. H., Kraberg, A., Bartsch, I., Boersma, M., Franke, H.-D., Freund, J., et al. (2010). Helgoland Roads, North Sea: 45 Years of Change. *Estuaries and Coasts* 33, 295–310. doi:10.1007/s12237-009-9228-y.
- Woese, C. R., and Fox, G. E. (1977). Phylogenetic structure of the prokaryotic domain: The primary kingdoms. *PNAS* 74, 5088–5090. doi:10.1073/pnas.74.11.5088.
- Woese, C. R., Kandler, O., and Wheelis, M. L. (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *PNAS* 87, 4576–4579. doi:10.1073/pnas.87.12.4576.

REFERENCES

- Worden, A. Z., Follows, M. J., Giovannoni, S. J., Wilken, S., Zimmerman, A. E., and Keeling, P. J. (2015). Rethinking the marine carbon cycle: Factoring in the multifarious lifestyles of microbes. *Science* 347, 1257594–1257594. doi:10.1126/science.1257594.
- Wu, Q. L., Zwart, G., Schauer, M., Agterveld, M. P. K., and Hahn, M. W. (2006). Bacterioplankton Community Composition along a Salinity Gradient of Sixteen High-Mountain Lakes Located on the Tibetan Plateau, China. *Appl. Environ. Microbiol.* 72, 5478–5485. doi:10.1128/AEM.00767-06.
- Wu, Y.-W., Simmons, B. A., and Singer, S. W. (2016). MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* 32, 605–607. doi:10.1093/bioinformatics/btv638.
- Wuchter, C., Abbas, B., Coolen, M. J., Herfort, L., van Bleijswijk, J., Timmers, P., et al. (2006). Archaeal nitrification in the ocean. *Proceedings of the National Academy of Sciences* 103, 12317–12322.
- Yentsch, C. M., Mague, F. C., Horan, P. K., and Muirhead, K. (1983). Flow cytometric DNA determinations on individual cells of the dinoflagellate *Gonyaulax tamarensis* var. *excavata*. *Journal of Experimental Marine Biology and Ecology* 67, 175–183. doi:10.1016/0022-0981(83)90088-6.
- Yilmaz, P., Yarza, P., Rapp, J. Z., and Glöckner, F. O. (2016). Expanding the World of Marine Bacterial and Archaeal Clades. *Front. Microbiol.* 6. doi:10.3389/fmicb.2015.01524.
- Yoch, D. C. (2002). Dimethylsulfoniopropionate: Its Sources, Role in the Marine Food Web, and Biological Degradation to Dimethylsulfide. *Applied and Environmental Microbiology* 68, 5804–5815. doi:10.1128/AEM.68.12.5804-5815.2002.
- Yooseph, S., Nealson, K. H., Rusch, D. B., McCrow, J. P., Dupont, C. L., Kim, M., et al. (2010). Genomic and functional adaptation in surface ocean planktonic prokaryotes. *Nature* 468, 60–66. doi:10.1038/nature09530.
- Zakem, E. J., Al-Haj, A., Church, M. J., van Dijken, G. L., Dutkiewicz, S., Foster, S. Q., et al. (2018). Ecological control of nitrite in the upper ocean. *Nature Communications* 9. doi:10.1038/s41467-018-03553-w.
- Zehr, J. P., and Ward, B. B. (2002). Nitrogen cycling in the ocean: new perspectives on processes and paradigms. *Appl. Environ. Microbiol.* 68, 1015–1024.
- Zhang, C. L., Xie, W., Martin-Cuadrado, A.-B., and Rodriguez-Valera, F. (2015a). Marine Group II Archaea, potentially important players in the global ocean carbon cycle. *Frontiers in Microbiology* 6. doi:10.3389/fmicb.2015.01108.

-
- Zhang, G., Liang, S., Shi, X., and Han, X. (2015b). Dissolved organic nitrogen bioavailability indicated by amino acids during a diatom to dinoflagellate bloom succession in the Changjiang River estuary and its adjacent shelf. *Marine Chemistry* 176, 83–95. doi:10.1016/j.marchem.2015.08.001.
- Zhang, Y., Jiao, N., Cottrell, M. T., and Kirchman, D. L. (2006). Contribution of major bacterial groups to bacterial biomass production along a salinity gradient in the South China Sea. *Aquatic Microbial Ecology* 43, 233–241. doi:10.3354/ame043233.
- Zhang, Y., Sun, Y., Jiao, N., Stepanauskas, R., and Luo, H. (2016). Ecological Genomics of the Uncultivated Marine Roseobacter Lineage CHAB-I-5. *Applied and Environmental Microbiology* 82, 2100–2111. doi:10.1128/AEM.03678-15.
- Zhou, J., Richlen, M. L., Sehein, T. R., Kulis, D. M., Anderson, D. M., and Cai, Z. (2018). Microbial Community Structure and Associations During a Marine Dinoflagellate Bloom. *Front Microbiol* 9. doi:10.3389/fmicb.2018.01201.
- ZoBell, C. E. (1946). *Marine microbiology, a monograph on hydrobacteriology*. Waltham, Mass.,: Chronica Botanica Company, Available at: <https://www.biodiversitylibrary.org/bibliography/10079>.

Titre : Effet des changements environnementaux sur la dynamique des communautés microbiennes : évolution et adaptation des bactéries dans le cadre du développement de l'Observatoire Microbien en Rade de Brest et Mer d'Iroise

Mots clés : écologie microbienne, suivi spatio-temporel, métagénomique, observatoire marin

Résumé : Les bactéries marines ont un rôle primordial au cœur du fonctionnement des écosystèmes marins. Dans ces environnements elles forment des communautés complexes extrêmement dynamiques dans le temps et dans l'espace dont la diversité globale reste encore peu connue. Comprendre les mécanismes qui gouvernent cette dynamique est essentiel pour mieux appréhender leur rôle dans les écosystèmes et anticiper l'influence de changements globaux et des pressions anthropiques. Cette thèse présente la mise en place d'un Observatoire microbien en Rade de Brest et Mer d'Iroise. La dynamique de ces communautés en lien avec l'environnement a dans un premier temps été étudiée par séquençage d'amplicons 16S. La présence d'un front saisonnier en mer d'Iroise structure profondément les communautés bactériennes qui vont varier avec les différentes masses d'eau. Mais un des paramètres les plus structurant de ces communautés seraient la distribution hétérogène du phytoplancton qui sélectionne des bactéries hétérotrophes adaptées à différentes concentrations en matière organique.

En Rade de Brest, l'analyse des 3 premières années d'un suivi temporel sur la station SOMLIT montre que les communautés bactériennes présentent une saisonnalité remarquable. Cette dynamique saisonnière présente un changement rapide lors de la mise en place des blooms printaniers avec la réponse rapide et intense de plusieurs taxa hétérotrophes. Enfin, une approche métagénomique nous a permis de reconstruire le génome (MAG) de la bactérie la plus abondante retrouvée sur ces deux suivis, appartenant à la famille des *Rhodobacteraceae*. L'analyse de son génome montre des caractéristiques inhabituelles pour cette famille, avec des traits typiques des « streamlined genomes ». Cependant, contrairement à ce qui est le plus généralement décrit pour les bactéries de ce type, elle est adaptée à un environnement riche en matière organique, représentant jusqu'à 20% de la communauté microbienne en période de blooms. Ces résultats soulignent l'importance de poursuivre l'exploration des communautés tant d'un point de vue de leur dynamique globale que de leurs fonctions.

Title : Effect of environmental variations on bacterial communities dynamics : evolution and adaptation of bacteria as part of a microbial observatory in the Bay of Brest and the Iroise Sea.

Keywords : microbial ecology, dynamic, spatio-temporal survey, metagenomic, marine observatory

Abstract : Marine bacteria play a key role in the functioning of marine ecosystems. In these environments they form complex communities that are extremely dynamic in time and space, whose global diversity is still largely unknown. Understanding the mechanisms that govern their dynamic is essential to better understand their role in ecosystems and anticipate the influence of global changes and anthropogenic pressures. This thesis presents the set-up of a microbial observatory in the Bay of Brest and the Iroise Sea. Bacterial community dynamics in relation to the environment were first studied using 16S amplicon sequencing. The presence of a seasonal front in the Iroise Sea deeply structures bacterial communities that will vary with the different water masses. But one of the most structuring parameters is the heterogeneous distribution of phytoplankton, which would select heterotrophic bacteria adapted to different organic matter concentrations

In the Bay of Brest, analysis of the first 3 years of temporal monitoring at the SOMLIT station shows that bacterial communities present remarkable recurrent seasonality. This seasonal dynamic presents a rapid change when spring blooms are implemented with the rapid and intense response of several heterotrophic taxa. Finally, a metagenomic approach has allowed us to reconstruct the genome (MAG) of the most abundant bacterium found on these two surveys, belonging to the *Rhodobacteraceae* family. Its genome shows unusual characteristics for this family, with features typical of streamlined genomes. However, contrary to what is generally described for such bacteria, it is adapted to an environment rich in organic matter, representing up to 20% of the microbial community during bloom periods. These results highlight the importance to explore communities both in terms of their overall dynamics and functionality.